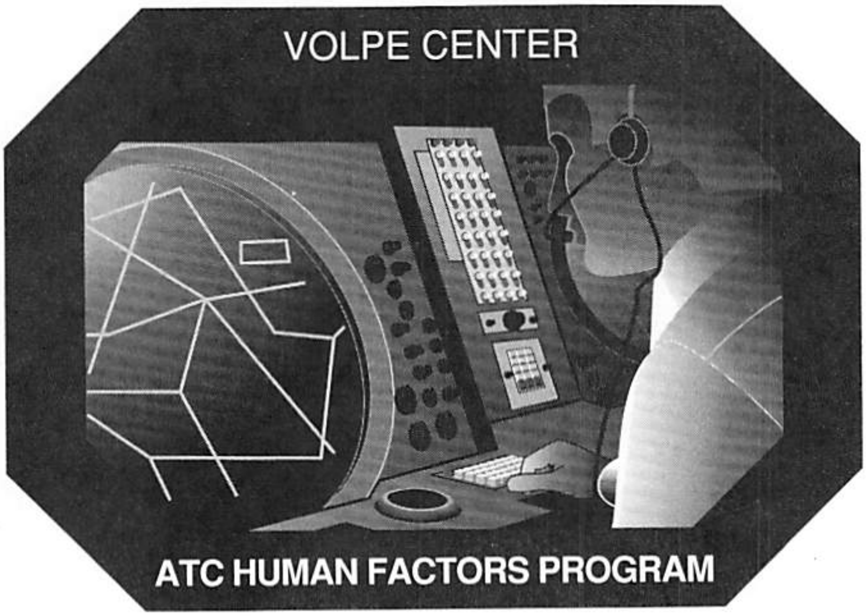
Chief Scientific and Technical Advisor for Human Factors
AAR-100

Human Factors in the Design and Evaluation of Air Traffic Control Systems

DOT/FAA/RD-95/3
DOT-VNTSC-FAA-95-3

Office of Aviation Research
Washington, DC 20591



U.S. Department of Transportation
Research and Special Programs Administration
John A. Volpe National Transportation Systems Center
Cambridge, MA 02142-1093

Edited by Kim M. Cardosi and Elizabeth D. Murphy

Final Report
April 1995

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
April 1995

3. REPORT TYPE AND DATES COVERED
Final Report

4. TITLE AND SUBTITLE
Human Factors in the Design and Evaluation of Air Traffic Control Systems

5. FUNDING NUMBERS
FA5L1/A5112

6. AUTHOR(S)
Kim M. Cardosi, Elizabeth D. Murphy* (Eds.)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
Research and Special Programs Administration
John A. Volpe National Transportation Systems Administration
Cambridge, MA 02142

8. PERFORMING ORGANIZATION REPORT NUMBER
DOT-VNTSC-FAA-95-3

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
Federal Aviation Administration
Research and Development Service
800 Independence Avenue, S.W.
Washington, DC 20591

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
DOT/FAA/RD-95/3

11. SUPPLEMENTARY NOTES
*CTA Inc.
Suite 800
6116 Executive Blvd.
Rockville, MD 20852

12a. DISTRIBUTION/AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This document presents human factors issues that should be considered in the design and evaluation of air traffic control (ATC) systems and subsystems. It provides background material on the capabilities and limitations of humans as information processors and discusses issues in: ATC automation, computer-human interface, workstation design, workload and performance measurement, controller team formation and activities, and human factors testing and evaluation. The goal of this material is to help air traffic controllers and other operations specialists identify potential problems by alerting them to known design flaws and providing them with information as to why some design options may be undesirable or operationally unsuitable. This document presents design goals based on human factors principles, standards, and guidelines. Some of these design goals are idealistic in an ATC operational setting. They are presented so that the operations specialists can identify key human factors issues and understand the implications of compromises, and where they must be made.

14. SUBJECT TERMS
human factors in air traffic control

15. NUMBER OF PAGES
760

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	= 2.5 centimeters (cm)
1 foot (ft)	= 30 centimeters (cm)
1 yard (yd)	= 0.9 meter (m)
1 mile (mi)	= 1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	= 6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	= 0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	= 0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	= 2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	= 4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	= 28 grams (gm)
1 pound (lb)	= 0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	= 0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	= 5 milliliters (ml)
1 tablespoon (tbsp)	= 15 milliliters (ml)
1 fluid ounce (fl oz)	= 30 milliliters (ml)
1 cup (c)	= 0.24 liter (l)
1 pint (pt)	= 0.47 liter (l)
1 quart (qt)	= 0.96 liter (l)
1 gallon (gal)	= 3.8 liters (l)
1 cubic foot (cu ft, ft^3)	= 0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	= 0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	= 0.04 inch (in)
1 centimeter (cm)	= 0.4 inch (in)
1 meter (m)	= 3.3 feet (ft)
1 meter (m)	= 1.1 yards (yd)
1 kilometer (k)	= 0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	= 0.16 square inch (sq in, in^2)
1 square meter (m^2)	= 1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	= 0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	= 1 hectare (he) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	= 0.036 ounce (oz)
1 kilogram (kg)	= 2.2 pounds (lb)
1 tonne (t) = 1,000 kilograms (kg)	= 1.1 short tons

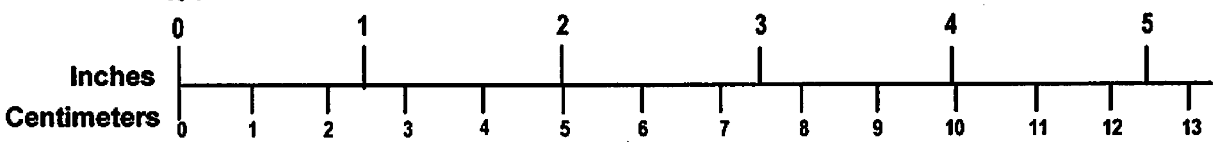
VOLUME (APPROXIMATE)

1 milliliter (ml)	= 0.03 fluid ounce (fl oz)
1 liter (l)	= 2.1 pints (pt)
1 liter (l)	= 1.06 quarts (qt)
1 liter (l)	= 0.26 gallon (gal)
1 cubic meter (m^3)	= 36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	= 1.3 cubic yards (cu yd, yd^3)

TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSION

°F	-40°	-22°	-4°	14°	32°	50°	68°	86°	104°	122°	140°	158°	176°	194°	212°
°C	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°

For more exact and or other conversion factors, see NBS Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 1/23/95

PREFACE

Air traffic control (ATC) specialists are often responsible for shaping the development and evaluation of new air traffic systems. This important task requires making many decisions about the design and operation of displays, controls, and supporting software functions. This handbook and accompanying checklist have been designed to help operations specialists address these human factors issues. This handbook provides background material on the role of human factors in the acquisition process, the capabilities and limitations of humans as information processors and the evaluation of displays and controls. It also includes discussions of issues of particular interest to air traffic control, such as the benefits and limitations of automation, and methods of workload assessment. Application of the information presented in this handbook will help to minimize the probability of human error in human-system interactions and increase the efficiency of human-system performance. These are two paramount goals of the FAA's *National Plan for Civil Aviation Human Factors: An Initiative for Research and Application* (March, 1995). This material is provided solely for guidance and is intended to be used by ATC specialists as they see fit.

This work was sponsored by the Federal Aviation Administration's Office of the Chief Scientific and Technical Advisor for Human Factors (AAR-100). We are grateful to Lawrence Cole, Glen Hewitt, Thomas McCloy, William White, and John Zalenchak for their support and technical advice.

During the first several months of this project, we conducted a survey to identify topics that should be included in this document. We are extremely grateful to the people who so graciously gave us the benefit of their time, thought, and experience.

Interviews were held in the following locations with the personnel listed, to whom we wish to express our appreciation:

Data Distribution Facility, Gaithersburg, MD:
Larry Roberts, NAS Implementation (AMNI) Seattle ARTCC

FAA Headquarters (ATR), Washington, DC:
Terry Bass, Terry Brown, and Terry Schomberg

Oklahoma City, OK:
FAA Academy: Claude Schuldt

Civil Aeromedical Institute (CAMI): Carol Manning, Mark Rodgers, Henry Mertens, Mark Touchstone

Denver ARTCC, Longmont, CO:
Mo Hart, Center TRACON Tower Automation System (CTAS)/ CTAS System Development Team

Seattle ARTCC, Auburn, WA:
Dave Taylor, Denise Harrell, John Warner, Peter Maunsell, Carl Jensen, Glen Wood, Tom Rieger, Mike Marler, Larry Roberts

Seattle-Tacoma (SEATAC) Tower:
Art Vail

We are especially grateful to Fred Heistuman (Seattle ARTCC, NAS Implementation Office) for his gracious support during our visit to the Center.

We also would like to thank the following members of the VSCS System Requirements Team for their very helpful feedback and their responses to our questionnaires: Jeff Call (ATR-320), Hoyt Diamond (Indianapolis ARTCC), Joel Hicks (ATR-320), Terry Jackson (Kansas City ARTCC), Richard Jensen (Seattle ARTCC), Edward Kunz (DSM), Chris McMahon (Boston ARTCC), Gary Nigro (AEA-512), Robert Potter (Denver ARTCC), and Charles Ullmann (SATCS).

Additional questionnaires were collected from the following individuals, to whom we wish to express our gratitude: Robert Greene, Paul Krois, Gary Powell, Wayne Tobey, and James Yohman (CTA), and Rod Hurlbert (Denver ARTCC).

A group discussion of human factors issues associated with system design and development was held with the following CTA human factors personnel, who support the VSCS System Requirements Team (SRT): Valerie Dykstra, James Epley, Mark Guidi, and Brian Legan. We are grateful for their enthusiastic participation and thoughtful comments.

During the months of preparing the text and checklist, we relied on reviews by Renate Roske-Hofstrand (NYMA), James Yohman (CTA), Robert Humbertson (Booz·Allen & Hamilton), Kelly Harwood (CTA), Rust Potter (FAA Technical Center), Leslie Carter (CTA), and Paula Van Balen (Harris Corporation). Discussions with these reviewers were extremely helpful and productive.

We thank Stephanie Levy (EG&G Dynatrend) for her painstaking construction of the document and endless hours of talented editorial and publication support. We also thank Sherwin Han (EG&G Dynatrend) for his review of the checklist.

We are also grateful to those who contributed photographs to the text: Paula Nouragas (FAA Technical Center), Ed Peters (Measurement Systems, Inc.), Pat Tomasetti (FAA Public Affairs), Tirey Vickers (Air Traffic Control Association), and Marilyn Welles (MITRE).

The writers wish to express their sincere thanks to Robert Hawley (Booz·Allen & Hamilton, Program Manager) and Jordan Multer (Volpe Center, Technical Project Monitor) for their managerial support and to John Crosby (Booz·Allen & Hamilton, Graphic Designer/Editor) who cheerfully kept track of important details and helped ensure that things ran smoothly.

Finally, even the most complex and efficient machine is only as good as its power source. Cynthia Smith Karon (Booz·Allen & Hamilton) was our power source. Her remarkable technical, organizational, and editorial talents combined with her dedication and attention to detail contributed significantly to the quality of this document. Her joyful energy and enthusiasm were contagious and remained undaunted, even through the most trying of circumstances. We are indebted to Cyndy in more ways than she will ever know.

ABOUT THE AUTHORS

Kim M. Cardosi received a Ph.D. in Experimental Psychology from Brown University in 1985 where her area of research was visual perception. She received a private pilot certificate in 1989. In 1988, Dr. Cardosi received the Research and Special Programs' Administrator's Award for Superior Achievement for her work in cockpit and cabin crew coordination. In 1993 she received the same award for work in human factors in air traffic control. Dr. Cardosi has conducted numerous studies on pilot-controller voice communications and simulation studies of prototype equipment for air traffic control. She is currently the manager of the air traffic control human factors program at the Volpe National Transportation Systems Center in Cambridge, Massachusetts.

Elizabeth D. Murphy received an M.A. in Industrial/Organizational Psychology from George Mason University in 1983 and is currently pursuing a Ph.D. in Cognitive and Experimental Psychology at the University of Maryland. Ms. Murphy has over 11 years of experience in human factors analysis and design of automated systems and environments for the Federal Aviation Administration (FAA) and the National Aeronautics and Space Administration (NASA). Her areas of expertise include cognitive psychology, air traffic control human factors, cognitive modeling, workload and performance, and computer-human interaction. She has co-authored more than 30 scientific papers and is currently a senior human factors specialist with CTA INCORPORATED in Rockville, Maryland.

Glen Hewitt is currently serving as a Scientific and Technical Advisor for Human Factors in FAA Research and Acquisitions. Previously, he conducted human factors support for system operational test and evaluation as a Senior Principal Scientist with the Atlantic Research Corporation. Prior to that, he served with the Department of Defense where he conducted manpower and force planning, modelling, and analyses, and was instrumental in the development and implementation of Manpower and Personnel Integration (MANPRINT) and Human System Integration (HSI) programs. In 1987, he fulfilled a Fellowship program in Human Factors and Operations Research with the RAND Corporation. He is the author of a number of articles, handbooks, and guides on human factors applications and programs. He is a graduate of both the Army's Command and General Staff College (Planning and Analysis) and the Navy's Naval Command College (Policy and Economic Decision Making). He is a Certified System Professional from the International Certified Computer Professionals. He holds a BS in Engineering from the United States Military Academy and an MS in Systems Management and Safety from the University of Southern California.

John S. Werner received his Ph.D. in Experimental Psychology from Brown University in 1979. He has conducted postdoctoral research at the Institute for Human Factors in Soesterberg, The Netherlands; the Department of Neurophysiology at the University of Freiberg, Germany; and the Institute of Ophthalmology at the University of London. Dr Werner is a Fellow of the American Psychological Association and the Optical Society of America. He has received awards from the American Academy of Optometry, the American Psychological Association, and the Alexander von Humboldt Foundation in Bonn, Germany. Dr Werner's primary research interests are the physiological mechanisms of human vision, particularly color vision, and their changes correlated with age (from birth to senescence) and exposure to light. He has co-authored several books and more than 80 scientific papers. He is currently a Professor of Psychology at the University of Colorado, Boulder.

Daniel J. Hannon received a Ph.D. in Experimental Psychology from Brown University in 1991 where he studied visual perception, with a particular emphasis on the perception of color. He conducted postdoctoral research on age-related declines in visual perception and cognition at Syracuse University. Dr. Hannon began working in cockpit and air traffic control human factors in 1993 as a research scientist for Battelle Memorial Institute and later as an engineering psychologist with the Volpe National Transportation Systems Center in Cambridge, Massachusetts. His current research pursuits include the development of cockpit displays for the depiction of terrain and aeronautical charts, and the use of color on air traffic control displays.

Donna Anastasi received an MS in Industrial and Systems Engineering from the Georgia Institute of Technology in 1986. Her area of expertise is user-system interface design and evaluation. She was the primary author of the Style Guide for the User-System Interface of the Air Force Mission Support System. Ms. Anastasi is currently a group engineer at CTA INCORPORATED in Bedford, Massachusetts and serves as the team leader for an on-site evaluation of a user-system interface at Lockheed Saunders.

Brian Hill received an MA in Human Factors and Applied Experimental Psychology from the University of Dayton in 1990. His areas of specialization include requirements analysis, operational testing and evaluation, site implementation, and training. He has worked on evaluations of the Initial Sector Suite (ISS) and Tower Control Complex (TCC). Mr. Hill is currently at CTA INCORPORATED in McLean, Virginia where he is a senior engineer serving as the technical manager of the team supporting the FAA's Operational Engineering Group for the Voice Switching and Control System (VSCS).

Jerry Guttman received an MA in Industrial psychology from Western Kentucky University in 1985. His areas of specialization include measurement of air traffic controller performance, computer-human interface and display design. He has co-authored a number of technical reports, primarily on airspace complexity and its effect on controller performance, and on laboratory evaluations of air traffic controller performance. While contributing to this document Mr. Guttman was at CTA INCORPORATED; he is currently a Technical Program Manager with Princeton Economics Research Incorporated (PERI).

Paola Amaldi received an MS in Cognitive Engineering from the University of Pittsburgh. She helped to develop controller performance and workload measures in support of the FAA's Multiple Parallel Approach Program. Ms. Amaldi is currently participating in a work-study program abroad while pursuing a Ph.D.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Air traffic control (ATC) specialists are often responsible for shaping the development and evaluation of new air traffic systems. This important task requires making many decisions about the design and operation of displays, controls, and supporting software functions. This document presents human factors issues that should be considered in the design and evaluation of air traffic control (ATC) systems and subsystems. It provides background material on the role of human factors in the acquisition process, the capabilities and limitations of humans as information processors and the evaluation of displays and controls. It also includes discussions of issues of particular interest to air traffic control, such as the benefits and limitations of automation, and methods of workload assessment. Application of the information presented in this handbook will help to minimize the probability of human error in human-system interactions, limit the consequences of these errors, and increase the efficiency of human-system performance. These are paramount goals of the FAA's *National Plan for Civil Aviation Human Factors: An Initiative for Research and Application* (March, 1995).

The introduction to this document, presented in Chapter 1, provides an overview of the scope of this document, how it was constructed, and how it should be used. Chapter 2 discusses the role of the science of human factors in the system acquisition process from the specification of requirements to the evaluation of prototype systems. The chapter concludes with information on how to develop a human factors plan.

Chapters 3 and 4 provide an introduction to visual and auditory perception. Since all of the critical information that controllers take in comes either through the eyes or the ears, it is important to understand these processes. The capabilities and limitations of human vision and how these abilities change with age are discussed in detail in Chapter 3. Topics include: acuity, form vision, flicker, depth perception, and color perception. Chapter 4 discusses the capabilities and limitations of the auditory system with a particular focus on speech perception and controller-pilot voice communications.

How we process the information that is taken in through our senses is discussed in Chapter 5. Because the controller's job centers on processing information, planning, and making decisions under time pressure, it is important that systems support and complement these critical processes. Topics discussed in Chapter 5 include: the time required for information processing activities, attention, memory and forgetting, problem solving, and decision making.

How the increased use of ATC automation can be expected to change the controllers' tasks is discussed in Chapter 6, *Issues in ATC Automation*. The purpose of this chapter is to assist the ATC specialists in their specification of requirements in terms of *which* tasks should be automated and *how* these functions should be automated. The differences between a "technology-centered" and a "user-centered" approach to automation is discussed as well as the potential benefits and drawbacks of automation.

A overview of issues in computer-human interface (CHI) is presented in Chapter 7. This chapter focusses on visual and auditory displays, and the devices and methods of data entry. It includes information on symbology, the use of color, flicker, visual and auditory alerts, keyboards, touchscreens, trackballs and other input devices, menus, formats for data-entry, and error messages.

How workload is defined and measured is critical in an evaluation. Chapter 8 discusses workload and performance measurement in the ATC environment. It surveys the methods currently used to measure workload and discusses the advantages and limitations of each of these methods. This information should be used not only to select the workload measures used in an evaluation, but also to interpret the results of workload evaluations.

An ATC workstation includes all the items located within the controller's work space: the main display and control console, auxiliary displays, communications equipment, work surfaces, seating, and storage. If designed properly, ATC workstations and facility environments can promote the controller's safety, health, job performance, and job satisfaction. Proper design results from an understanding of the operational realities and the application of basic human factors principles. Chapter 9 examines the key human factors issues in workstation and facility design.

Chapter 10 discusses human factors testing and evaluation. It provides information on: how human performance is measured, the different types of tests available (e.g., questionnaires, laboratory experiments, simulation studies), how to determine which is the most appropriate type of test, and how test results should be analyzed and interpreted.

Because each chapter is meant to be able to stand alone and not require knowledge of previous chapters, a thorough reader will notice some redundancy. Where a specific topic is covered in more detail in another chapter, it is noted.

Most chapters in this document conclude with a list of checklist items. These items are compiled in the accompanying document, *Human Factors Checklist for the Design and Evaluation of ATC Systems*. The goal of these checklist items is to point air traffic controllers and other operations specialists to questions that they may wish to consider in their evaluation of new systems or subsystems, or a new component of an existing system. The numbers in parentheses at the end of each checklist item refer to the section of the handbook that discusses the issue. This mapping allows the checklist user to learn about the basis for the item, why it is important, and the implications of compromise. This material is provided solely for guidance and is intended to be used by ATC specialists as they see fit.

CHAPTER 1. INTRODUCTION

Elizabeth D. Murphy

> "At the heart of human factors is the human in the air traffic control system and the interactions between human and system. The human affects the system in terms of safety, efficiency of performance, capacity, and adaptability to unusual circumstances. The system affects the human in terms of roles and functions, job satisfaction, health, and morale."
>
> Hopkin, 1988

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1. INTRODUCTION

When new air traffic control (ATC) equipment is under consideration, the Federal Aviation Administration (FAA) charters at least one team of operational experts to advise system designers and developers. Typically, members of these user teams are Air Traffic Control Specialists (ATCSs) from field facilities and from the Air Traffic Plans and Requirements Service (ATR) at FAA Headquarters.[1] These user teams are the primary audience for the material presented in this handbook and evaluation checklist.

1.1 PURPOSE

Each team chartered by the FAA has an important job to do. The team's job is to provide operational input to the design and evaluation of a new ATC system. **This handbook and checklist are designed to assist these teams in accomplishing their objectives by pointing them to human factors questions that should be asked during system design and evaluation.** These materials have been developed to provide guidance only; there is no requirement to comply with this guidance.

1.1.1 How should this handbook and checklist be used?

Controllers are operations specialists or "subject-matter experts" (SMEs) in ATC. As representatives of the front-line "operators" or "end users" of automated systems installed in ATC facilities, controllers are well qualified to provide expert input in two key areas:

- Identification of operational requirements for new systems.
- Evaluation of design products for operational suitability.

Design products include, but are not limited to the following: concepts documented in reports, paper mockups, prototypes, simulations, human-computer dialogue designs, workstations, tasks and procedures, lighting plans, and furniture. If a particular design product is found to be operationally unsuitable by the team, it must be re-designed and re-evaluated.

Modifications to a design after it has been tested and evaluated is costly and time consuming. **The earlier in the development process that human factors issues are identified, the easier and less costly they will be to address.** Identifying potential problems areas early can also help to focus the operational testing and evaluation and make these tests more effective at identifying potential operational limitations.

"One of the lessons we learned, early on in our modernization program, was that ignoring human factors in our major acquisitions can cost us dearly, both in the expense of re-engineering and in schedule delays. We've made it a requirement that human factors must be systematically integrated at each critical step in the design, testing, and acquisition of any new technology introduced into the air traffic control system."

Del Balzo, 1993, p. 3

This handbook can help operations specialists identify potential problem areas early by alerting them to known design flaws and by providing them with information to why some design options may be undesirable or operationally unsuitable. This text presents human factors principles, standards, and guidelines. Some of the guidelines may seem idealistic. They are not necessarily meant to serve as system requirements; requirements must be determined by operations specialists who are familiar with the proposed system and the environment in which it will be used. Rather, **this handbook and checklist present design goals based on current human factors knowledge. These ideals do not preclude compromise. However, where compromises must be made, the implications of these compromises must be clearly understood.** Each checklist item is a (necessarily succinct) statement of a design goal. At the end of each checklist item is a number in parentheses. This number refers to the section of the text that discusses this issue. The supporting text will help the team to understand why the issue is important and to anticipate the possible effects of compromise.

The objective of this handbook is not to transform controllers into human factors engineers, but to raise their level of awareness of human factors concepts, terminology, and methods. Ideally, the operations specialists should be supported by human factors experts, who work with them to identify operational requirements, plan and conduct evaluations, and document the team's findings. Designers and human factors engineers working with controllers need to become as familiar as possible with the ATC domain. They need to understand controllers' goals, control tasks and procedures, constraints on ATC, and the aviation environment in which controllers operate. The first step is to talk to controllers and elicit their insights into the strengths and weaknesses of the current and proposed systems. At all times, designers and human factors specialists must try to put themselves in the controller's shoes. They need to consider all the tasks that the controllers must perform at the same time. Above all, designers must build systems that serve the controller, rather than systems that simply add to controller workload.

1.2 DEFINITIONS

Key terms that need to be defined include **human factors engineering, systems,** and **systems approach.** Other terms are defined in later chapters.

1.2.1 What is human factors engineering, and why is it important in the design of ATC systems?

> "Human factors discovers and applies information about human abilities, limitations, and other characteristics to the design of tools, machines, systems, tasks, jobs, and environments for safe, comfortable, and effective human use."
>
> Chapanis, 1983, p. 12

Human factors engineering is the discipline that applies knowledge of human capabilities and limitations to the design of technological systems. Human factors engineering is also known as ergonomics.[2] Ergonomics is concerned with designing products for people to use and often focuses on design of tools, equipment, and systems for use in work environments. You may have seen or heard the term, ergonomics, in connection with the design of automobile interiors and office furniture. Ergonomics also plays an important role in the design of systems for air traffic control.

The primary human factors concern is that design products be compatible with human capabilities and limitations. Such compatibility is critical for ease of training, ease of use, and ease of maintaining design products for use in ATC environments. Compatibility may also be critical for safety and efficiency. Incompatibility in these areas makes a design product operationally unsuitable.

In defining human factors as a discipline, it may be useful to say what human factors is not. Human factors is not a medical discipline, although it is concerned with the effects of design on people's health and well being; it is not stress management, although it is concerned with identifying and reducing sources of high stress in work environments; it is not genetic engineering. (Notice that "human factors" is used with a singular verb because the phrase is the name of the discipline, just like "physics" or "economics" is the name of a discipline.) Human factors is a close cousin of engineering psychology.

> "Machines exist only for one purpose, and that purpose is to serve people."
>
> Chapanis, 1983, p. 14

Human factors is a discipline that combines features of many disciplines including psychology, engineering, anthropology, sociology, and linguistics. In basic terms, human factors is concerned with designing things for people's use (Chapanis, 1983). One branch of human factors takes what we know about people (e.g., the capabilities and limitations of our visual and auditory systems, attention, and decision making), and applies that knowledge to the engineering or design of tools and automated capabilities for use in work environments. The broader focus of human factors is on people at home, at work, and at play as they interact with the products of good and bad design, such as video cassette recorders (VCRs), automobile interiors, aircraft cockpit displays, or ATC workstations. (See Norman, 1991; Sanders and McCormick, 1993). In each case, **the goal of human factors specialists is to minimize the potential for design-induced error by ensuring that the equipment is suitable for the user population, the users' tasks, and the users' environment.** The goals of human factors engineers are the same as the users' - to assure the usability and operational suitability of design products for use in ATC environments.

1.2.2 In what sense are we using the word "system"?

A system is a set of components that work together to achieve a goal or a set of goals. In the biology of organisms, we have the circulatory system, the digestive system, the nervous system, and so forth. In the wider environment, we have ecosystems that can be severely disturbed by foreign substances, such as massive amounts of oil. The family is a system, that is, a network of relationships. The kinds of systems we focus on in this handbook are those made up of people and their tools.

One person and a hammer comprise a basic construction system. At the most basic level, the system for controlling air traffic is made up of people and the "things" designed to help them achieve their goal: the safe, orderly, and expeditious

"To be effective, human factors considerations must be introduced at the start of system design."

Chapanis, 1983, p. 14

flow of air traffic. These things include sensing devices, hardware, software, maps and charts, procedures, communication equipment, organizational structures, and so forth.

In everyday conversation, the word "system" is often used to refer only to the inanimate parts of the overall system, that is, the computer components and the software. In a broader sense, it is useful to remember that controllers, supervisors, Airway Facilities (AF) personnel, and FAA management are all part of "the system."

1.2.3 What do we mean by "the systems approach" to design?

Briefly, the systems approach considers each component of the system in relation to the others. It focuses on quantifying relationships between components, including the human, and testing their interactions to predict system productivity. It is indirectly concerned with designing the human-machine system to make the best use of human capabilities and to compensate for human limitations. (See Sheridan, 1988, for a discussion of the systems approach.)

Beyond the Systems Approach.

Thinking of the human as a system component (Rasmussen, 1980) can give the impression that the person is simply another cog in the system. What we need, instead, is a broader perspective, a user-centered approach, that values the person's unique capabilities and puts the user of technology in the driver's seat. We also need to include the capabilities of working teams that exist formally or informally within the ATC system. The most obvious team of this kind is the pair of controllers operating a busy sector, with a clear assignment of responsibility to each and a reliance on both verbal and non-verbal communication between them. The working relationship between these two controllers is as much a part of the system as is the radar equipment.

As we widen our view, we see that the ATC team can include controllers at the adjacent sectors, the area supervisor, adjacent areas, and adjacent regions. Finally, we see that all the parts are related to each other on a national scale that makes up the National Airspace System (NAS). It takes another small step to see that pilots, air crews, and their organizations are parts of an even more encompassing team, the aviation system.

Applying the systems approach to integrating human factors into FAA programs is an agency objective (see the FAA's *National Plan for Aviation Human Factors*, 1991, and the *National Plan for Civil Aviation Human Factors: An Initiative for Research and Application*, 1995). Too often, new subsystems are considered in isolation, rather than parts of an operational whole. This results in a proliferation of individual sub-systems that can "talk" to the controller but not to each other.

It is important to remember that the FAA is designing systems for use not only by individual controllers but also by ATC teams in the field. Design products must support the work of ATC teams as well as the work of individuals. ATC system design must be user-centered, that is, it must be based on the operational requirements defined to make the best use of human capabilities while compensating, to the extent possible, for human limitations.

1.3 KEY POINTS

Several key points are threaded through this handbook. They provide the underlying rationale for the importance of considering human factors in the design of ATC systems.

1.3.1 What are key issues in the human factors of ATC systems design?

Designing a complex, automated ATC system requires designers to address a wide range of engineering issues that relate to hardware, software, and the operational environment. Jobs, tasks, and procedures must also be designed. The basic human factors question considered in this handbook is, "How can the hardware, software, and working environment best be designed for effective, efficient use by trained operational personnel?" It is important to understand that there is no one "right" answer to this question.

> "Design has to start with the user."
>
> Chapanis, 1983, p. 14

Designing for the ATC workforce requires an understanding of the needs and requirements of the workforce. These needs and requirements are not solely for automated tools and capabilities to support job performance, but also for more intangible, though no less important, sources of job satisfaction and peer recognition. **The key challenge to designers is to modernize the controller's tools while retaining an active, involved role for the controller in the system.**

The design process can be focused on meeting this challenge if important human factors questions are, first, recognized as crucial to system success and, second, built into the iterative cycles of design, development, and evaluation. Nine key human factors questions are listed as follows:

- Will the functions performed by the controller provide an involved, active role that fosters a sense of accomplishment and pride?
- To perform a particular function, what information does the controller need, when, and in what form?
- How should the information be presented to controllers?
- Have human capabilities and limitations been considered in mapping information to ATC displays and in mapping human responses to controls?
- Is the computer-controller interface self-descriptive, controllable, suitable for the operational task, consistent with user expectations, and error tolerant?
- Are human decision-making abilities used appropriately?
- Do controllers perceive themselves to be in control of the automated system?

- Are workstations and facility environments designed for the range of individual differences in the ATC workforce?

- Are design products usable, operationally suitable, and acceptable to the ATC workforce?

These questions are based on the human factors literature (e.g., Bainbridge, 1987; Dzida, 1989; Fitts, 1951; Harwood, 1994; Hopkin, 1982, 1988; ICAO, 1993; Norman and Draper, 1986; Sanders and McCormick, 1993; Wickens, 1992; Wise, Hopkin, and Smith, 1991). A detailed set of basic human factors questions is presented in the latest edition of ***Human Factors in Engineering and Design*** (Sanders and McCormick, 1993, pp. 750-752).

Supported by evidence from disasters such as the Three Mile Island nuclear accident, Eastern Airlines Flight 401's slow descent into the Everglades, and the Challenger explosion, this handbook begins by assuming that human factors engineering is not a frill. Without effective attention to human factors, technological disasters will continue to pile up (Roush, 1993). Human factors engineering is integral to system success. As shown in Table 1-1, human factors engineering contributes to numerous important system design elements.

Table 1-1. Major Benefits of Human Factors Engineering
From "Introduction" by A. Chapanis in *Human Factors Considerations in Systems Design*, C. M. Mitchell, P. M. Van Balen, and K. Moe (Eds.), (NASA Conference Publication 2246) pp. 17-20, 1983.

• Reduction of error • Increased safety • Increased system reliability • Reduced time for training • Improved maintenance • Reduction of fatigue and physical stress • Increased efficiency and productivity • Improved work environments • Increased human comfort • Reduction of monotony and boredom	• Improved system design and development • More effective product testing • Improved selection and training programs • More usable system documentation (e.g., user's manuals) • Easier, more user-centered system installation

1.3.2 What are this handbook's key points about human factors in the design of ATC systems?

The key points, elaborated in this handbook, can be summarized as follows:

- Human error is inevitable. It cannot be automated or designed out of a system, The challenge to designers is not only to help prevent errors, but also to design systems that detect, contain, and limit the consequences of human error.

- The human operator must remain "in the loop", that is, have an active, involved role in the path of control action and feedback.

- To be most useful and effective, systems must be compatible with controllers' natural approaches to organizing information, building and maintaining situational awareness, planning, strategizing, and acting.

- To be successful, systems must take human limitations into consideration, being careful, for example, not to overload or underload controllers' capacities in memory and attention.

- Automated aiding must be designed to support and extend controllers' capabilities, not to replace them.

- A user-centered approach to design will meet the objectives of usability and operational suitability far more readily than will a technology-driven approach.

These points are addressed from different perspectives in the various chapters. Every effort has been made to draw implications for design and evaluation from the basic human factors material.

1.4 SOURCES

1.4.1 What are the major sources of the human factors information presented in this handbook?

Major sources of information for this handbook included studies in areas of basic research, such as vision and decision making; studies in applied research in aviation-related environments; and guidelines documents. Many studies and documents directly related to human factors in ATC were identified through a comprehensive literature review conducted through the Crew Systems Ergonomics Information Analysis Center (CSERIAC) at Wright-Patterson Air Force Base (Gravelle and Hecht, 1992). This vast literature is represented in the references and suggestions for further reading, which appear at the end of each chapter.

In the development of this human factors handbook, the authors consulted the material in another FAA human factors handbook (Cardosi and Huntley, 1993), which applies human factors to flight deck certification. In turn, we have applied a great deal of the same background material and the same principles to ATC.

Because specifications for ATC systems require design products to be in compliance with human factors standards, we have sampled the standards in many places. In particular, we frequently cite MIL-STD-1472D (DoD, 1989). This standard sets forth *Human Engineering Requirements for Military Systems, Equipment and Facilities*. These requirements are widely accepted as applying to various operational settings, such as aircraft flight decks, nuclear power plants, spacecraft ground control rooms, manned spacecraft, and ATC facilities. This standard should be part of any human factors resource library. A checklist tied to the numbered paragraphs of MIL-STD-1472D is available from CSERIAC (Lockheed, 1993). Another standard we consulted was the *American National Standard for Human Factors Engineering of Visual Display Terminal Workstations* (ANSI, 1988). Many other additional sources are cited in the references.

1.4.2 How do citations in the text map to the references?

Sources of information are generally cited in the text by the last name of the author and the date of publication, enclosed within parentheses at the end of a sentence. It is not necessary to "read" these citations unless you are especially interested in finding a more detailed treatment of the topic. The full citations for source material are given in the references at the end of each chapter. These are alphabetized by author's last name.

1.5 USE OF THIS HANDBOOK AND CHECKLIST

1.5.1 How should this handbook be used?

It is not necessary to read this handbook from cover to cover. Instead, each chapter can be read independently of the others. If time does not permit reading a whole chapter, you can key in on the questions or topics of interest by scanning the section headings and the questions provided in the left margins.

Locating Topics. Other ways to find topics of interest are to consult the index at the back of this handbook, the main table of contents in the front, or the table of contents at the beginning of each chapter. Cross references within chapters refer you to related material in other sections or other chapters. The checklist's references to sections can also help you find background material on each checklist item.

1.5.2 How should the checklist be used?

Checklists are located at the end of selected chapters, and compiled into a single checklist in the companion volume, titled, *Human Factors Checklist for the Design and Evaluation of Air Traffic Control Systems.*

Checklist Format. Checklist items are generally stated so that they represent qualities or attributes that design products should have. The response section allows evaluators to indicate whether the system addresses the item in a satisfactory or unsatisfactory way. There is also a section for notes.

The numbers in parentheses at the end of each checklist item refer to the section of this handbook that discusses the issue. This mapping allows the checklist user to learn about the . basis for the item, why it is important, and the implications of compromise. Checklist items marked with and "**E**" indicate items that must be assessed with equipment and/or by referring to the specifications documentation.

Checklist as a Multi-Faceted Tool. The primary purpose of the checklist is to point air traffic operations specialists (and other operations specialists) to human factors issues that they may wish to include in their consideration of a new system or subsystem, or a new component of an existing system. Responses to the checklist items can help to focus group discussions and identify issues that should be addressed in every stage of the acquisition process, from the development of system requirements to formal operational testing. Operations specialists may wish to use some of the items as a basis for identifying human factors issues that should be formulated into appropriate requirements and specifications. In this way, use of the checklist can also support the development of a Human Factors Plan (HFP) as required by FAA Order 1810.1F, Acquisition Policy (3/19/93). For more information on how to write a human factors plan see Chapter 2, *Human Factors in Systems Acquisition.*

The checklist is intended to add structure and objectivity to the selection and evaluation phases of acquisition. It is **not** meant to serve as a comprehensive assessment or to replace usability testing. The checklist can only examine individual components of a system and point to broader issues (such as how these components fit together, the uses of automation, etc.). In many cases, the ability of the checklist to identify potential problems will be entirely dependent on the person using the checklist. Where checklist items are general or broad, an intimate knowledge of the system and how the person will use the system, is required to make the connection between the intent of the item and specific system attributes or functions. Many of the checklist items are objective and precise and can be answered with observations alone (e.g., "the user can adjust symbol size"). However, other items are more general and the answer may require objective testing (e.g, "the meanings of auditory displays are readily apparent"). Also, some of the items are idealistic; they represent the ideal based on current human factors knowledge. They are not offered as system requirements or standards, nor do they preclude compromise; where compromises must be made, however, the implications should be clearly understood.

Checklist as a Flexible Tool. A customized checklist should be created for each system under consideration. In its electronic form, a checklist can be created by selecting appropriate items and deleting items that do not apply to a specific system. Customized checklists can be created in one of two ways. Items can be selected, either individually or in sections, from the master checklist (the entire collection of checklist items). Alternatively, checklist items can be reviewed and selected through a search by key words (e.g., automation, keyboards) that have checklist items associated with them.

1.6 DOCUMENT OVERVIEW

Figure 1-1 provides an overview of key human factors issues and topics treated in this handbook. (Detailed discussion of other important human factors topics is beyond the scope of this handbook, although other topics are touched on occasionally.)

Moving out from the center of the figure, you will find key topic areas, such as Human Factors in System Acquisition and Human Capabilities and Limitations. As you continue moving outward from this band of the figure, you will find the chapters where particular topics are addressed. For example, the role of human factors in system acquisition is discussed in Chapter 2. The next three chapters discuss human capabilities and limitations that must be understood in order to ensure that systems are designed to be compatible with the user and help to compensate for human limitations. Visual and auditory perception are discussed in chapters 3 and 4. Chapter 5 explores topics in information processing such as attention, memory, and decision-making.

As you move clockwise around the outer perimeter of the figure, you will find issues such as the potential effects of ATC automation, (Chapter 6); issues in computer-human interface (CHI) design and evaluation are examined in Chapter 7; recommendations on the measurement of controller workload and performance are provided in Chapter 8; guidance on workstation and facility design and evaluation is set forth in Chapter 9; and methods of human factors testing and evaluation are discussed in Chapter 10.

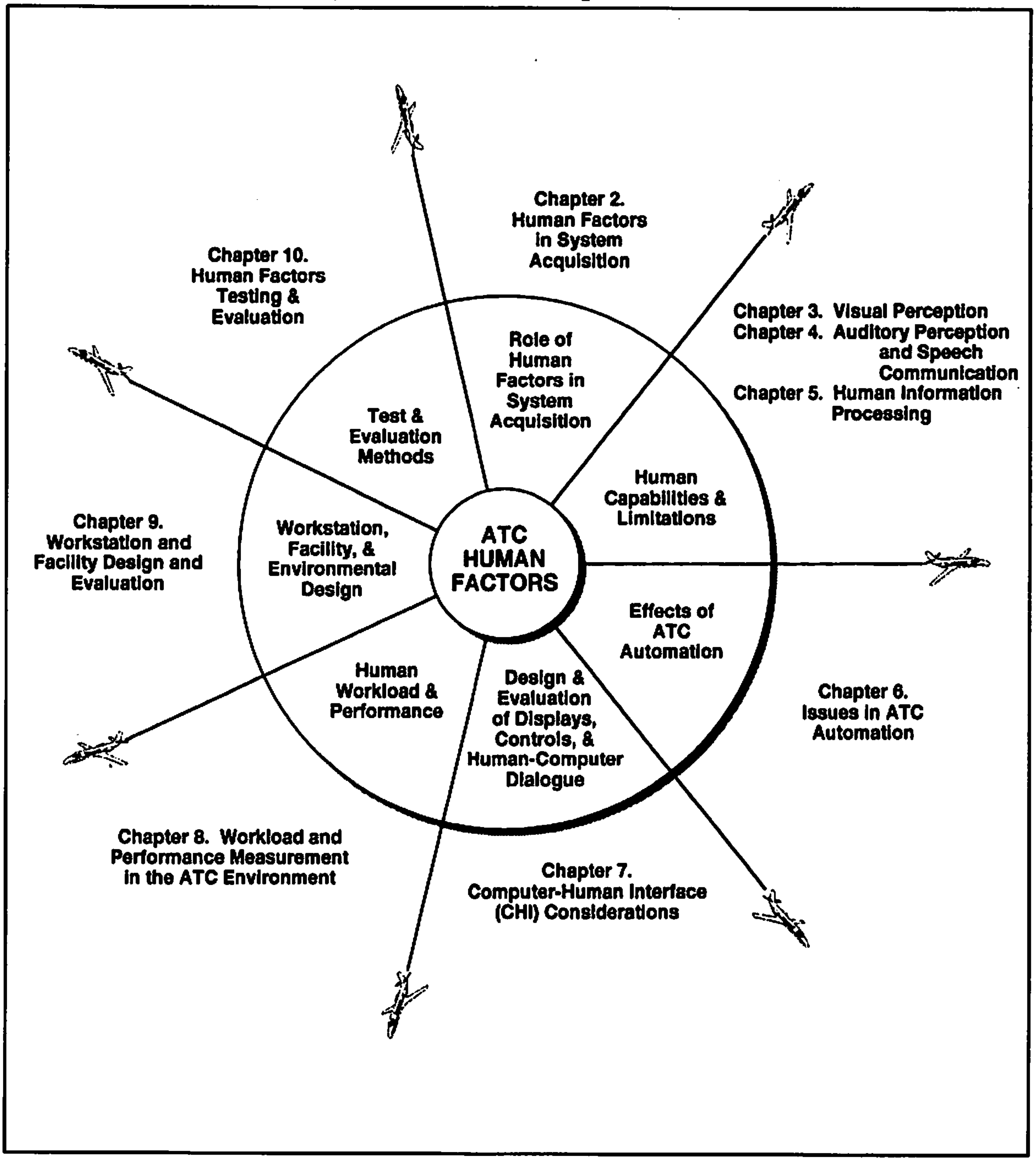
Figure 1-1 does not mean to imply that the topics shown are an exhaustive set, just that these are the key topics covered in this handbook.

1.7 CALL FOR COMMENTS

Users of this handbook and checklist are encouraged to submit comments and make suggestions for improvement (e.g., topics that you would like to see included in a later edition). Please address comments or requests for additional copies to:

Kim Cardosi, DTS-45
U. S. Department of Transportation
Research and Special Programs Administration
John A. Volpe National Transportation Systems Center
55 Broadway, Kendall Square
Cambridge, MA 02142

Figure 1-1. Overview of Key Human Factors Topics Addressed in this Handbook



1.8 REFERENCES AND SUGGESTIONS FOR FURTHER READING

American National Standards Institute (ANSI). (1988). *American national standard for human factors engineering of visual display terminal workstations* (ANSI/HFS Standard No. 100-1988). Santa Monica, CA: Human Factors Society.

Bainbridge, L. (1987). Ironies of automation. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.), *New technology and human error*, (pp. 271-283). New York: Wiley.

Cardosi, K. M., and Huntley, M. S. (Eds.). (1993). *Human factors for flight deck certification personnel* (DOT/FAA/RD-93/5), Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Chapanis, A. (1983). Introduction to human factors considerations in systems design. In C. M. Mitchell, P. M. Van Balen, and K. Moe (Eds.), *Human factors considerations in system design* (NASA Conference Publication 2246, pp. 9-24). Washington, DC: National Aeronautics and Space Administration-Goddard Space Flight Center.

Del Balzo, J. M. (1993). Speech to the ICAO Flight Safety/Human Factors Conference. Reprinted in D. Beringer (Ed.), *The Flyer*, June newsletter of the Aerospace Systems Technical Group of the Human Factors and Ergonomics Society, 2-3.

Department of Defense. (1989). *Human engineering requirements for military systems, equipment and facilities* (MIL-STD-1472D). Washington, DC: DoD.

Dzida, W. (1989). The development of ergonomic standards. *SIGCHI Bulletin, 20(3)*, 35-43. (Published by the Association for Computing Machinery (ACM) Special Interest Group on Computer-Human Interaction, 11 West 42nd St., New York, NY 10036.)

Federal Aviation Administration. (1991). *The National plan for aviation human factors.* Washington, DC: Department of Transportation/Federal Aviation Administration.

Federal Aviation Administration. (1995). *The National plan for civil aviation human factors: An initiative for research and application.* Washington, DC: Department of Transportation/Federal Aviation Administration.

Fitts, P. (Ed.). (1951). *Human engineering for an effective air-navigation and traffic-control system.* Washington, DC: National Research Council.

Gravelle, M., and Hecht, J. N. (1992). *Search & summary: Air traffic control and human performance* (CSERIAC-SS-92-003). Wright-Patterson AFB, OH: Crew System Ergonomics Information Analysis Center.

Harwood, K. (1994). Defining human-centered system issues for verifying and validating air traffic control systems. In J. A. Wise, V. D. Hopkin, and P. Stager (Eds.), *Verification and validation of complex systems: Human factors issues* (NATO Advanced Science Institute, Series F: Vol. 110, pp. 115-129). New York: Springer-Verlag.

Hopkin, V. D. (1988). Air traffic control. In E. L. Weiner and D. C. Nagel (Eds.), *Human factors in aviation* (pp. 636-663). New York: Academic Press.

Hopkin, V. D. (1982). *Human factors in air traffic control* (NATO AGARDograph No. 275). Neuilly sur Seine, France: NATO, Advisory Group for Aerospace Research and Development.

International Civil Aviation Organization. (1993). *Human Factors Digest No. 8: Human factors in air traffic control* (ICAO Circular 241-AN/145). Montreal: ICAO.

Lockheed Human Factors Engineering. (1993). *The MIL-STD-1472D checklist.* Sunnyvale, CA: Lockheed Missiles and Space Co. (Available from the CSERIAC Program Office, Wright-Patterson AFB, OH: 45433-6573: (513) 255-4842).

Norman, D. A. (1991). *The Design of everyday things.* New York: Doubleday/Currency.

Norman, D. A., and Draper, S. W. (Eds.). (1986). *User-centered system design.* Hillsdale, N.J.: Erlbaum.

Rasmussen, J. (1980). The Human as a system component. In H. T. Smith and T. R. G. Green (Eds.), *Human interaction with computers* (pp. 67-96). New York: Academic Press.

Roush, W. (1993). Learning from technological disasters. *Technology Review, August/September,* 50-57.

Sanders, M. S., and McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). New York: McGraw-Hill.

Sheridan, T. B. (1988). The System perspective. In E. L. Wiener and D. C. Nagel (Eds.), *Human factors in aviation* (pp. 27-51). New York: Academic Press.

Wickens, C. (1992). *Engineering psychology and human performance* (2nd ed.) New York: HarperCollins.

Wise, J. A., Hopkin, V. D., and Smith, M. L. (Eds.). (1991). *Automation and systems issues in air traffic control* (NATO Advanced Science Institutes Series F: Computer and Systems Sciences, Vol. 73). New York: Springer-Verlag.

1.9 ENDNOTES

1. Following customary usage, we refer to ATCSs as "controllers" and to the various Air Traffic Services as "Air Traffic."

2. Ergonomics comes from two Greek words, "ergon" and "nomos," that together mean the laws of work. Ergonomists (or human factors engineers) are interested in discovering the laws of work and in designing jobs and work environments for the benefit of workers and their productivity.

CHAPTER 2. HUMAN FACTORS IN SYSTEMS ACQUISITION

Glen Hewitt

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 2. HUMAN FACTORS IN SYSTEMS ACQUISITION

2.1 INTRODUCTION

Human error remains the most common contributing factor in aviation accidents and incidents, yet strategies for mitigating their impact are well known and widely documented. Acquisition handbooks and system engineering guides identify equipment design, training and personnel selection, procedures, and organizational structures and management as causal of human-system interface discrepancies. Because each application of new technology poses unforeseen difficulties in these interface areas, technology must be centered on the needs, capabilities and limitations of the user rather than the capabilities of the technology. Such an approach to system acquisition and engineering, as emphasized in the FAA's *National Plan for Civil Aviation Human Factors: An Initiative for Research and Application,* places the user (operator and maintainer) at the center of design and development considerations. Failing to consider the performance of the human component of the system increases program cost and schedule, and jeopardizes system technical performance. However, **ensuring that system designs are human-centered, will lead to enhanced system performance, reduced safety risks, lower implementation and life cycle costs, and a higher probability of program success.** It is never efficient to play "catch-up" with human factors in any acquisition.

Qualified human factors support during system design and acquisition provides an engineering perspective to the incorporation of information about human capabilities and limitations related to how people receive and process information, solve problems, and interact with the system and its components. Where sufficient information is not already available to address a concern, the human factors discipline also offers an arsenal of methods for obtaining time-tested and objective performance measures. Tools and methods used to analyze system performance with the human component as an integral part can help to provide reliable information on which to base developmental and operational decisions, efficiently.

One of the main goals of human factors engineers is to ensure that the users' interests are considered in all stages of system acquisition and development. In this sense, **the goal of the user and the human factors specialists are the same - to ensure that the system is usable and operationally suitable to the user, the task, and the environment.** However, the role of selecting the optimal human-system interface should not be the sole proprietorship of the user for three reasons. First, system design decisions [including selection of design in nondevelopmental items (NDI) and commercial off-the-shelf (COTS) acquisitions] always involve complex trade-offs among optimal performance, cost, and competing technological alternatives. These decisions require program management and engineering expertise from a multitude of disciplines, including that of the human engineering community. Second, decisions and alternatives related to human performance may be counter-intuitive unless scrutinized and analyzed by those skilled in quantifying the relationships. For example, the decision to assign some system diagnostic tasks to an operator or a maintainer may require detailed analyses of life cycle cost, alternative training strategies, and personnel selection criteria. Third, users are not always the best judge of what will provide the best operational performance. For example, users may want many more system features than they will frequently use (such as excess display information) which can lead to distracting clutter, confusion during emergencies, and non-standard interfaces that complicate training and supervision. Having human factors engineers involved in the acquisition process helps to ensure that the system is easy to use, unlikely to induce errors, and tolerant of common and uncommon human performance anomalies.

Human factors need to be considered early in the process, such as when the requirements are identified. Requirements that are written with systematic and explicit consideration of human factors issues lay the foundation for ensuring that these issues are addressed at the earliest possible stage of development (or procurement). **Identifying potential human factors problems early has the advantage of remedying situations when they are easier and less costly**

to fix. Too often, human factors engineers aren't involved in the process until the developmental or operational testing stage. Addressing human factors problems at this stage can lead to costly changes in design, delays in implementation, or degraded performance. While managers in charge of programs are often, and understandably, reluctant to incur additional costs (such as those associated with baselining human-system performance or establishing human-in-the-loop performance criteria), ignoring human factors concerns can prove even costlier when errors are induced by inadequate design.

2.1.1 What is the relationship between human factors engineering and system performance?

Years of experience in human factors engineering has led to an understanding of the many interactions that contribute to system (operational) performance. Below are five axioms of performance.

1. Human Performance Affects System Performance. Despite continual advances in various technologies, the characteristics and aptitudes of people who operate and maintain systems change very little over time. The constancy of the human component implies a limiting factor for system performance as well as design considerations for the engineers. Systems must be designed (or selected from competing designs) to complement human capabilities and mitigate their limitations. **A well-designed system will enhance human performance, while a poorly-designed system will degrade it.**

2. Skill is a Function of Aptitude and Training. While the characteristics and aptitudes of people that operate and maintain systems change very little over time, what may change is the skill that people bring to the work environment. This skill is dependent upon how those raw aptitudes (which can provide a basis for selection) have been honed to perform certain tasks. Formal training is the most common method for increasing the level of complex skills. Experience (which is another form of training) is necessary to maintain these skills. Performance enhancement (or degradation) as a result of training (or lack of it) can be quantitatively measured to predict the impact of familiarity and exposure to the system over time.

Training must be considered from the beginning of the requirements phase. Training strategies depend upon training limitations. How many controllers will need to be trained? How much time and expense can be incurred in the training? How much experience is brought to the features of the new system, and will that experience help or hinder the operator's performance? How many different types of training must be developed to accommodate different initial skill levels? The human factors engineer provides systematic measures to identify the rate of skill accretion or skill decay as a function of familiarity with the system components such as the equipment, procedures, configurations, and operational scenarios. Finally, while training is often used to moderate technological, cost or other design limitations, it should never be used as an alternative to proper human factors planning and human-system design.

3. Performance can be Measured by Time and Accuracy. Performance is a matter of degree. Many methods have been devised to help evaluate human and system performance. The most objective and robust are those that determine how well system objectives are achieved and how much time is required to do so. How well a system objective is achieved is measured in terms of response accuracy and the time required to do so is called "response time". For example, a set of performance measures may indicate the effectiveness of a specific alert. One measure would provide information about how long it takes an operator (e.g., controller or airways facilities technician) to notice the alert and respond to it (i.e., response time). Another measure would provide information about the frequency that the operator selected and performed the appropriate response (i.e., accuracy). Measures of system performance should always include collection of data that can be related to these terms.

4. Tasks Are Determined by Equipment Design. In an ideal world, all tasks would be equally easy to perform and have an equal impact or consequence upon system performance. However, how a system is designed determines the variety of tasks required to use it. A new system, or an addition of a subsystem to an existing system, can change the way people do their jobs. It can eliminate the need to perform certain tasks; it can change the way these tasks are accomplished; and it can create new tasks that were not performed with the existing system. It is important to know (in advance of deployment) how changes in the system affect the operator and maintainer tasks. For example, a poorly designed system may create the need for unduly complex procedures to make the system perform properly. **Critical tasks (those that impact upon system performance the most) are determined by system design and must be analyzed adequately.**

The tasks that have the greatest impact on system performance need to be identified before the system can be evaluated. Identifying these critical tasks requires knowledge about how the system is to be used, maintained, supported, and how people will be performing these tasks in the intended environment. Once the critical tasks are identified, then the measures that will be used for evaluation (i.e., to determine whether these critical tasks can be accomplished successfully in the new system) can be selected. **System design alternatives and trade-offs can create or eliminate critical tasks that are easy or difficult to execute.**

5. Task Performance Is Affected by the Equipment Designers. Equipment designers need to consider the capabilities and limitations of the people who will use, maintain, and support the system that they are developing. **How well, or poorly, the design takes into account the background, experience level, and training of the users will determine how well people will be able to use the system.** In this way, the designer of the system can enhance or degrade the performance of the user. For those systems developed for broad applications (especially NDI and COTS hardware and software), selection of the vendor determines the design selection, but the same principles apply.

2.2 EARLY AND CONTINUOUS FOCUS UPON HUMAN FACTORS CONSIDERATIONS

2.2.1 When and how should human factors be considered in the acquisition process?

The performance of a system is a function of the equipment (hardware and software), the performance of the people associated with the equipment, and the environment in which the equipment is placed. Since the human performance element is a crucial component, human factors issues must be considered and documented throughout the acquisition process - from the analysis and determination of requirements and system specifications through operational testing and evaluation to deployment. As Figure 2-1 shows, early consideration of human factors issues will result in potential problems being detected earlier, and resolved easier, than if human factors planning is delayed. Furthermore, providing a "lifecycle" human factors perspective will ensure that the new system or sub-system is fully integrated into the working environment, that the proposed training and staffing for the new system is effective and efficient, that the expected level of performance will be achieved, and that necessary enhancements to the system are objectively measured and documented. Lifecycle human factors support provides economical, as well as operational, benefits. Figure 2-2 depicts the costs of resolving human factors issues as a function of the phase of acquisition in which these issues are addressed. Initially, it is more expensive (in terms of time and funding) to deal with human factors considerations than to ignore them. However, an initial human factors investment pays high dividends, in terms of costs and schedule, in later stages of acquisition where changes are more costly and difficult to make.

The successful application of human factors to an acquisition program requires a holistic approach where the operator and maintainer are viewed as an integral component of the system. To achieve such an approach entails explicit, systematic, and iterative consideration of the human-system interactions. While human factors discrepancies and

Figure 2-1. The Relation between Human Factors Planning and the Phase in which Human Performance Issues are Detected and Resolved

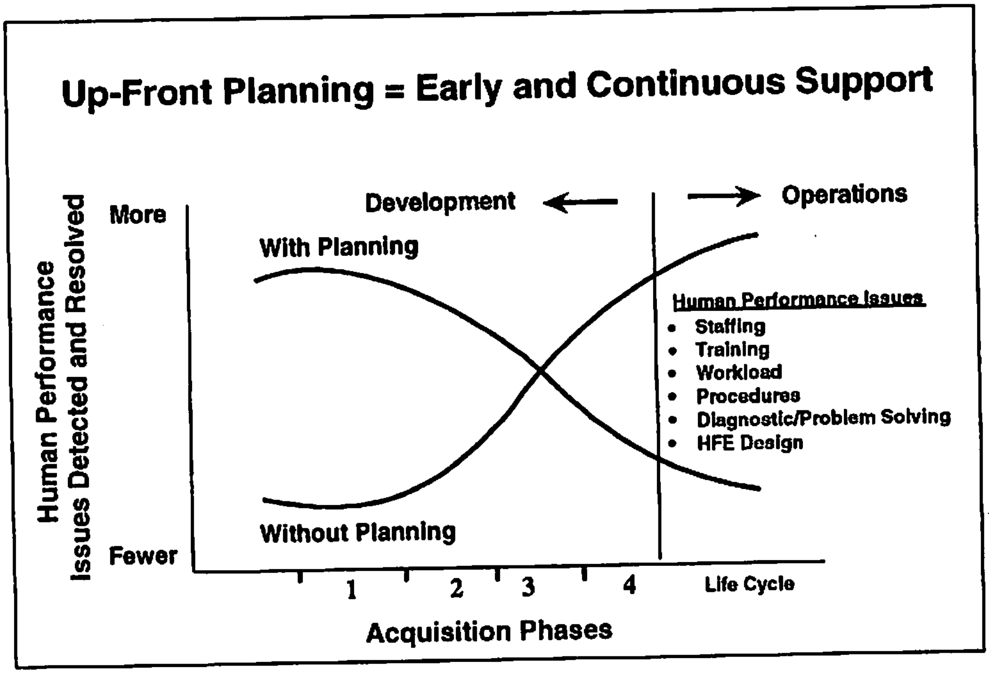
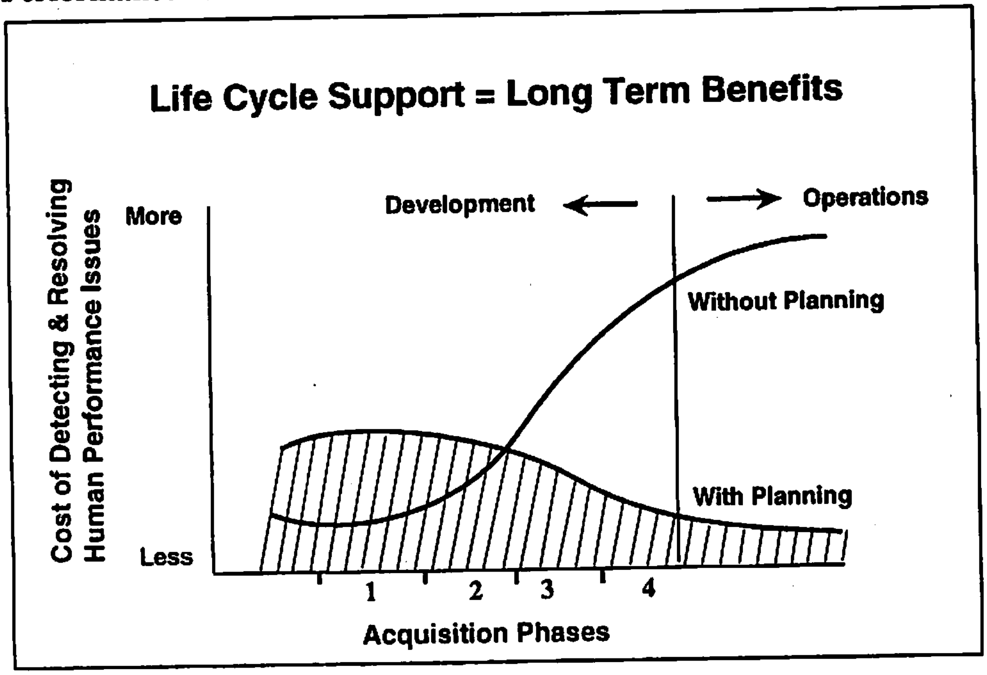


Figure 2-2. Effects of Early Human Factors Planning on the Cost of Resolving Human Performance Issues



problems are likely to arise during any phase of the acquisition, the means and methods to identify and resolve the human performance implications of equipment design, training, procedures, and organizational and management alternatives must be planned and initiated at the earliest stage and continuously monitored and executed at each stage of system development. This systematic and coordinated approach requires many interfaces with other engineering disciplines and acquisition participants. Some of the more important coordination activities occur in the following areas:

Systems Engineering: Human factors engineers should participate in the development process to help examine trade-offs, evaluate alternatives, predict the consequences of system changes, and ensure that human capabilities and limitations are considered in design enhancements. They can also assist in identifying the personnel, skills, and expertise needed at each phase of the acquisition to ensure proper system development.

Training: Human factors engineers can work with training specialists to coordinate strategies and issues that impact staffing and personnel selection requirements, human and system performance, task standards and operational conditions.

Configuration Management: Human factors engineers can help to identify human performance implications of proposed design changes to the current configuration of the system and the impact of configuration variations.

Test and Evaluation: Human factors specialists should assist in developing and executing the plans and procedures for testing and evaluation. They should also participate in identifying the critical issues to be addressed in the test and the objective measures and criteria that will be used to determine whether the system meets operational goals. In addition to planning and conducting the tests, they can help to ensure that the data that result from the tests are analyzed and interpreted appropriately.

Integrated Logistics Support: Human factors specialists can help to identify what human resources will be necessary to operate and maintain the system over its life cycle.

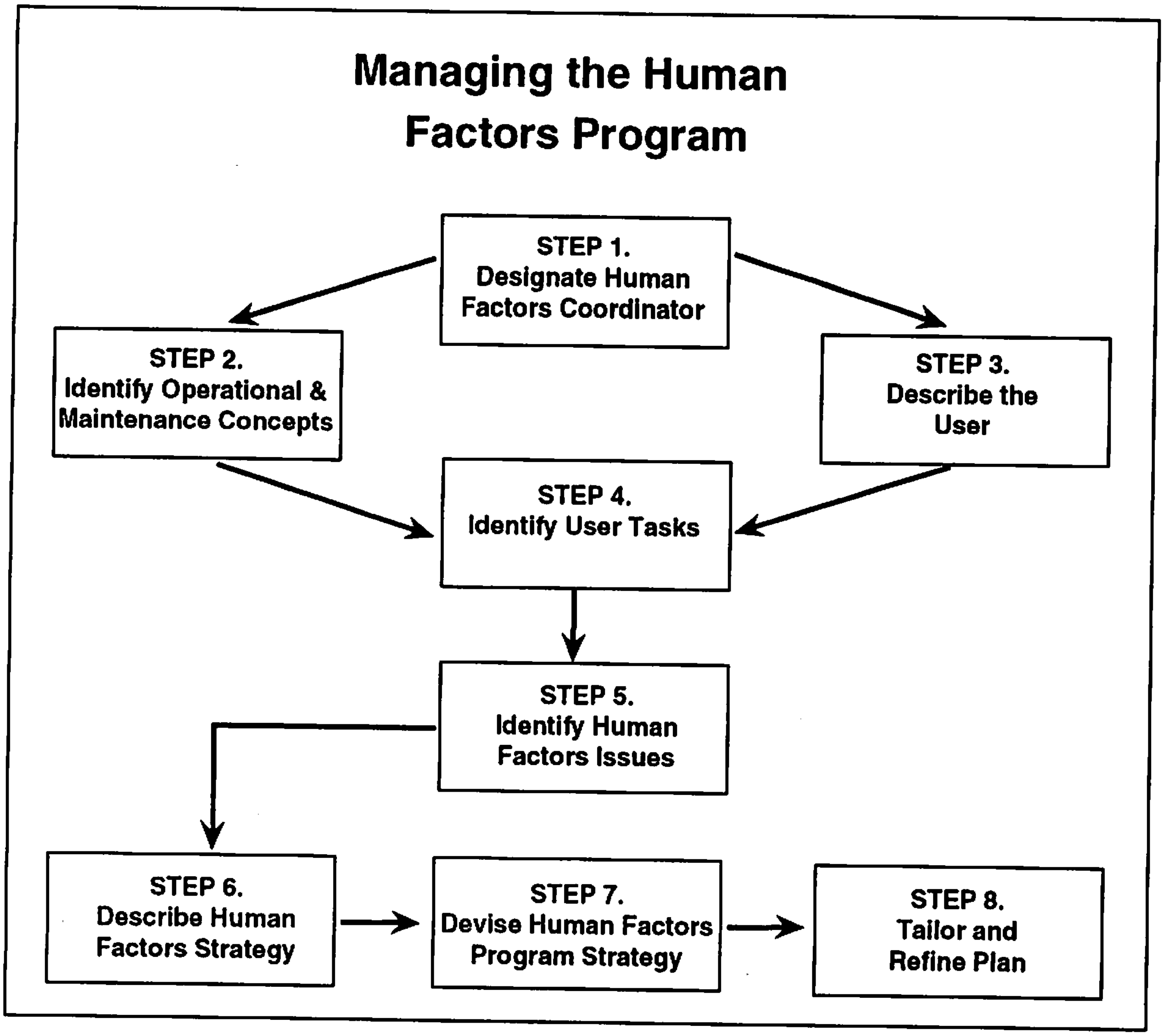
2.3 DOCUMENTING HUMAN FACTORS CONSIDERATIONS-DEVELOPING A HUMAN FACTORS PLAN

2.3.1 How do you develop a human factors plan?

FAA Order 1810.1F, Acquisition policy (3/19/93), requires that a Human Factors Plan be developed for system acquisitions. The purpose of a human factors plan is to ensure that human factors considerations are fully incorporated into the system development. As of this writing, it is the responsibility of the Integrated Product Team (IPT) to initiate the human factors plan. This plan addresses the human performance and human resource parameters for program and design alternatives. The plan is first developed during Phase 1 and updated during each subsequent acquisition phase. The initial human factors plan outlines the issues, tasks, and strategies associated with human considerations in the operation, maintenance, and support of system options. Subsequent updates to the plan further define and refine the human parameters of the program and ensure identification and remediation of human factors problems and issues in the program. The human factors plan should be designed to be a living document; it should be tailored to the specific program requirements, procurement strategy, key decision point, and acquisition phase as well as customer considerations of the program. It imposes only the necessary and reasonable requirements to achieve the objective effectiveness of human performance during system operation, maintenance, and support; and the efficient use of personnel resources, skills, training, and funds.

The following steps (illustrated in Figure 2-3) are necessary to develop a Human Factors Plan:

Figure 2-3. Steps Necessary to Develop a Human Factors Plan



Step 1: **Designate a Human Factors Coordinator.** Appoint someone to coordinate human factors who will be integral to the IPT. This person will develop, direct, and monitor the human factors program and its activities for the system acquisition. This person will be the focal point for overseeing the human factors planning, coordinating, monitoring, and execution.

Step 2: **Identify Operational and Maintenance Concepts.** Describing how, and under what conditions, the system will be used and maintained is necessary to identify the human resource requirements (e.g., the type of personnel required, the number of labor hours that will be required to operate/maintain the system, the training that they will need to receive). Such a description is also necessary to identify the user functions or tasks (especially those that are critical to operations). These operational and maintenance concepts should include:

Location, physical environment, and workspace - For example, do the space limitations and ambient lighting conditions (which pose a range of lighting conditions from intense sunlight glare to dark adaptation) impose human-system performance and design constraints and limitations?

Operational conditions and limitations for the system - For example, is the system expected to perform equally well for all operating conditions (such as bad weather, poor visibility)?

Operational scenarios in the employment of the system - For example, will the system be employed similarly at the extremes of high and low workload periods?

Step 3: **Describe the Users.** Describe the operators and maintainers of the system. This may include air traffic control specialists, traffic management specialists, airways facilities technicians, or others supporting the equipment. It is important to identify the users sufficiently to understand the skills of the users, the practices and procedures they are accustomed to following, the other duties that the users will perform concurrently with using the system, and the resources (e.g., equipment, tools, references, job aids) with which they will be performing these duties. Therefore, this section of the plan describes the relevant characteristics of these users that may affect system performance. These may include: biographical data (such as average age or length of job experience), training history, anthropometry, aptitudes, task-related experience, supervisory relationships, and organizational structure.

Step 4: **Identify the User Tasks.** Focus early on what operators, maintainers, and supervisors will have to do. It is particularly important to identify any new functions or tasks the controller will perform with the new system that he/she did not perform with the old equipment as well as tasks that will be performed differently. Examples of tasks that may be included are: create, enter, delete, or modify (update) a flight plan; identify which sector accepted an interfacility hand-off; display traffic management information; display weather information; predict the traffic situation at a specified time; and display the predicted results of selecting a specific resolution to a potential conflict. This section of the plan should consider the following:

Functions/tasks to be performed by the system as described in operational concepts;

Predecessor systems and equipment and the functions/tasks they required; and

Configurations of the system that cause variations in functions or tasks.

Step 5: **Identify Issues.** Having described what people must do (and under what conditions), identify the potential risks to (and opportunities for) human and system performance as well as resource costs (e.g., selection, staffing, training). For example, this could include: whether or not additional staffing could be required; time and resources available for training users; space limitations in the work environment; whether or not the system could impose any additional workload on the operator; and other areas where human-system performance risks reside. Human factors issues monitored during the program development should consider:

Key human engineering design goals (for the individual user and for the system);

Potential problem areas, constraints, and resource limitations; and

Critical unknowns.

As the system develops and more information becomes available, these goals should be refined in terms that directly relate to performance. All design goals, such as "The system shall be easy to use and minimize the probability of human error" need to be operationally defined. This specifies how they will be measured and determined to be acceptable. For example, a design goal of "minimizing cognitive workload" may be operationally defined in terms of the memory loads imposed by the system on the operator, the number of mental transformations of data that the system requires of the operator, or in terms of how well the task is performed (time and accuracy) when such memory loads are imposed.

Measures of effectiveness or measures of performance that will be used to determine that the system is acceptable should also be specified at this time. For example, an evaluation of an auditory alarm might include the following criteria: users should be able to recognize a specific auditory alarm and respond to it within 10 seconds with an error rate less than 5%. Eventually, specific performance criteria should be described for every critical function or task for which the system will be used. If the measures to be used can be specified early, they should be included in program documentation (e.g., specifications, Statement of Work, test plans).

Step 6: **Describe Human Factors Program Tasks and Activities.** List and describe the tasks and activities that must be accomplished during the execution of the human factors program. There are two dimensions to this:

Key human factors engineering activities from the specification of requirements to the operational evaluation and beyond. For example, this could include the procedures that will be used to identify the tasks (e.g., a task analysis), a description of the evaluations and how they will be conducted, and a description of how the results of the evaluation will be analyzed. (Depending on the issues to be investigated, the evaluation procedures can range from a user survey to a full mission simulation. The method should be appropriate to the issues and resource limitations. How the results will be analyzed e.g., what statistics will be performed on the data, should be planned with the test so that it is appropriate to the test.) This section should also include efforts necessary (e.g., studies, analyses, demonstrations, etc.) to answer the unknowns and resolve the higher risks.

Activities to integrate human performance in the program (in addition to designating a Human Factors Coordinator and writing a Human Factors Plan). This includes what the government must do to direct and monitor the human factors effort and how the program's contractual efforts will be maintained, influenced, and assessed. For example, it is necessary to consider: what are the reporting channels for human-system performance concerns and discrepancies, what relationship does the human factors coordinator have with other system engineering personnel, how will the human factors representative work with the contractor, what must be done to put the proper relationships in place, and what actions must be accomplished to set the actions in motion.

Step 7: **Devise a Strategy.** Develop a concept and approach for achieving the objectives of the Human Factors Program. This strategy should establish how the Human Factors Program will be controlled and fulfilled.

Step 8: **Tailor and Iterate the Planning.** Because each system acquisition program is unique in its pace, cost, size, complexity, and human-system interaction, the human factors planning will vary from program to program. Also, as development occurs, some system and human factors requirements will change. The planning steps and their sequence should be tailored as necessary. Human factors planning must be iterative. The Human Factors Plan should be a living document that changes as the program develops. Adjust as appropriate -- but focus on how the government manages the human factors efforts and how the human performance and human resource issues are resolved.

2.4 CONTEXT OF HUMAN FACTORS PLAN

2.4.1 What information should the human factors plan contain?

Intent. The intent of the human factors plan is to focus the government's attention on human factors concerns, issues, tasks, analyses, objectives, and strategies that will integrate human performance considerations throughout the program. The plan is not intended to provide specifications or work statements (although it should lead the government to include certain specifications or work requirements in contractual documentation). It is intended to specify how the government will direct and control the identification and resolution of human performance issues in the program.

Use. The plan should describe the efforts to be undertaken to ensure that human factors issues are considered at every stage of acquisition (including NDI and COTS). It should begin to sort out the high-priority issues, refine them, separate what is known from what needs to be known, translate unknowns into contractual tasks, shape managerial requirements, and stimulate input to program documentation. As a living document, it should follow the progress of issue resolution during design, development, and acquisition activities.

Emphasis. Emphasis in the human factors plan should be placed on government controls, products of the human factors effort, and performance requirements:

Government Controls. Deciding how the government will control the human factors endeavors entails outlining the people, methods, and process, that the government will employ to get the work done. How the government will ensure that human factors are integrated in the design of the system and in each step of the acquisition process should also be discussed.

Products. Determining the products of the human factors effort entails identifying what needs to be known and how that information will be acquired. The more specific these tasks/activities are, the sharper the human factors focus will be in the documentation (request for proposal or statement of work). Identifying specifics will lead to increased efficiency of the developer/vendor human factors effort, and thus, a higher quality, less costly response.

Performance Requirements. The plan should assist in determining the performance requirements of the system with people in the loop. System performance requirements (with the human involved) should be derived from source documentation to include the Mission Analysis information, acquisition planning documents (such as the Acquisition Plan), life cycle cost documents (such as the Cost-Benefit Analysis), the Operational Requirements Document (ORD), and other documents containing concepts for operation and maintenance.

The tasks and analyses that need to be conducted to support the definition and evaluation of performance requirements should be specified in the plan. Human factors requirements should be expressed in terms of time to accomplish, tolerable errors, time to train, numbers and types of people available, etc. These requirements should not be arbitrary or capricious; they should be derived from the system performance that is expected when the equipment is employed.

Performance requirements with the human in the loop are the basis for including human factors engineering in all stages of acquisition (e.g., mission need statements, proposal preparation and source selection, system design activities, testing, documentation, and program reporting). These requirements are not only needed to convey human factors system requirements to contractors for full development projects but, under an NDI or COTS strategy, they can serve as criteria for product selection during market surveys, etc. They will also serve as the basis for assessing training development and for establishing test and evaluation acceptance criteria.

2.5 CONTENTS OF HUMAN FACTORS PLAN

2.5.1 Is there a suggested format for the human factors plan?

The recommended content and format of the Human Factors Plan (as summarized in Table 2-1) is:

BACKGROUND: Provide a brief description of the program, the equipment (including concepts for operation and maintenance), its schedule, and the people (target audience) who will be affected by the operation and maintenance of the system.

Program Summary. Provide a brief description of the program (including relevant concepts for operation and maintenance).

Program Schedule. Provide an overview of the program schedule.

Target Audience. Identify the population that will be affected during the operations and maintenance of the system. Include a description of any relevant demographics, biographical information, training background, aptitudes, task-related experience, anthropometric data, physical qualifications, organizational relationships, and work space requirements. (Lengthy descriptions may be included in an appendix.)

Guidance. Summarize any decisions, direction, or previous guidance that will impact the human factors approach or results.

Constraints. Identify the known or anticipated limitations (e.g. technical, manpower, training time available) that will affect human performance, personnel resources, training, and human factors engineering goals and system requirements. It is useful to be as specific as possible about what the constraints are and how they might impact the program. For example, there may be an assumption that no additional staffing will be required (e.g., that the system will be operated and maintained with existing air traffic and airways facilities personnel).

If so, this should be viewed as a constraint. Therefore, detailed assessments will be required to determine training requirements, impacts on current functions, etc. Such assumptions are likely to affect job design, workload, etc. Time available for training is another resource constraint that could affect operator, maintainer, or certification tasks and system design. Availability of simulation capabilities may provide a constraint to the testing and analysis. While general constraints would be described in this section, it may be more appropriate to discuss constraints that apply to specific issues in the section below.

ISSUES AND ENHANCEMENTS. List and describe the problems, concerns, deficiencies, risks, and opportunities to be addressed by human factors efforts during the system development. (If this list and description of issues become lengthy, the details may be included in an appendix.) It may be useful to divide the human factors issues into areas (or domains). These domains might include personnel resources (e.g., manpower and staffing), training, human engineering, and safety (including health). Other organizational structures for the issues are acceptable.

Issue Description. Describe the issue or problem background, importance, and consequence.

Objectives. Identify the objectives to be met, obstacles to be overcome, and the planned solution. This section also provides performance measures and criteria that will be used to evaluate resolution of the issue.

The objectives should provide quantifiable operational measures. For example, if alerts are to be sounded, it is important to determine how quickly and with what error rate a person must recognize and react to such an alert. The human factors plan should specify the performance thresholds that make a difference in accomplishing the task, function, or mission.

If the human performance thresholds are unknown, then the human factors plan should identify a task for the developer (or an approach for the government) that will produce the required information. This information will need to be available early in the development (or procurement) to influence the requirements, design, development, and testing.

The major human factors performance requirements and criteria may be reflected in existing systems, baselines, or processes that are being replaced. For example, one major human factors challenge associated with the program may be to accommodate the new "XYZ System" without degrading co-existing systems during the transition period.

Actions. Identify the actions to be taken in remediation of the issue and current status of the issue.

ACTIVITIES. Provide a list and description of each activity (e.g., tasks, studies, analyses) to be performed during the acquisition in support of resolving the issues and controlling the human factors program. This section should identify what the government and the system developer must do to satisfy human factors concerns.

Activity Description. For each phase, describe the activities to be performed; the rationale (i.e., reasons for the activity to be conducted); the technical information needed, data requirements, and data sources; the estimated resources (e.g., time, personnel, funding) required to complete the activity; and the agency expected to perform the activity. This section should address those tasks, analyses, and studies that must be done by the developer or vendor in support of the human considerations. For example, one deliverable should be a Human Factors Program Plan (such as that addressed in MIL-STD-46855) which delineates the contractor's approach in terms of what and how human factors will be executed to meet contractual requirements. Another example of an activity, a Functional Analysis, may need to be conducted to determine how functions should be allocated to the equipment and the users of that equipment. A Task Analysis may be done to delineate tasks for which operators and maintainers will be responsible. A Manpower Assessment may be considered to verify that there is no additional staffing required. There are others.

Activity Schedule. Display the activities to be undertaken and their relationship to each other and to other significant program activities and events. This paragraph should show how things that need to be done in support of the human factors effort are to be integrated in the acquisition program, especially in relation with other program activities and with the key decision points. It should identify the relationships (feeds, dependencies) with other efforts (e.g., market surveys, training, test and evaluation).

STRATEGY: The strategy should be addressed from the top down, but also built from the "ground" up. That is, it should be derived from the major concerns and issues and describe the approach for human factors in the program, but it should address specifics such as schedules, tasks, guidance, constraints, and objectives. This is the place to identify the concepts used to guide the government's control of the human factors effort.

Goals and Requirements. Identify the major human factors performance objectives necessary to achieve compatibility and suitability with the operational and maintenance concepts. A "motherhood and apple pie" approach will not prove helpful in defining the human factors activities. It is necessary to be specific and address key questions such as, "What objectives does the government wish to achieve?" and "How will the government accomplish these objectives?"

Approach. Describe the general approach to be taken to achieve the human factors goals and requirements, meet customer operational needs, and resolve major issues.

Identify who will be responsible for the human factors effort and what direction they are going to take. Discuss how the program office will proceed with the human factors effort. Consider the following questions: To what degree will the program office use contractor support to monitor or assess the system developer's design? How will the government's human factors representative participate in the program (conducting market surveys, preparing the Statement of Work, evaluating offerors during source selection, identifying Critical Operational Issues and Criteria, conducting preliminary design reviews, briefing program reviews, collecting data during testing, etc.). What will the representative do in resolving the major issues that will or may emerge? Will research need to be done? How much government simulation or analysis will be required? How will this be assessed? Who will do it? Is there to be an effort directed toward modeling human performance? Are mockups to be developed? What is the human factors role in them? Is useful information available from other programs or agencies? What analyses might need to be done by the developer of the system? What coordination is necessary to link up with the logistics or training people and their efforts to avoid redundancies and to capitalize upon work already performed? Will experience with other systems be useful in identifying the general scope of the issues before they are clarified and resolved by the contractor for the specific FAA environment? What processes will be used to fulfill the human factors program objectives?

References. This section should identify relevant references (i.e., published text or other sources of important information) needed for the full understanding of the human factors plan. If there is a lengthy list, it may be addressed in an appendix.

REVIEW, APPROVAL, and DISTRIBUTION. This section identifies how the human factors plan is to be administratively handled. If the IPT is responsible for the plan, then the IPT will coordinate with (and provide copies of the plan to) the appropriate engineering representatives of the IPT, program sponsor, and associated organizations.

Review/Approval. This paragraph indicates the review process established for the human factors plan. If no other arrangement is made, review may be assisted by organizational representatives of the human factors engineering discipline. The paragraph should indicate where and when the initial plan (and all updates) will be submitted for review/approval. Normally, the cognizant development Integrated Product Team member or Service Director (as applicable) is appropriate.

Distribution. This paragraph should show where copies of the human factors plan (initial and updates) will be forwarded. As a minimum, the plan should be provided to system engineering representatives of the program and those representing disciplines with critical interface.

Table 2-1. Human Factors Plan Content and Format

HUMAN FACTORS PLAN CONTENT AND FORMAT		
Headings		**Content**
Background	Program Summary	* Brief description of the program * Concept of operation and maintenance
	Program Schedule	* Overview of system acquisition schedule
	Target Audience	* Identify the user and maintainer * Demographics * Biographical data * Previous training * Aptitudes * Task-related experience * Anthropometric data * Physical qualifications * Organizational relationships * Work space requirements (Appendix if data are lengthy)
	Guidance	* Summarize any guidance received
	Constraints	* State if additional staffing is required by the new system * State whether an existing job series will be used or a new one created * Post limits on the amount of time that can be afforded for training * Establish standards on the working conditions that will be acceptable when the new system is fielded
Issues and Enhancements	Issue Description	* Describe the issue or problem background, importance, and consequences or task to be done to support the acquisition
	Objectives	* Identify Human Factors Program objectives * Provide performance measures and criteria in terms of time and accuracy to perform tasks to evaluate resolution of issue * When human performance thresholds are known, identify tasks for the developer to be done early enough in the acquisition to influence requirements and system engineering * Identify the actions to be taken to resolve each issue * Show the current status of each issue

Table 2-1. Human Factors Plan Content and Format (continued)

(Table continued from previous page)

	Actions	* Identify actions to be taken to resolve issues * Show current status of each action
Activities	Activity Description	* Identify any tasks, studies, or analyses that must be performed to resolve the issues (e.g., Human Factors Program Plan per MIL-STD-46855. Functional Analysis to support equipment vs. people allocation of functions. Task Analysis to produce a specific operator and maintainer task list)
	Activity Schedule	* By acquisition phase, describe the human factors tasks in terms of who, what, when, and how (resources) * Identify feeds to and dependencies on NAILS, training, and test and evaluation programs
Strategy	Goals and Requirements	* Strategy should be derived from the major concerns, issues, schedule, tasks, guidance, constraints, objectives, and approach for the Human Factors Program * Answer the question, "What objectives does the government wish to achieve?" * Answer the question, "How will the government accomplish these objectives?"
	Approach	* Define who will be responsible for the Human Factors Program * Set out the extent of contractor support required * Define how the Human Factors Coordinator will support the acquisition team
	References	* Identify relevant references needed for a full understanding of the human Factors Plan (Use appendix if appropriate.)
Review, Approval & Distribution	Review/Approval	* Identify administrative handling procedures * Identify update schedule and procedure * Identify review and approval authority and procedures
	Distribution	At a minimum, the Human Factors Plan and each update should go to system engineering representatives and personnel with oversight responsibilities.

2.6 INFORMATION REFERENCES

2.6.1 Where can I get additional information?

The material provided in this chapter will be useful in identifying human factors considerations that should be addressed in a human factors plan. Other references for developing a human factors plan include:

- FAA Order 9550.8, Human Factors Policy

- Paragraph 4-9, FAA Acquisition Order, 1810.1F, dated March 19, 1993.

- Chapter 9, FAA Acquisition Guide for Program Managers, dated April 1994.

- MIL-STD-46855, Human Engineering Requirements for Military Systems, Equipment, and Facilities, dated May 26, 1994.

- MIL-STD-1472, Human Engineering Design Criteria for Military Systems, Equipment, and Facilities.

2.7 ASSISTANCE

For more information or assistance, call the Office of the Chief Scientific and Technical Advisor for Human Factors, AAR-100, (202) 267-7125, FAX (202) 267-5797.

APPENDIX 2A. CHECKLIST ITEMS

1. Responsibilities for who will develop and control the human factors work are specifically designated (2.3.1).

2. Methods for coordinating human factors concerns and considerations among integrated product team members and contractor personnel are established (2.3.1).

3. The processes and procedures for how the government will direct, control, and monitor the human factors efforts are described (2.3.1, 2.4.1, 2.5.1).

4. The operators and maintainers of the system are described in plan (2.3.1).

5. The user functions/tasks are described in detail (2.3.1).

6. The system objectives for personnel resources, training, workload, ergónomics, and safety are identified (2.4.1).

7. Key design goals are operationally defined (i.e., described in terms of how they will be measured and evaluated) (2.3.1).

8. Parameters to be used as criteria against which the system will be evaluated are identified (2.3.1, 2.4.1).

9. The tasks and analyses that need to be conducted to support the definition and evaluation of system performance requirements are specified (2.4.1).

10. The system constraints on personnel resources, training, ergonomics, and safety are described (2.5.1)

11. Critical known issues and work to be done to address system performance requirements are identified (2.4.1, 2.5.1).

12. Critical "unknowns" are listed (to be answered/assessed as more information becomes available) (2.3.1, 2.5.1).

13. A feasible schedule is proposed for accomplishing the human factors work (2.5.1).

CHAPTER 3. VISUAL PERCEPTION

John S. Werner

> "Vision appears to be an immediate, effortless event. To see the surrounding environment, it seems we need only open our eyes and look around. However, this subjective feeling disguises the immense sophistication of the human...visual system and the great complexity of the information-processing tasks it is able to perform in an apparently effortless manner."
>
> Yuille and Ullman (1990)

TABLE OF CONTENTS

CHAPTER 3. VISUAL PERCEPTION

LIST OF FIGURES AND TABLES

CHAPTER 3. VISUAL PERCEPTION

The purpose of this chapter is to provide some basic information about vision, our primary sensory channel. This information is meant to increase understanding and assist in documenting the perceptual basis for the guidelines presented for visual displays. It will provide the basis for understanding the properties of visual displays that can be helpful to the controller and those that will be operationally unacceptable.

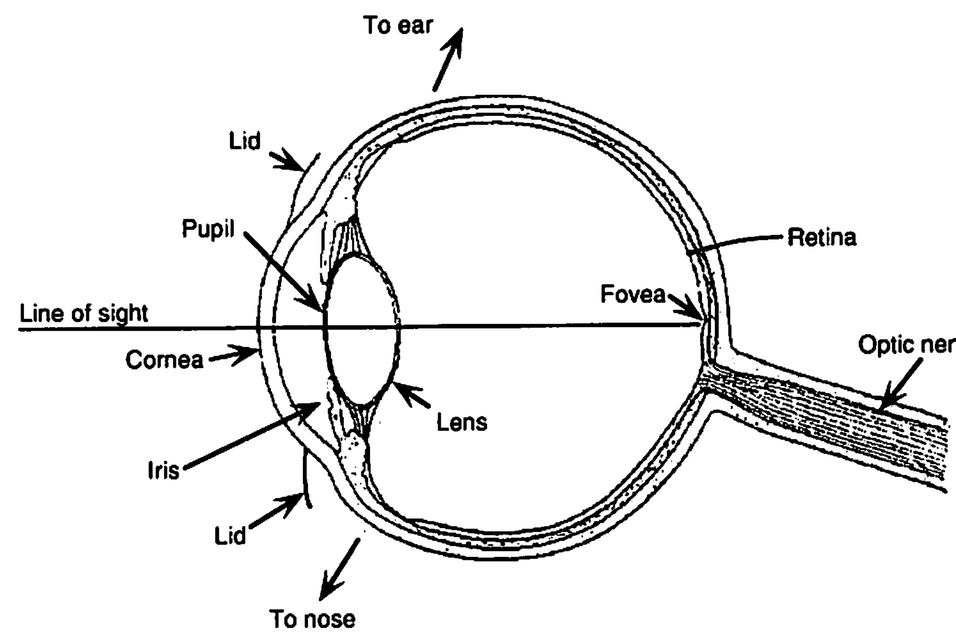
3.1 BASICS OF VISUAL PERCEPTION

Some might argue that controllers serving on acquisition teams do not need to know a great deal about the basic visual processes, how the eye works, and so forth. On the other hand, knowing something about these topics can be helpful in critiquing visual display designs. Accordingly, this section takes the reader on a short tour of key topics that are related to the design and evaluation of visual displays. Such displays should be designed for compatibility with the human visual system.

3.1.1 What are the major features of the eye, and what are their functions?

Figure 3-1 is a diagram of the human **eye**. The eyeball is surrounded by a tough white tissue called the **sclera**, which becomes the clear **cornea** at the front. As light travels through space in waves of different lengths, it passes through the cornea and continues on through the **pupil**. The pupil is a hole formed by a ring of muscles called the **iris**. It is the outer layer of the iris that gives our eyes their color. Contraction and expansion of the iris opens or closes the pupil to adjust the amount of light entering the eye. Light then passes through the lens and strikes the **retina**. The retina includes the receptors that convert light energy into nerve signals that the brain can interpret. One part of the retina, called the **fovea**, contains the highest density of receptors. When we look at an object, we move our head and eyes so that the reflected light representing the image of interest will fall on the fovea.

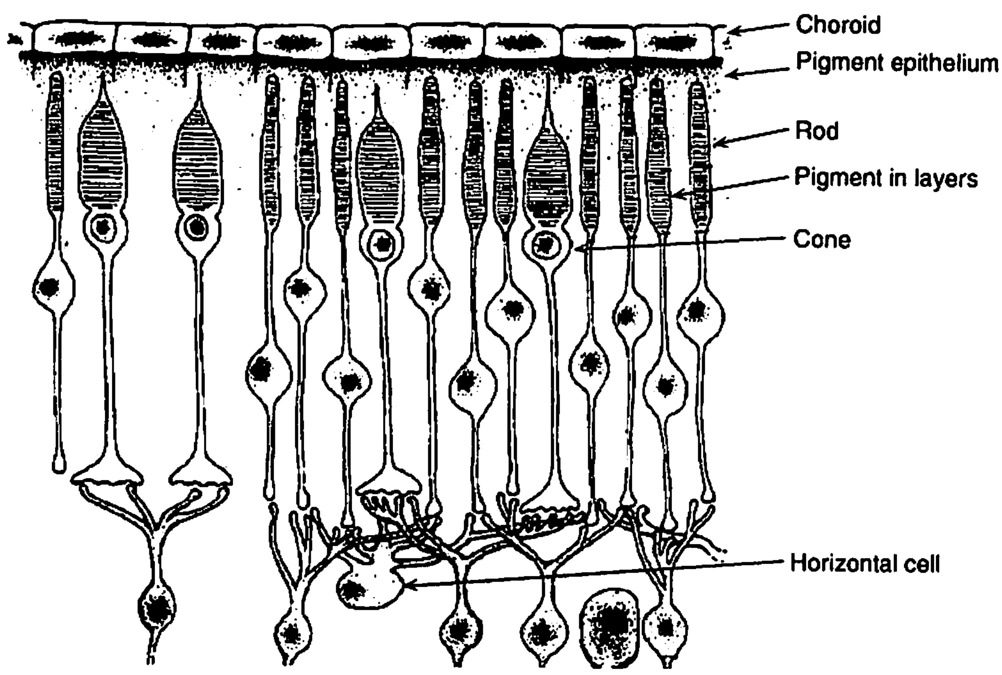
Figure 3-1. Major Features of the Human Eye
From *Visual Perception* by T. Cornsweet, 1970. Copyright 1970 by Academic Press. Reprinted by permission of Harcourt Brace & Company.



At any instant, the eye can focus on objects only if the objects fall within a limited range of distance. The human eye uses a somewhat flexible **lens** in the eye to adjust the focus of light entering the eye. When the shape of the lens is changed, the light passing through it will be bent at a different angle. This process is known as **accommodation.** The response time required for accommodation is about 360 milliseconds (Campbell and Westheimer, 1960). This time is long enough to produce a noticeable blur when shifting focus. It will be most noticeable when shifting focus from a distant object to a display panel as the lens of the eye will have to bend significantly to adjust to the change in distance. The combination of a single eye movement and the accompanying accommodation of the lens of the eye can take about one-half second.

The two types of receptors in the retina are called **rods** and **cones.** There are approximately 6 million cones and 125 million rods in the human retina. After passing through the outer portions of the eye, light energy strikes these receptors and causes chemical changes to occur in them, which in turn causes them to send nerve impulses to the brain and other receptors. Figure 3-2 illustrates the various cell types located in the retina.

Figure 3-2. Diagram of Rods and Cones
From *Sensation and Perception* by Coren, Porec, and Ward, 1984, p. 64. Copyright 1984 by Academic Press. Used by permission of Harcourt Brace & Company.



Rods and cones have special features that aid our vision under various lighting conditions. Cones are relatively insensitive to changes in light intensity, are the primary receptors for vision during the day, and are responsible for color vision. Cones are most densely packed in the fovea and provide us with our best visual acuity, or the ability to see fine details. Outside the fovea, where the density of the cones decreases, there is a corresponding decrease in visual acuity. The density of rods is greatest about 20 degrees from the fovea and decreases toward the periphery. The periphery has many more rods than cones, but a careful reading of the figure shows that there are as many as 7,500 cones per square mm even in the peripheral retina.

Rods are very sensitive to small changes in light intensity and are the primary receptors for night vision. Rods are most densely packed towards the edges of the retina. In fact, the periphery of the retina has many more rods than cones. Because rods are unable to detect colors, vision with rods (as in night vision) is very similar to viewing a black and white television. Colors are interpreted by rods as different shades of gray or levels of brightness.

3.1.2 What is the importance of visual angle and how is it measured?

Visual angle is the area of the retina that is covered or subtended by the visual image of a physical object. Visual angle is measured in degrees. The size of a visual angle depends on the object's size and distance from the observer.

Consider what happens when we look at an object, say an airport control tower. Imagine the tower as many points of light, and imagine that we are looking at the light coming from the top of the tower. When we focus on the tower, our cornea and lens bend the light so that an image is formed at the back of the eye, on the retina. The optics of the eye bend the light so that the image of the tower on the retina is upside down and reversed left to right.

Smaller and smaller objects at closer and closer distances could all subtend similar visual angles. For example, a radar facility, an airport terminal, a Cessna, and a baggage cart could all cover approximately the same area of the retina, depending on our distance from them. At differing distances, two objects of the same size will produce different visual angles. As the distance between an object and the cornea doubles, the size of the image produced by the object is halved.

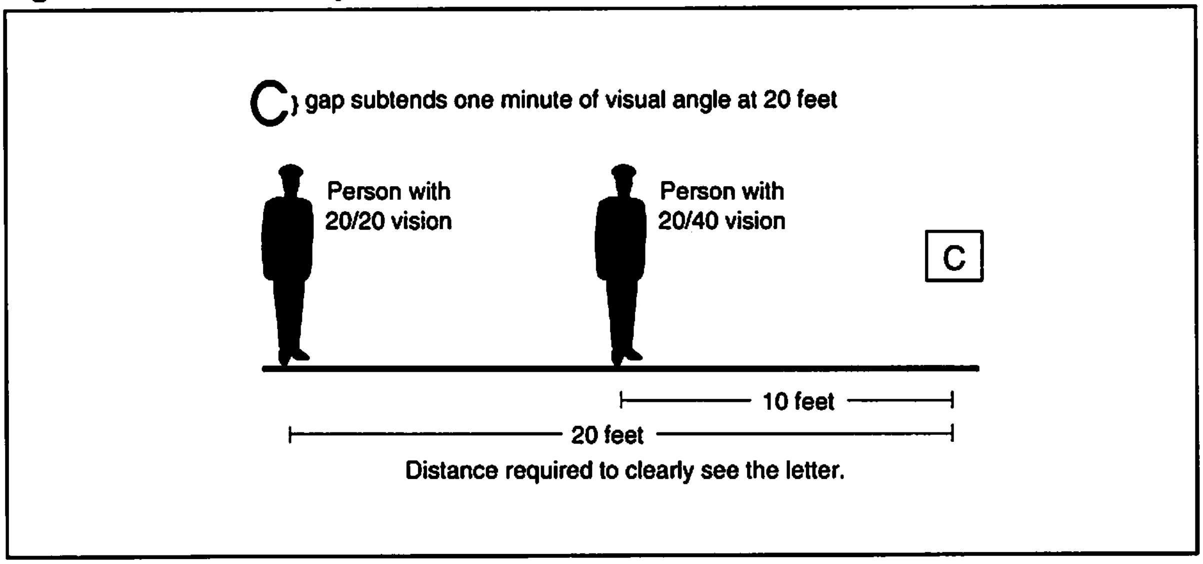
The number of degrees in visual angle x is calculated by: arctan (size/distance). By definition, one degree of visual angle equals 60 minutes of arc, and one minute of arc equals 60 seconds of arc. For purposes of illustration, the visual angle x of your thumb nail at arm's length is about two degrees (2°).

Visual angle is an important human-factors consideration in the design of visual displays and workstations. For example, knowing the visual angle of text (given a certain level of contrast), can help predict the accuracy with which controllers can read the text. We will return to it in the discussion of recommended sizes for symbols and alphanumerics on visual displays (Chapter 7, Computer-Human Interface [CHI] Considerations). Visual angle will also figure in the discussion of recommended distance between visual display terminals (VDTs) and seated controllers (Chapter 9, Workstation and Facility Design and Evaluation).

3.1.3 What is visual acuity and how is it measured?

Visual acuity is defined in terms of the smallest detail that an observer can see. This is measured by the familiar eye chart found in an optometrist's office. At a distance of 20 feet, an individual with 20/20 vision can clearly see a letter that has a gap that subtends a visual angle of one minute. An acuity of 20/40 means that, at a distance of 20 feet, an individual can clearly see the gap that would have a visual angle of one minute at a distance of 40 feet. Thus, a person with 20/40 vision has poorer visual acuity and requires either a larger target or a closer reading distance than does a viewer with 20/20 acuity. This is illustrated in Figure 3-3. In many states, a person is legally blind if acuity is 20/400 or worse.

Figure 3-3. Visual Acuity: Measurement

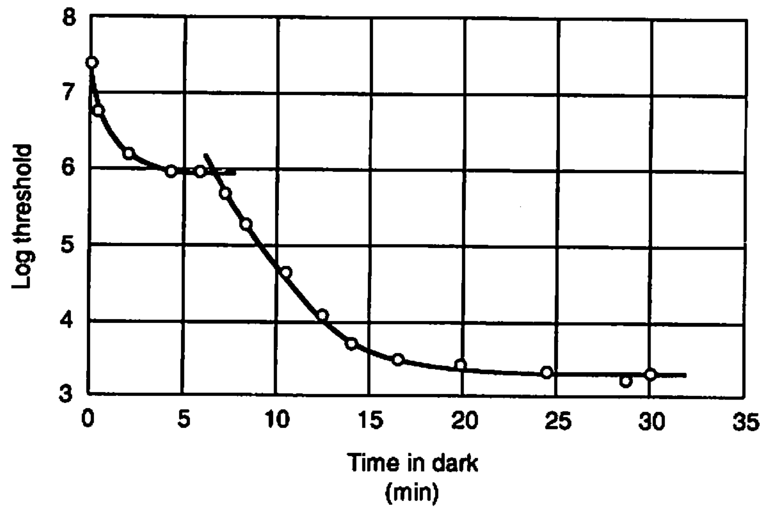


3.1.4 How does visual acuity change with ambient light levels?

Most of us have groped around in a dark movie theater until our eyes have adjusted to the dim level of illumination. This process is called **dark adaptation**, and it occurs, in part, because our receptors need time to achieve their maximum sensitivity (i.e., minimum threshold). If we were to measure the minimum amount of light required to see after we enter a darkened room, we would be measuring our threshold as a function of time in the dark. The resulting dark adaptation curve would look like the one in Figure 3-4. This curve indicates that the eye becomes progressively more sensitive in the dark.

Notice that the curve shown in Figure 3-4 has two distinct phases. The first phase, which lasts about seven minutes, is attributed to the cone system; the second phase is attributed to the rod system. When we first enter the dark, our cones are more sensitive than the rods, and more light is required to see anything. After about seven minutes, the rods are more sensitive than the cones, and we are beginning to adapt to the dark. It takes about 20 minutes for us to fully adapt to the dark.

Figure 3-4. Minimum Amount of Light Needed to See in the Dark
From *Vision and Visual Perception*, edited by C. H. Graham, 1965, p. 75. Copyright 1965 by John Wiley & Sons. Reprinted by permission.



Visual acuity also varies with luminance when ambient illumination is high. We have seen that in the scotopic range (vision in the dark), visual acuity depends on rods and is very poor. As light intensity increases, visual acuity depends more on cones and improves dramatically. Even after cones "take over," visual acuity continues to vary with light intensity. When a stimulus is moving or the display is vibrating (as in turbulence), visual acuity may be considerably reduced.

3.1.5 What is contrast sensitivity, and why is it important for form vision?

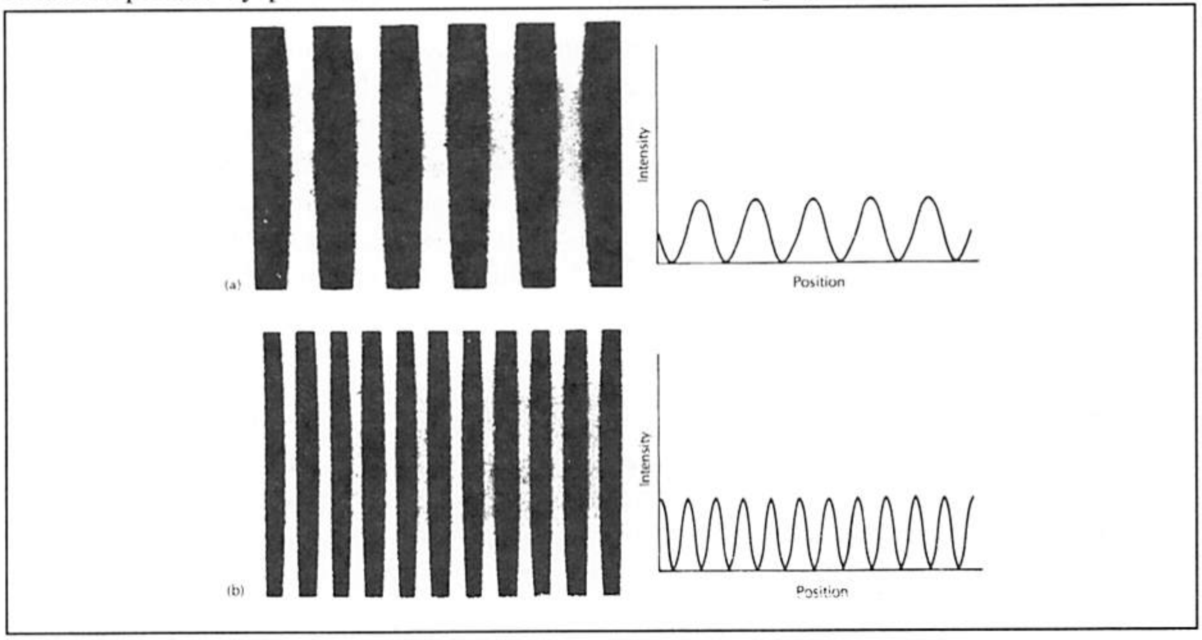
Even an individual who has 20/20 acuity may not have perfect vision. While measures of visual acuity indicate one's ability to resolve fine detail, they tell us very little about the ability to distinguish larger forms, such as faces.

Whether the object is an aircraft ID on a scope or a truck on an airport taxiway, the forms of all objects are defined by contrast. It is, therefore, important to characterize the sensitivity of the visual system to contrast. One approach is to measure contrast sensitivity using grating stimuli in which the luminance is varied sinusoidally, as illustrated by Figure 3-5. If one were to measure the intensity of the stimuli on the left, by passing a light meter across it, the sinusoidal luminance profile on the right would be found. The profile of the stimuli could be characterized by the contrast, which was defined above by the difference between the luminance maximum and minimum, divided by the average luminance.

The frequency of oscillation of the sine wave is defined in terms of the number of cycles per degree of visual angle (cpd). For example, the stimulus on the top of Figure 3-5 has a lower spatial frequency than the one on the bottom.

Figure 3-5. Differences in Contrast Resulting from Low and High Spatial Distributions of Luminance

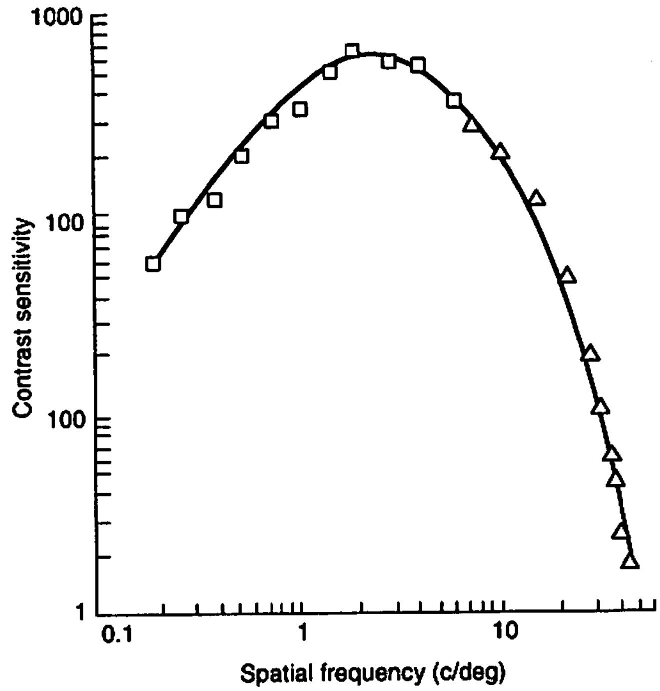
From *Visual Perception* by T. Cornsweet, 1970, p. 314. Copyright 1970 by Academic Press. Reprinted by permission of Harcourt Brace & Company.



Contrast threshold is measured by determining the minimum contrast required for detection of a grating having a particular spatial frequency (usually generated on a CRT display). For example, it represents how much contrast is required for the observer to be able to see a bar pattern, rather than a uniform gray field. Contrast sensitivity is the reciprocal of contrast threshold. Thus, the **contrast sensitivity function** represents the sensitivity of an individual to sine-wave gratings plotted as a function of their spatial frequency. Figure 3-6 shows a typical contrast sensitivity function (Campbell and Robson, 1968). These data were obtained with a set of static sine-wave gratings (like those in Figure 3-5), but contrast sensitivity functions vary as a function of luminance, temporal characteristics of the grating stimuli (e.g., flickering versus steady), and stimulus motion characteristics (e.g., drifting versus stationary gratings). The shape of the contrast sensitivity function also varies with the individual observer and the orientation of the grating. For example, many individuals are more sensitive to vertical and horizontal gratings of high spatial frequency than to oblique (45° or 135° from horizontal) gratings (Appelle, 1972).

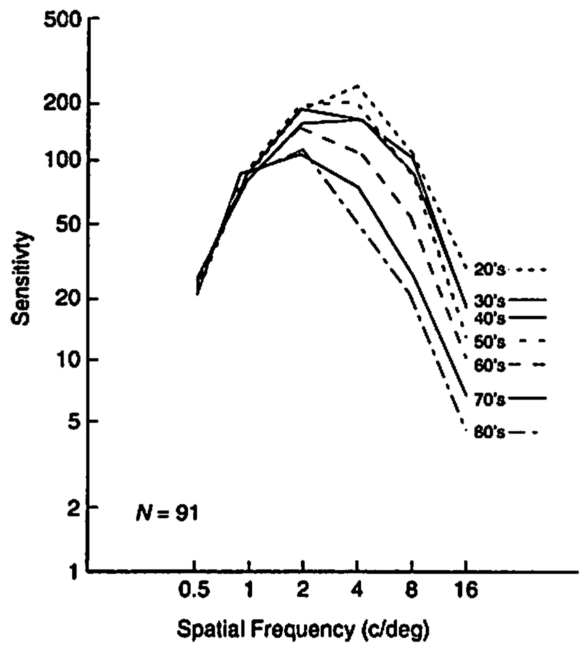
It can be deduced from the contrast sensitivity function that we are not equally sensitive to the contrast of objects of all sizes. High spatial frequency sensitivity is related to visual acuity; both are a measure of resolution, or the finest detail that can be seen. When spatial vision is measured by an optometrist or ophthalmologist, only visual acuity is typically measured. While a more complete evaluation of spatial vision would include contrast sensitivity measurements over a range of spatial frequencies, it is the high frequency sensitivity that is most impaired by optical blur (Westheimer, 1964). Thus, high frequency sensitivity is what is improved by spectacle corrections.

Figure 3-6. Contrast Sensitivity Function
From "Application of Fourier Analysis to the Visibility of Grating" by Campbell and Robson in *Journal of Physiology, 197*, 1968, p. 136. Copyright 1968 by the Physiological Society (Oxford). Reprinted by permission.



Average contrast sensitivity for various spatial frequencies are shown in Figure 3-7 plotted as a function of age. These data represent averages from 91 clinically normal, refracted observers (Owsley, Sekuler, and Siemsen, 1983). Age-related declines in contrast sensitivity, like declines related to decreased retinal illuminance, are most pronounced at high spatial frequencies. This is partially due to the fact that the lens transmits less light and the pupil is smaller in elderly observers.

Figure 3-7. Contrast Sensitivity as a Function of Spatial Frequency for Different Age Groups
From "Contrast Sensitivity Throughout Adulthood" by C. Owsley, R. Sekuler and D. Siemson, 1983, p. 694, *Vision Research, 23.* Copyright 1983 by Pergamon Press, Inc. Reprinted by permission.



The contrast sensitivity function has several areas of application. First, as a predictor of visual performance, the contrast sensitivity function may be more useful than traditional measures of visual acuity. The visual acuity chart varies only the size of the stimuli to evaluate spatial vision while contrast sensitivity testing requires variation in both size and contrast. The importance of this additional information was illustrated by researchers who conducted an experiment with experienced pilots and an aircraft simulator (Ginsburg, Evans, Sekular, and Harp, 1982). The simulated visibility was poor and half of the simulated landings had to be aborted due to an obstacle placed on the runway. Performance was measured by how close the pilots flew to the obstacle before aborting the landing. Pilot responses on this task (times required to abort the landing) varied considerably. Individual differences in performance were not well correlated with visual acuity but were well predicted by individual variation in contrast sensitivity. Thus, contrast sensitivity testing may be more useful than traditional measures of visual performance for predicting responses in complex settings, including ATC.

A second application of the contrast sensitivity function is for predicting the visibility of complex patterns presented on displays. It may not be feasible to test every unit of symbology directly, but knowing the contrast sensitivity function, it may be possible to make some predictions using a spatial frequency analysis of the stimulus. For example, if we know the mark-up of a particular symbol and its contrast sensitivity, we can predict which features will be easier to see from a given distance.

3.1.6 What is flicker and what affects our ability to detect it?

If a light is turned on and off in rapid succession, we will experience a sensation we call flicker. If the frequency of oscillations, measured in cycles per second (cps or Hz) is high enough, the flicker will no longer be perceptible. This explains how fluorescent lamps appear to be steady even though they are going on and off at 120 Hz. The frequency of oscillations at which flicker is no longer perceptible is know as the **critical flicker fusion (CFF)** frequency. CFF will differ among observers and will differ for the same observer in different situations.

Several factors affect our ability to detect flicker. One such factor is the light level. As luminance increases, flicker is easier to detect. At high light levels, CFF may occur at 60 Hz for some observers.

Whether or not a display will appear to flicker may also depend on whether we are viewing the display straight on (i.e., centrally) or peripherally. A display that appears steady when you look directly at it may appear to flicker when viewed out of the corner of your eye. This is because the cells that predominate in the periphery of the retina (i.e., the rods) are more sensitive to flicker than the cone cells in and near the fovea.

The age of the observer also affects one's ability to detect flicker. Flicker sensitivity declines rather markedly as a function of increasing age. This is to be expected, at least in part, because the light transmitted by the lens decreases with age and flicker sensitivity is dependent on light level. However, it is still not clear whether these changes in flicker sensitivity are secondary to changes in light level alone or whether there are additional neural changes that contribute to the age-related change in CFF (Weale, 1982). In any event, this sensitivity to flicker as a function of age means that a display may appear steady to one observer, while a younger observer may see the same display as flickering.

For these reasons, it is not wise for controllers to evaluate flicker solely by looking at the display. **In order to avoid the appearance of flicker for all observers under all light levels, displays should have a refresh rate of at least 65 Hz.**

3.1.7 In what other ways does age affect vision?

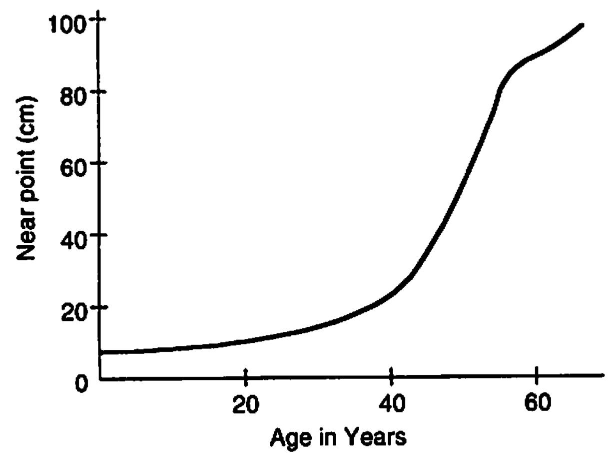
Aging affects visual capabilities in several ways. As we get older, the flexibility of our lenses decreases. This loss of flexibility increases the time required to focus on near objects after we have been looking at distant objects. Corrective lenses can counteract some of these effects. In addition to losing flexibility, the lens further yellows with age, decreasing its ability to transmit light and its ability to detect light at short wavelengths (in the blue range).

Aging and Loss of Accommodation

The flexibility of the eye lens decreases with age and thereby limits the ability to accommodate, both in terms of the amount of change in the lens and the time required to respond to changes that occur when shifting fixation from far to near objects (Weale, 1982). The loss in accommodative ability, known as **presbyopia,** is often quantified in terms of the near point, or the closest distance at which an object can be seen without blur. As illustrated by Figure 3-8, the near point increases with advancing age. By about age 40, the near point is such that reading can only be accomplished if the print is held at some distance or if reading glasses are used.

Some individuals require one lens correction for their distance vision and a different correction for their presbyopia. This can be accomplished by bifocal lenses—lenses which require the individual to look through different parts in order to properly focus near and far objects.

Figure 3-8. Relationship Between Age and Near Point Focus
From "Physiological Effects of Aging" by E. P. W. Helps in *Proceedings of the Royal Society of Medicine, 66*, 1973, pp. 815-818. Copyright 1973 by the Royal Society of Medicine. Reprinted by permission.



Optical Transmission of Light and Aging

The various optical components of the eye—the **ocular media**—shown in Figure 3-1 are not completely transparent. The lens of the eye, in particular, has a yellowish color. It absorbs quite strongly at the short wavelengths of the visible spectrum (around 400 to 450 nm) and even more strongly in the ultraviolet portion of the spectrum from 300 to 400 nm. The ability of the lens to transmit light decreases markedly with advancing age. The average 70-year-old eye transmits about 22 times less light at 400 nm than does the eye of the average 1-month-old infant. This difference between young and old diminishes with increasing wavelength.

Because the lens increases its absorption with age, the visual stimulus arriving at the receptors will be less intense with age. In addition, for stimuli with a broad spectrum of wavelengths, there will be a change in the relative distribution of light energy because the short wavelengths will be attenuated more than will the middle or long wavelengths. Since the stimulus at the retina is changing with age, **there will be age-related decreases in the ability to detect short wavelengths of light.** The amount of light absorbed by the lens will also directly influence our ability to discriminate short wavelengths (blue hues). Thus, the large range of individual variation in the lens leads to large individual differences in discrimination of blue hues and in how a specific blue light or symbol will appear to different observers. The legibility of blue symbols on a display could easily differ for observers with normal color vision, depending on their ages.

While an increase in the absorption of light with advancing age is considered normal, some individuals experience an excessive change which leads to a lens opacity known as a **cataract**. A cataractous lens severely impairs vision and is typically treated by surgical removal and implantation of a plastic, artificial lens. These artificial lenses eliminate the ability to accommodate, but in most cases of cataract, the individual is more than 55 or 60 years of age and has lost this ability anyway.

3.1.8 What is depth perception, and what factors affect a controller's ability to perceive depth?

The ability to perceive the positions of objects in space is called depth perception. There are two major classes of cues that we use to perceive depth. **Monocular depth cues** provide information about depth that can be extracted using only one eye. **Binocular depth cues** rely on an analysis of slightly different information available from each of the two eyes.

Monocular Depth Cues

If you close one eye and look around, you will probably not be confused about the relative distances of most objects. Your perception of distance in this case is based on monocular cues, which are even more powerful than some of the binocular cues to depth (Kaufman, 1974).

The **size** of objects can sometimes indicate their relative depth. If several similar items are presented together, the larger items will be judged as closer. This makes sense because, in fact, the size of an object's image on the retina becomes progressively smaller as it moves away.

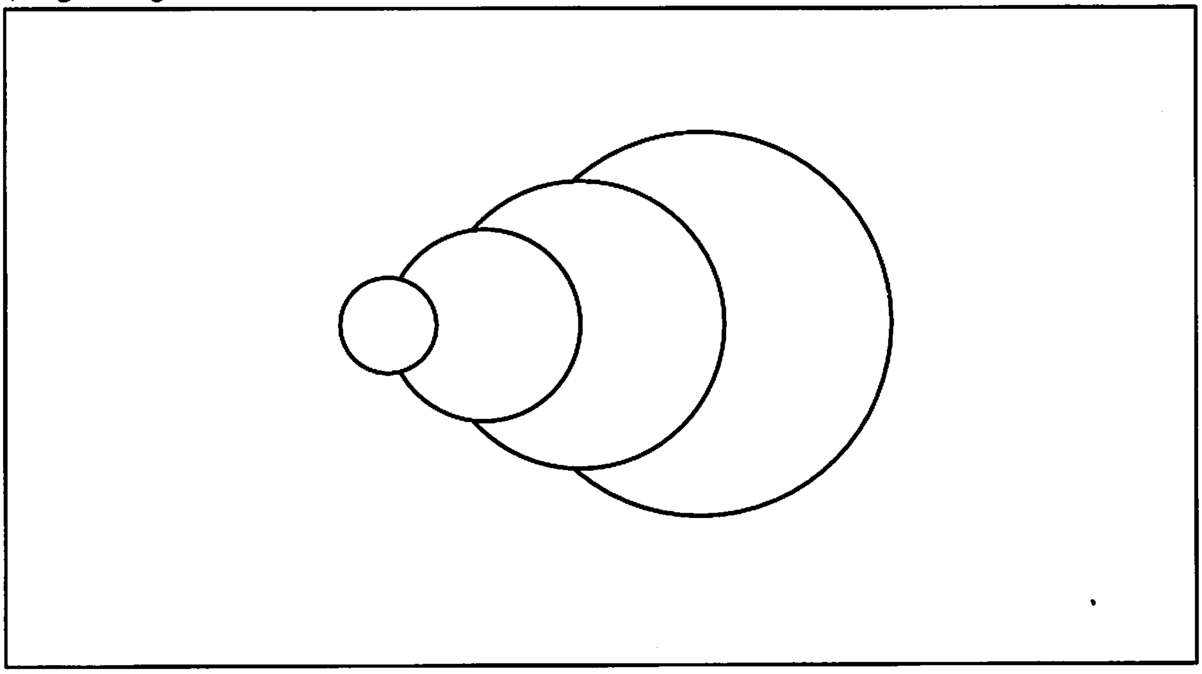
The ability to infer distance from image size often depends on familiarity with the true size of the objects. At great distances, such as looking down from an airplane, we perceive objects to be smaller than when they are near. In this situation, our familiarity with objects and their constancy of size serve as a source of information about distance. Although from the air a house seems like a toy, our knowledge about the actual size of houses informs us that the house is only farther away, not smaller.

The relation between size and distance can lead not only to faulty inferences about distance but assumptions about distance can also lead to faulty inferences about size. When we are misinformed about distance, our perceptions of size and shape will be affected. You have probably noticed, for example, how much larger the moon appears when it is low on the horizon than high in the evening sky. This is called the **moon illusion.** The change in the moon's appearance is only slightly affected by atmospheric phenomena; by far the greatest effect is perceptual. Our retinal image of the moon is the same size in both positions. You can prove this by holding at arms length a piece of cardboard just large enough to block the moon from view. The same piece of cardboard blocks the moon at the horizon and at its zenith equally. Though they look different, they measure the same. The moon illusion seems to be caused by inaccurate distance information about very far objects (Kaufman and Rock, 1962). Because we see intervening objects on the earth's surface when we look at the moon near the horizon, our internal distance analyzers apparently cue us that the moon is farther away than when it is at its zenith. An object analyzed as more distant has to be larger to produce an image of the same size. Thus we perceive the moon as larger on the horizon than when it is at its zenith.

The relationship between size and distance is important to understanding not only harmless illusions, such as the size of the moon, but also in situations of more significance. As mentioned above, changing fixation from a head-up display (HUD) to distant objects often requires a change in the state of accommodation. Change in the focus of the eye is accompanied by a change in the apparent visual angle of distant objects. Thus, when a pilot shifts fixation from a HUD to a distant surface in the outside world, the objects in the distance may appear smaller and more distant than they really are (Iavecchia, Iavecchia, and Roscoe, 1988). While the resultant spatial errors in perception are temporary, researchers believe that such errors could introduce a significant safety hazard under some conditions (Iavecchia et al. 1988).

Any ambiguity about relative distance in relation to size can be rectified when one object partially occludes another, as shown in Figure 3-9. We perceive the partially occluded object as being more distant. This cue to depth is called **interposition.**

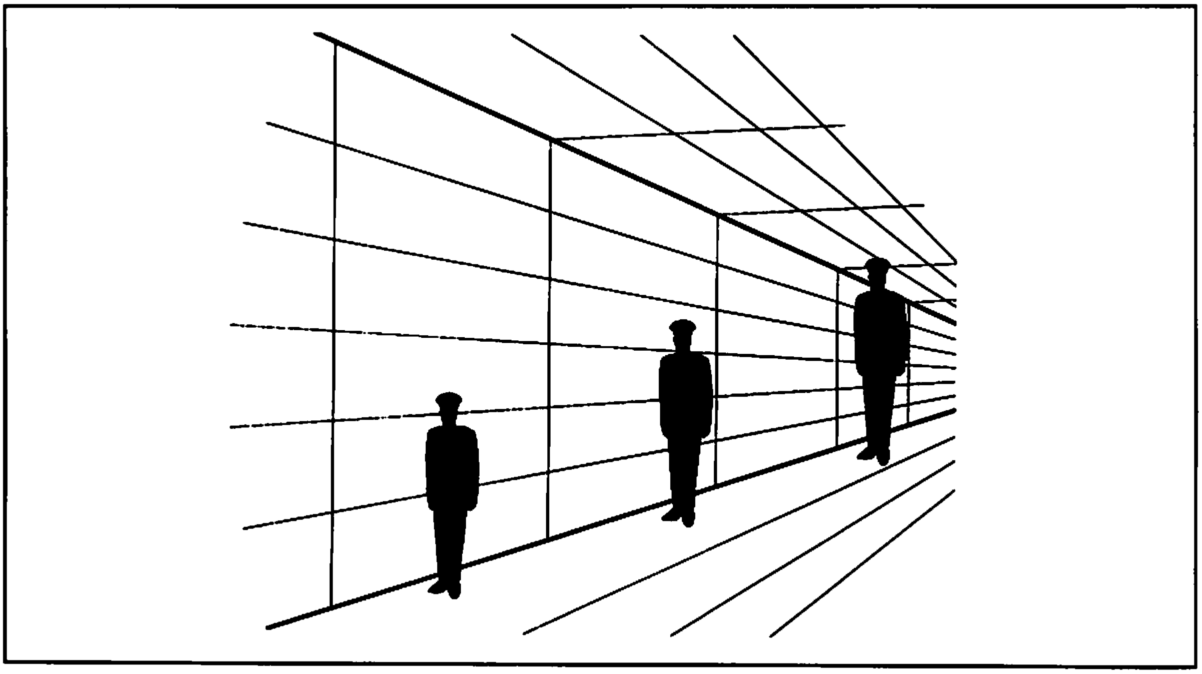
Figure 3-9. Illustration of Interposition as a Monocular Cue for Distance
(Original figure)



If a distant object is not partially occluded, we may still be able to judge its distance using **linear perspective.** When you look at a set of parallel lines, such as railroad tracks going off into the distance, the retinal images of these lines converge because the visual angle formed by two points parallel to another decreases as the points are farther away.

This cue to depth is so powerful that it may cause objects of the same size to be perceived as different, as in Figure 3-10. If you look at a textured surface, such as a lawn, two blades of grass the same distance apart would be separated by a smaller distance in the retinal image the farther away they are because they cover a smaller visual angle. Most surfaces have a certain pattern, grain, or **texture,** such as pebbles on the beach or the grain of a wood floor. Whatever the texture, it becomes denser with distance. This information can provide clear indications of distances (Newman, Whinham, and MacRae, 1973).

Figure 3-10. Illustration of How Linear Perspective Makes the Same Size Objects Appear to be Different Sizes
Adapted from *Perception* by R. Sekular and R. Blake, 1985, p. 238. Copyright 1985 by Knopf. Adapted by permission of McGraw-Hill.



Of special relevance in aviation is the depth cue known as aerial perspective. As light travels through the atmosphere, it is scattered by molecules in the air such as dust and water. The images of more distant objects are thus less clear. Under different atmospheric conditions, the perceived distance of an object of fixed size may vary. For example, an airport will appear farther away to a pilot on a hazy day than on a clear day. Novice and student pilots are likely to misjudge their distance to the airport on hazy days. Similarly, to a tower controller, aircraft may appear further away on a hazy day than on a clear day.

Some monocular cues to depth are not static, but depend on relative movement. When we are moving, objects appear to move in relation to the point of fixation. The direction and speed of movement is related to their relative distances. Objects that are more distant than the point of fixation appear to move in the same direction as the observer. Objects in front of the point of fixation appear to move opposite to the direction of the observer. You can demonstrate this by holding two fingers in front of you at different distances and then observing their relative displacement as you move your head back and forth. The difference in how near and far objects move, called **motion parallax**, is probably our most important monocular source of information about distance. Motion parallax occurs from any relative motion–moving the whole body, the head, or the eyes.

Motion perspective is a phenomenon related to motion parallax. It refers to the fact that as we move straight ahead, the images of objects surrounding the point of fixation tend to flow away from that point. Imagine walking through the stacks of books in a library. If the observer were to back up, the flow pattern would contract rather than expand. These optic flow patterns carry information about direction, distance and speed, and are believed to be an important depth cue used by pilots to land planes (Regan, Beverly, and Cynader, 1979).

Stereopsis

Because the two eyes are separated by about 3 inches, the visual fields are slightly different for the two eyes. In the region where the two eyes have overlapping visual fields, they will receive slightly different images of objects. This is easily verified by placing your finger six inches to a foot away from your face and then looking at your finger's position with reference to the background scene with one eye and then the other. Your finger will appear to be in two different positions. This happens because the two eyes have different angles of view. The ability to judge depth using this **retinal disparity** is known as **stereopsis.**

There are many implications of this ability. For example, there are several ways to demonstrate stereoscopic depth from two-dimensional images. It has been shown that if one image is presented to one eye and another image to the other eye through a stereoscope, the images can be fused and a three-dimensional image can be seen (Wheatstone, 1838). Today, 3-D movies are created by projecting two (disparate) images on a screen. Separation of the images is made possible by projecting them with polarized light of orthogonal orientations. If the viewer has polarizing glasses, the two images will be separately projected to each retina, fused, and perceived as three dimensional. Other technologies make three-dimensional imaging possible without special glasses. While this technology is promising, much more work is needed before three-dimensional CRT displays are practical for air traffic control.

There are also other practical applications. For example, if a counterfeit dollar bill is placed on one side of a stereoscopic viewer and a genuine dollar bill on the other, the two can be compared and differences of 0.005 mm can be detected because they will stand out in depth. Other virtues of stereovision are well known to aerial surveyors and experts in aerial surveillance. Under optimal conditions, stereoscopic depth can be used to resolve displacement in depth of about 2 sec of arc. This corresponds to a difference that is smaller than the diameter of a single cone receptor.

Stereoacuity varies with the distance of the object. Beyond about 100 feet, retinal disparity diminishes so greatly that this cue to depth is not useful. Thus, it is sometimes noted that routine aspects of flying an airplane do not require stereopsis, but it is helpful when moving the plane into the hangar (DeHaan, 1982). Similarly, when a tower controller is judging the distance of an aircraft to the airport, he or she is using depth cues other than stereopsis.

Binocular Rivalry

If the scenes presented to each eye are very different, such as a vertical grating presented to the right eye and a horizontal grating presented to the left eye, the visual system does not fuse the images. Rather, views of the two scenes may alternate from one eye to the other or a mosaic that combines portions of the two images may alternate. This is known as **binocular rivalry** and can occur whenever the images presented to each eye are too different to be combined. Apparently, the visual system attempts to match the images from the two eyes and when this cannot be done, one of the images or at least portions of one image are suppressed.

During a person's early life, the images to the two eyes may be chronically discordant due to the two eyes being improperly aligned, a condition known as **strabismus.** If this condition is not corrected in early childhood, the input from one of the eyes may become permanently suppressed and the individual will be **stereoblind,** that is, incapable of using stereoscopic cues to depth. Whether due to strabismus or other causes, about 5 to 10 percent of the population is stereoblind (Richards, 1970).

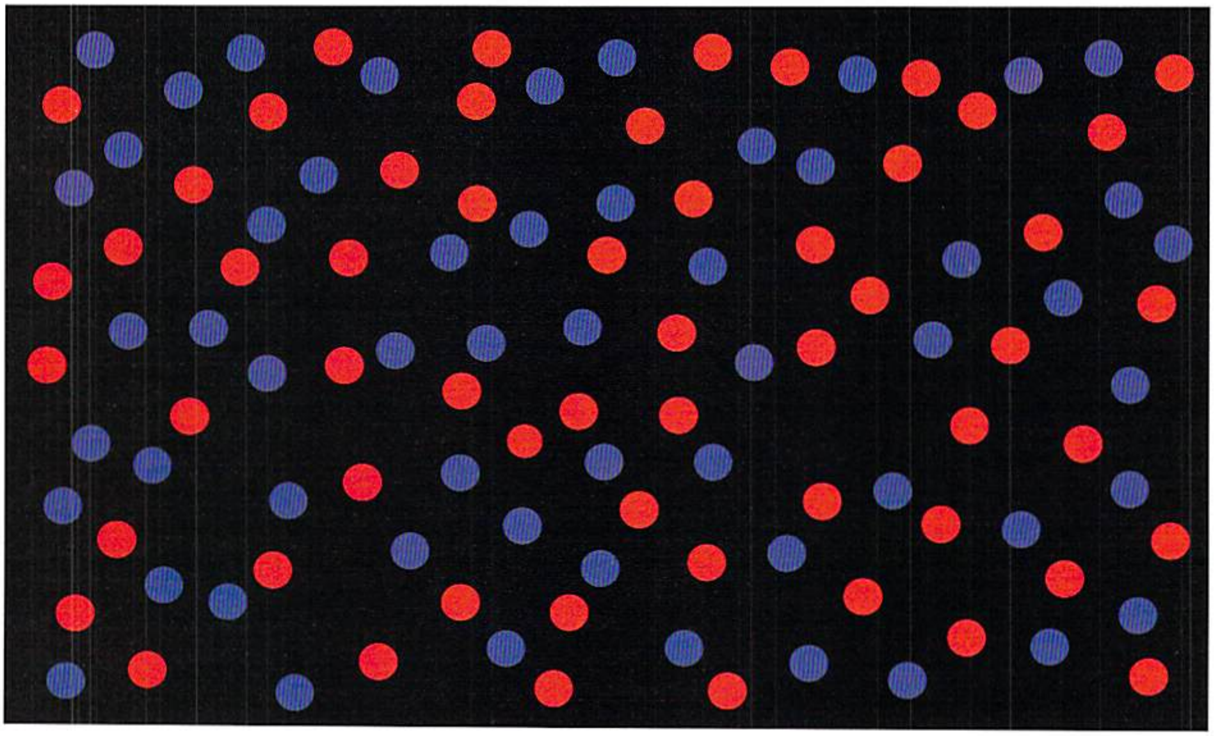
Color Stereopsis

When deeply saturated colors are viewed on a display, it sometimes appears that the different colors lie at different depths. This phenomenon, known as color stereopsis or chromostereopsis, is illustrated in Figure 3-11. The effect is most clearly seen with colors that are maximally separated in the spectrum. On displays, red may appear to be nearer than blue. Color stereopsis is due to retinal disparity arising from chromatic dispersion by the optics of the eye. Short wavelengths are imaged more nasally than long wavelengths and the resultant retinal disparity leads to the perception that the different colors are at different depth planes. Display operators can minimize this effect, if necessary, by using less saturated colors or brighter backgrounds (Walraven, 1985).

Implications for Displays

Stereopsis provides a little used channel for presenting information on visual displays. By using retinally disparate images, it is possible to create more realistic portrayals of the external environment than would be possible on displays carrying only monocular information. Applications of stereo imagery to controller displays are being discussed, but are far from serious consideration.

Figure 3-11. An Illustration of Color Stereopsis
(Original figure)



3.2 COLOR PERCEPTION

Color can significantly enhance search and identification of information on visual displays. It is more effective than shape or size in helping to locate information quickly (Christ, 1975). The attention-getting nature of color facilitates search while at the same time providing a good basis for grouping or organizing information on a display which may help display operators segregate multiple types of information and reduce clutter. For example, in one experiment, when the number of display items was increased from 30 to 60, search time increased by 108 percent when only one color was used, but increased by only 17 percent for redundant color-coded displays (Carter, 1979).

There are severe constraints on the effective usage of color information. The attention-getting value of a color is dependent on its being used sparingly. Only a limited number of colors should be used in order to avoid overtaxing the ability of an observer to classify colors. If each color is to have meaning, only about six or seven can be utilized effectively.

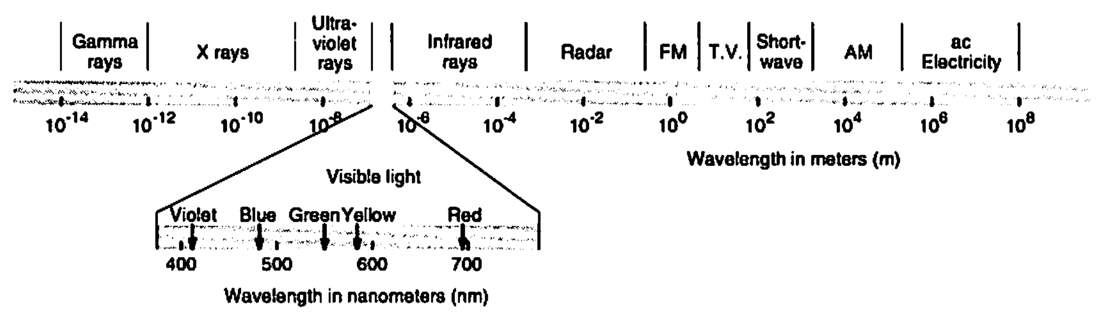
3.2.1 How do we perceive color?

As with sound energy, the movement of light energy through space is in a sinusoidal pattern. Sound waves were described in terms of their frequency, but light waves are more commonly described in terms of the length of the waves (i.e., the distance between two successive peaks). This description is equivalent to one based on frequency because wavelength and frequency are inversely related. As seen in Figure 3-12, the electromagnetic spectrum encompasses a wide range, but our eyes are sensitive only to a small band of radiation which we perceive as light. Normally, we can see light with wavelengths between about 400 and 700 **nanometers** (nm; 1 nm is one billionth of a meter). Most naturally occurring light sources emit many wavelengths (or a broadband of the spectrum), but in a laboratory, we use specialized instruments that emit only a narrow band of the spectrum called **monochromatic** lights. If a person with normal color vision were to view monochromatic lights in a dark room, the appearance would be violet at 400 nm, blue at 470 nm, green at 550 nm, yellow at 570 nm, and red at about 680 nm. Note that this description is for one set of conditions. The appearance can change for the same monochromatic lights when viewed under other conditions.

Color is defined by three properties: brightness, hue, and saturation. It would be convenient for engineers if these three psychological properties were related in one-to-one correspondence to physical properties of light, but they are not.

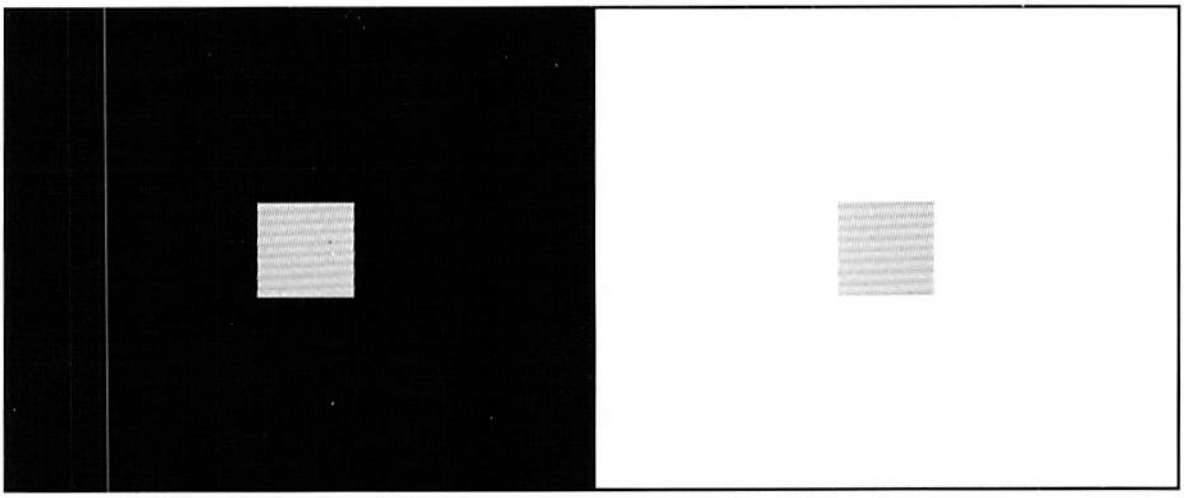
Figure 3-12. Regions of the Electromagnetic Spectrum and Their Corresponding Wavelengths

From *Sensation and Perception* (2nd ed.) by S. Coren, C. Porac, and L. M. Ward, 1984. Copyright 1984 by Academic Press. Reprinted by permission of Harcourt Brace & Company.



Imagine that you are sitting in a dark room viewing a moderately bright monochromatic light of 550 nm. A person with normal color vision would say it is yellowish green. If we increased the number of quanta the light emits, you would say that the light is now brighter. What you experience as **brightness** increases with the light intensity, but before you conclude that brightness depends only on light intensity, look at Figure 3-13, which demonstrates **simultaneous brightness contrast**. The two central patches are identical, but their brightness is influenced by the surroundings. All things being equal, brightness increases with intensity, but it is also affected by other factors.

Figure 3-13. An Illustration of Simultaneous Brightness Contrast
(Original figure)



As we increase the intensity of our 550 nm light, you will probably notice that what appeared as green with just a tinge of yellow now has a much more vivid yellow component. You might say that the color has changed, but this change in appearance is described more precisely as a change in hue. **Hue** refers to our chromatic experience with light, such as redness and greenness. Many people think that particular wavelengths produce definite hues, but this is not entirely correct. Wavelength is related to hue, but one must consider other variables as well, such as intensity. In our example, a single wavelength produced somewhat different hues at different intensities.

A third change in the appearance of our 550 nm light as we increase the intensity is that the tinge of whiteness that was detectable at low intensities has now become clearer. The whiteness or blackness component is another dimension of our color experience, known as **saturation.** A light with little white is said to be highly saturated and appears vivid; a light with more whiteness is less saturated and appears more "washed out."

Thus, there are three dimensions of color experience: brightness, hue, and saturation. These dimensions are not uniquely related to quanta and wavelength. As we increased the intensity of the light, yet kept the wavelength constant, we saw a clear change in brightness, but also a change in hue and saturation.

3.2.2 What is the incidence of color blindness in the general population?

Approximately 9 percent of the population has a congenital color vision deficiency of some type. This is commonly known as **color blindness** but is more appropriately referred to as color deficiency. The most common form of color deficiency is the inability to discriminate red and green wavelengths. This deficiency affects 2.1 percent of the male population and 0.03 percent of the female population. Red-green color deficiency is inherited and is carried by the sex chromosomes, which is why it is many times more likely to occur in males than in females.

The incidence of all types of color deficiency varies across populations. It is present in about 8 percent of Caucasian males, 5 percent of Asian males, and 3 percent of Black and Native American males. Although they are most commonly inherited, color vision deficiencies do not have to be genetically acquired. They can be acquired through disease (e.g., glaucoma or diabetes), injury (e.g., stroke) or as a side effect of the uses of certain drugs (e.g., streptomycin).

Furthermore, approximately 3 percent of private pilots, 2 percent of commercial pilots, and 1 percent of airline transport pilots are known to have some form of color vision deficiency (Society of Automotive Engineers, 1988). Similar figures are probable true for controllers, since they take the same kind of color vision tests that commercial pilots take.

Individuals with abnormal color vision are often good at naming colors. People with such deficiencies learn to use other cues to discriminate colors; they learn, for example, that on a stop light, red is on top. Many color deficient observers could name the colors in most aircraft cockpits without having learned position cues. This does not, however, imply that they can process the colors normally.

Discriminating between the colors may not be normal, especially under conditions in which the colors are desaturated ("washed out"). Search and reaction times are also impaired in color deficient observers. Cole and Macdonald (1988) demonstrated this using cockpit displays with redundant color coding (i.e., the meaning of the display symbols are coded by color and by another cue, such as shape).

Finally, it is important to know that color vision tests are only valid when administered under the proper conditions. The proper illumination of the tests can be obtained with specialized lamps, but because of their expense they are not always used. Failure to use the proper illuminant may result in misdiagnosis or failure to detect a color deficiency. Many of these testing considerations are summarized in a review by the Committee on Vision of the National Research Council (1981).

3.2.3 What other factors affect color appearance?

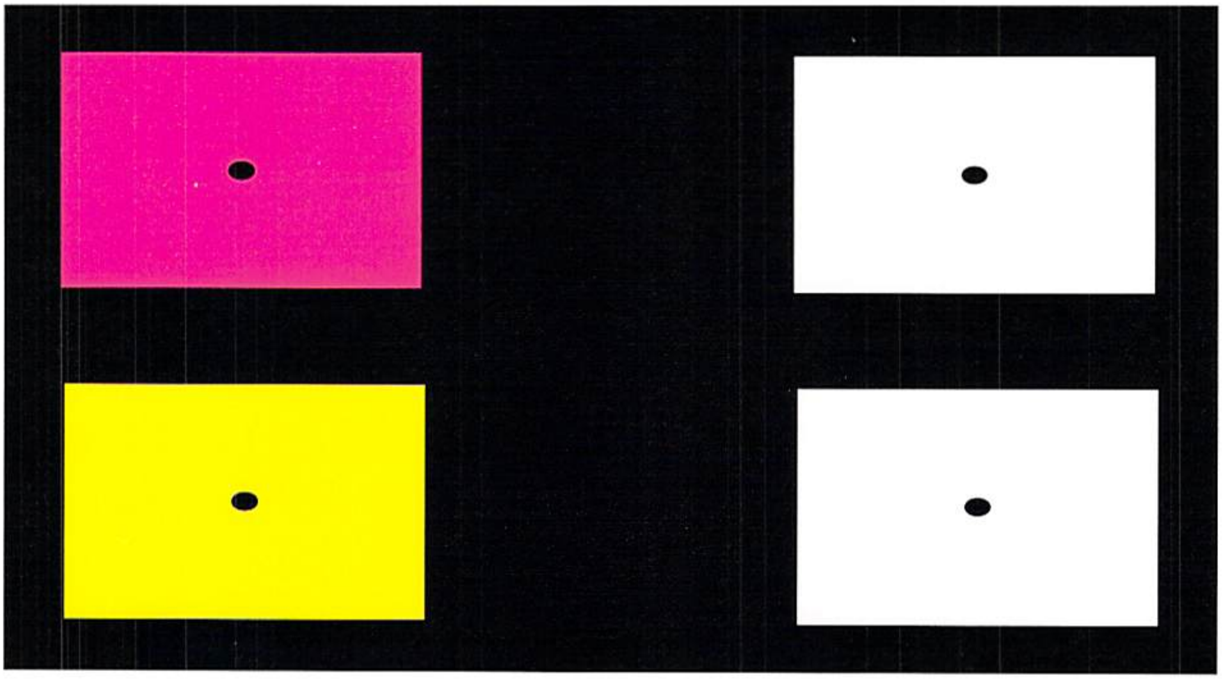
Perception of color is affected by several other factors in addition to those we have already discussed. Other colors can product **contrast effects**, of which there are two major categories: **successive contrast** and **simultaneous contrast.** Sometimes, in a phenomena known as **assimilation**, the brain's processing and interpretation of colors produces an apparent blending of a background color and an interfaced pattern of a different color. The broad distribution of wavelengths in the working environment due to **ambient illumination** can alter the perceptual state of control room personnel and produce changes in computer displays. These effects can be controlled or eliminated through informed design of visual displays and facility environment.

Contrast Effects

The appearance of a color can be altered by another color next to it or another color seen just before or after it. As we scan a scene, we view colors with an eye that has been tuned from moment-to-moment through exposure to preceding and surrounding colors. These contrast effects are dependent on the intensity, duration, and sizes of the stimuli. Here we will illustrate and describe contrast effects, but for detailed summaries of the literature see Graham and Brown (1965) or Jameson and Hurvich (1972).

- **Successive Contrast**: Figure 3-14 illustrates a temporal color-contrast effect. Fixate on the dot on the red square for a while and then shift your gaze to one of the white squares on the right. You will see an **afterimage** of the color complementary to, that is, opposite, that in the picture. In this case, the white square will look greenish after staring at the red square. After staring at the yellow square on the left, the white square on the right will appear bluish.

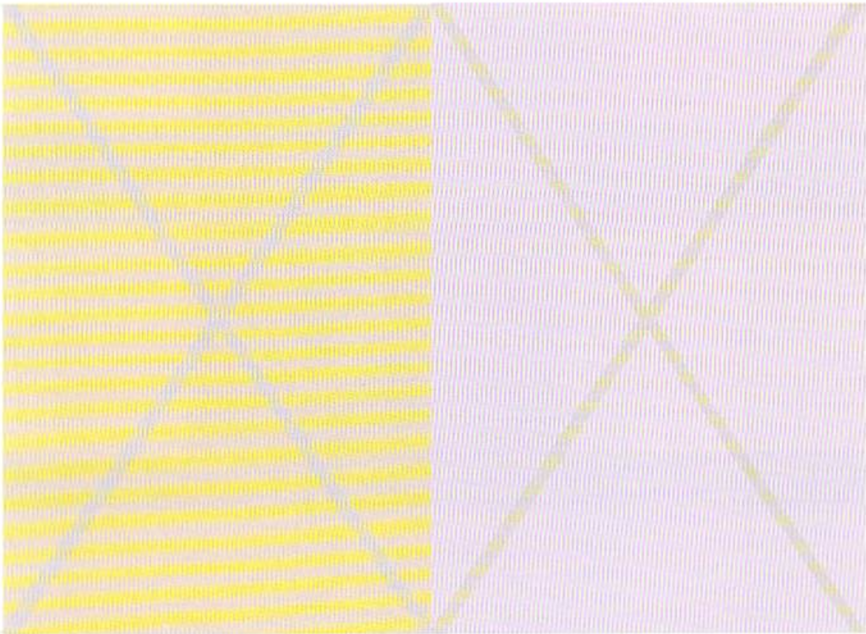
Figure 3-14. A Demonstration of Successive Color Contrast
Adapted from *Color Vision* by L. M. Hurvich, 1981.



Simultaneous Contrast

Figure 3-15 illustrates a spatial, color-contrast effect. The thin bars in the two patterns are identical, but they look different when surrounded by different colors. This is called **simultaneous color contrast** because it occurs instantaneously. The color induced into the focal area is opposite to that of the surround. This is attributable to opponent processes that operate over space; the neural activity in one region of the retina produces the opponent response in adjacent regions. While the effect noticed here is primarily from the surround altering the appearance of the bars, the opposite also occurs.

Figure 3-15. A Demonstration of Simultaneous Color Contrast
From *Interaction of Color* by J. Albers, 1975. Copyright 1975 by Yale University Press. Reprinted by permission.



Through simultaneous contrast, we can experience many colors that are not seen when viewing spectral lights. For example, the color brown is experienced only under conditions of color contrast. If a yellow spot of light is surrounded by a dim white ring of light, it will look yellow. As the luminance of the surround is increased (without changing the luminance of the center), there will be corresponding changes in the central color. First it will look beige or tan, then light brown, followed by dark brown (Fuld, Werner, and Wooten, 1983). If the ring is still further increased in luminance, the central spot will look black.

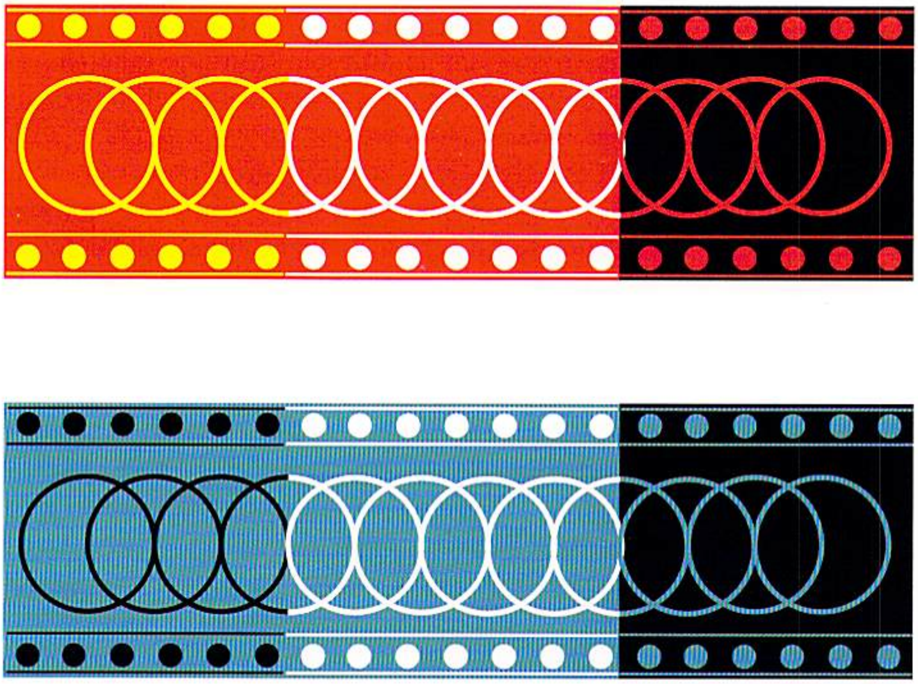
The color black is different from the other fundamental colors because it arises only from the indirect influence of light. That is, like brown, the color black is a contrast color and is only perceived under conditions of contrast. Any wavelength can be used in the center or surround and, if the luminance ratio is sufficiently high, the center will appear black (Werner, Cicerone, Kliegl, and Della Rosa, 1984).

The appearance of a color can be altered by preceding or surrounding colors that are only momentarily in the field of view. For example, white letters may be tinged with yellow when viewed on a blue background or tinged with green when the observer has adapted to a red background. These effects of chromatic adaptation can be altered to work in favor of color identification or detection. For example, detection of a yellow stimulus may be enhanced by presenting it on a blue background.

Assimilation

Sometimes a pattern and background of different colors will not oppose each other as in simultaneous contrast, but will seem to blend together. This is known as **assimilation** and is illustrated by Figure 3-16. Here we see that the saturation of the red background of the top left and center looks different depending on whether it is interlaced with white or black patterns, even though the background is physically the same in the two sections. The lower illustration shows the effect of assimilation with a blue background. Assimilation is not well understood, but it is known that it cannot be explained by light scatter from one region of the image to another. The phenomenon arises from the way in which colors are processed by the brain.

Figure 3-16. A Demonstration of Assimilation, the Bezold Spreading Effect
From *An Introduction to Color* by R. M. Evans, 1948, p. 192. Copyright 1948 by John Wiley & Sons.



Ambient Illumination

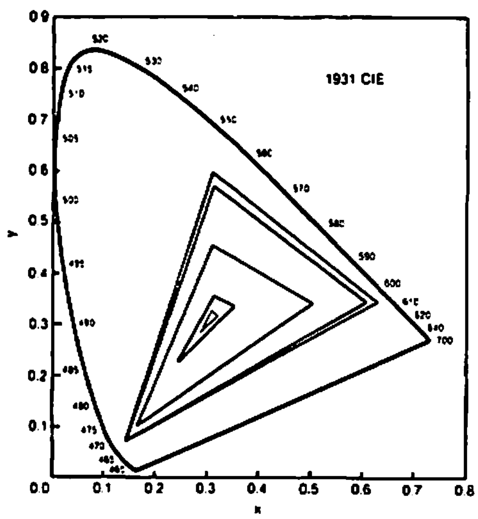
Ambient illumination refers to all of the light present in the immediate environment. It can refer to the sunlight shining in an air traffic control tower, fluorescent office lights, or any source of light that illuminates the workplace. Since most ambient lights contain a broad distribution of wavelengths, the receptors are not adapted as selectively as in laboratory experiments. One important consideration in evaluating changes in ambient illumination under natural conditions is that, in addition to altering the perceptual state of an observer, there often can be substantial changes in the display itself.

CRT screens typically reflect a high percentage of incident light. The light emitted from the display is therefore seen against this background of ambient light. Figure 3-17 shows how sunlight alters the spectral composition of the colors available on a display. As sunlight is added to the display, the gamut of chromaticities shrinks, as illustrated by the progressively smaller triangles (Viveash and Laycock, 1983). To an observer this would be experienced as a desaturation or "wash out" of the display colors as well as a shift in hue that accompanies changes in saturation, called the Abney effect. (See Kurtenbach, Sternheim, and Spillmann, 1984.) Some colors that were previously discriminable may no longer be so. Finally, not illustrated by the figure is the substantial reduction in luminance contrast with increasing ambient illumination. **Fairly low and consistent illumination levels are recommended for optimum color discrimination.**

Some visual displays on aircraft are automatically adjusted in their luminance by sensors that respond to the ambient illumination (e.g., all CRTs on Boeing 757 and 767). This is an important innovation, and indeed consistent with FAA recommendations (RD-81/38,II, p. 47) that alerting signals be automatically adjusted according to the ambient illumination level. However, manual override control is also recommended (RD-81/38,II, page 73) to compensate for individual differences in sensitivity, adaptation, and other factors such as use of sunglasses.

Figure 3-17. Chromaticity Diagram Showing How the Color Gamut of a Display Decreases with Increasing Sunlight

Adapted from "Computation of the Resultant Chromaticity Coordinates and Luminance of Combined and Filtered Sources in Display Design" by J. P. Viveash and J. Laycock, 1983, *Displays, 4*. Copyright 1983 by Butterworth Heinemann Ltd. Reprinted by permission.



We have seen that perception of a fixed stimulus will be changed as a function of many variables including the intensity, surrounding conditions, temporal parameters, and state of adaptation of an observer. If color is a redundant code, these problems, as well as loss of color due to aging of the display, will have substantially less impact on operator performance.

What colors should be avoided on a CRT display?

The choice of colors can be facilitated by considering the physiological principles by which colors are perceived. Colors that are barely discriminable at low ambient conditions may not be at all discriminable at high ambient conditions because of a physical change in the color gamut. This means that choosing the set of colors to be used on a display that will be used in direct sunlight should be done very carefully. The colors should be far apart perceptually so that they are not confusable, even when "washed out" by the sunlight. To color vision experts, this means that the colors should be distant from each other on the CIE chromaticity diagram (a more objective standard than perceptual distinctiveness).

The colors of small symbols and text also need to be chosen carefully. **The use of blue can be particularly problematic for displaying characters requiring good resolution.** The blue phosphors on many displays only produce relatively low luminances, but the main difficulty is a physiological problem in processing short wavelengths. One problem is that the cones most efficient at responding to short-wavelength light are distributed sparsely across the retina and contribute very little to detail vision. Short-wave cone signals are not used in defining borders or contours (Boynton, 1978). In addition, focusing of short-wavelength stimuli is not as easily achieved as for middle- and long-wave stimuli, making blue a color to avoid in displaying thin lines and small symbols.

Furthermore, for the central 2 degrees of the retina, the number of short wave-length cones, or S-cones, present is very small. This has important implications for color discrimination of centrally viewed images. If the visual angle of the image is small and the image is centrally located on the retina, color discriminations that depend on S-cones will be impaired. Color pairs that could be affected include yellow and white or blue and green. Red and green discriminations are not affected since their perception does not depend on S-cones.

A major advantage of blue and yellow is that our sensitivity to these colors extends further out in the visual field than our sensitivity to red and green. Blue hues also provide good contrast with yellow. Thus, while blue may be a good color to avoid when legibility is a consideration, it may be a good color to use for certain backgrounds on displays.

3.2.4 What other factors make it easier or more difficult to identify colors accurately?

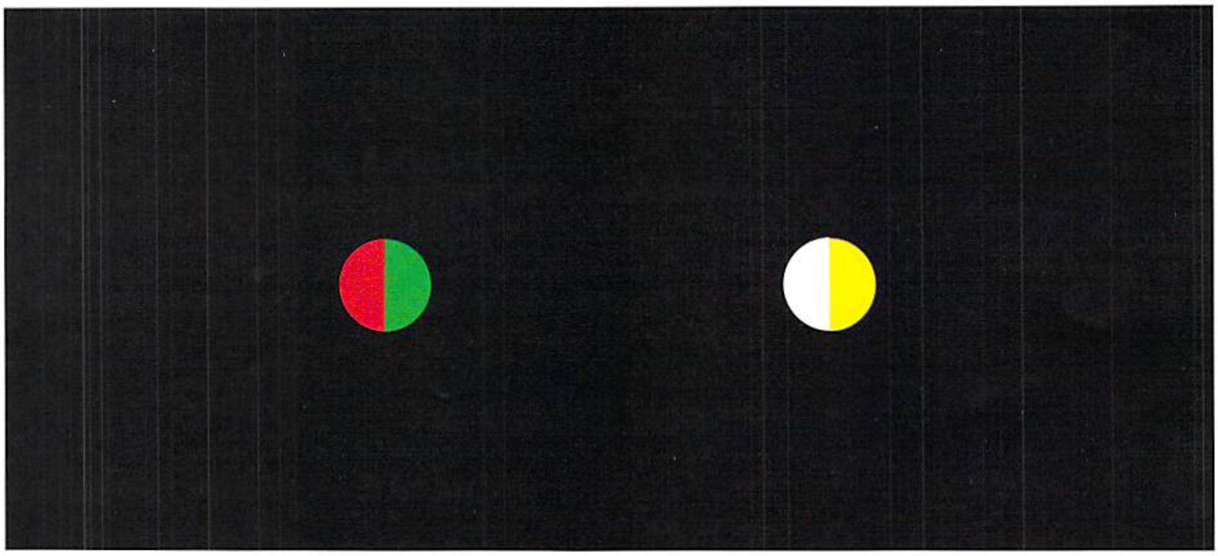
Several factors influence the controller's ability to discriminate colors include:

- The size of the symbols or text
- The number of colors used
- Central versus peripheral viewing
- The age of the observer.

The Size of the Symbols or Text

The size of the symbol or text will affect color appearance. With certain small fields, even normal individuals behave like individuals who completely lack some of the cells necessary for normal color vision. This is illustrated by Figure 3-18. When viewed close, so that the visual angle of each circle subtends several degrees, it is easy for an individual with normal color vision to tell the difference between (i.e., discriminate) the yellow and the white semicircles, and the red and the green semicircles. Viewed from a distance of several feet, however, the yellow and white semicircles seem to merge into one yellowish circle, while the red and green semicircles are still distinct. This is called **small-field tritanopia**, because tritanopes are individuals who completely lack S cones. A tritanope would not be able to discriminate the yellow from the white in Figure 3-18, even close up. Even with the small field condition, the red-green pair is still discriminable because S cones are not necessary for this discrimination. Thus, the small-field effect is limited to discriminations that depend on S cones.

Figure 3-18. Colors (Yellow and White) Not Discriminable at a Distance Due to Small Field Tritanopia
(Original figure)



The Number of Colors Used on a Display

One often hears of displays that are capable of presenting a large number of colors. In some applications, such as map displays, it may be useful to access a large color palette. However, if colors must be identified, not just discriminated, a large color palette may be of little value. For colors to be identified reliably, they must be distinct under a wide range of viewing conditions. The maximum number that fulfills this requirement is probably not greater than six. Of course, in applications that do not require absolute identification (e.g., cartography), the number of discriminable colors that can be used will increase. The number of colors might also be increased when they are only used to reduce clutter and need not be specifically identified.

Central Versus Peripheral Viewing

The way a colored symbol or text appears to us will depend, in part, on whether we are looking directly at it, viewing it out of the corner of our eye, or somewhere in between. Within the central 10 degrees, the observer is responsive to all the basic colors: red, green, yellow, blue, black, and white. As we move out from the center, sensitivity to red and green diminishes. Objects that were previously described as reddish yellow and bluish green are now simply seen as yellow or blue. With further eccentricity, the yellow and blue zones diminish and color responses are limited to black and white. Thus, the accuracy with which we can identify colors in a display depends on whether we are looking at them directly or viewing them peripherally.

The Age of the Observer

We have already seen that the lens of the eye becomes more opaque and yellow with age. This results in an increased difficulty in seeing the color blue, as subtle blue shading will appear as white. In addition to the changes in the lens, our ability to distinguish between subtly different colors decreases with age. This is due to the fact that there is a reduction of approximately 25 percent in cone sensitivity for each decade of life (Werner and Steele, 1988).

The abilities and limitations of the human visual system have important implications for the design and use of visual displays. Specific issues to be considered in the design and evaluation of visual displays are discussed in Chapter 7, Computer-Human Interface (CHI) Considerations.

For more technical information on color displays for ATC, see Appendix A, Human Factors Considerations for Color Displays for ATC.

3.3 REFERENCES AND SUGGESTIONS FOR FURTHER READING

Albers, J. (1975). *Interaction of color.* New Haven, CT: Yale University Press.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The Oblique effect in man and animals. *Psychological Bulletin, 78,* 266-278.

Berson, B. L., Po-Chedley, D. A., Boucek, G. P., Hanson, D. C., Leffler, M. F., Wasson, R. L. (1981). *Aircraft alerting systems standardization study (Volume II: Aircraft alerting system design guidelines,* Report No. DOT/FAA/RD-81/II). Washington, DC: Department of Transportation/Federal Aviation Administration.

Boynton, R. M. (1978). Color in contour and object perception. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception.* New York: Academic Press.

Cardosi, K. M., and Huntley, M. S. (Eds.). (1993). *Human factors for flight deck certification personnel* (DOT/FAA/RD-93/5). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Campbell, F. W., and Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551-566.

Campbell, F. W., and Westheimer, G. (1960). Dynamics of accommodation responses of the human eye. *Journal of Physiology, 151,* 285-295.

Carter, R. C. (1979). Visual search and color coding. *Proceedings of the Human Factors Society 23rd Annual Meeting,* (pp. 369-373). Santa Monica, CA: Human Factors Society.

Christ, R. E. (1975). Review and analysis of color coding research for visual displays. *Human Factors, 17,* 542-570.

Cole, B. L., and Macdonald, W. A. (1988). *Defective color vision can impede information acquisition from colour coded video displays.* Department of Aviation Report No. 3, Melbourne, Australia.

Committee on Vision, National Research Council. (1981). *Proceedings for Testing Color Vision. Report of Working Group 41.* Washington, DC: National Academy Press.

Cornsweet, T. N. (1970). *Visual perception.* New York: Academic Press.

DeHaan, W. V. (1982). *The Optometrist's and ophthalmologist's guide to pilot's vision.* Boulder, CO: American Trend Publishing.

Evans, R. M. (1948). *An Introduction to color.* New York: Wiley.

Fuld, K., Werner, J. S., and Wooten, B. R. (1983). The possible elemental nature of brown. *Vision Research, 23*, 631-637.

Ginsburg, A. P., Evans, D. W., Sekuler, R., and Harp, S. A. (1982). Contrast sensitivity predicts pilots' performance in aircraft simulators. *American Journal of Optometry & Physiological Optics, 59*, 105-108.

Graham, C. H. (1965). Some fundamental data. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 68-80). New York: Wiley.

Graham, C. H., and Brown, J. L. (1965). Color contrast and color appearances: Brightness constancy and color constancy. In C. H. Graham (Ed.), *Vision and Visual Perception* (pp. 452-478). New York: Wiley.

Helps, E. P. W. (1973). Physiological effects of aging. *Proceedings of the Royal Society of Medicine, 66*, 815-818.

Hurvich, L. M. (1981). *Color vision.* Sunderland, MA: Sinauer Associates.

Iavecchia, J. H., Iavecchia, H. P., and Roscoe, S. N. (1988). Eye accommodation to head-up virtual images. *Human Factors, 30*, 689-702.

Jameson, D., and Hurvich, L. M. (1972). Color adaptation: Sensitivity, contrast, after-images. In D. Jameson and L. M. Hurvich (Eds.) *Handbook of sensory physiology* (Vol. VII/4, pp. 568-581). Berlin: Springer-Verlag.

Kaufman, L. (1974). *Sight and mind: An Introduction to visual perception.* New York: Oxford University Press.

Kaufman, L., and Rock, I. (1962). The moon illusion. *Scientific American, 207*, 120-132.

Kurtenbach, W., Sternheim, C. E., and Spillmann, L. (1984). Change in hue of spectral colors by dilution with white light (Abney effect). *Journal of the Optical Society, A, 1*, 365-372.

Newman, C. V., Whinham, E. A., and MacRae, A. W. (1973). The Influence of texture on judgment of slant and relative distance in a picture with suggested depth. *Perception & Psychophysics, 14*, 280-284.

Owsley, C., Sekuler, R., and Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research, 23*, 689-699.

Regan, D., Beverly, K., and Cynader, M. (1979). The visual perception of motion in depth. *Scientific American, 241*, 136-151.

Richards, W. (1970). Stereopsis and stereoblindness. *Experimental Brain Research, 10*, 380-388.

Sekuler, R., and Blake, R. (1985). *Perception*. New York: Knopf.

Society of Automotive Engineers (1988). *Human engineering considerations in the application of color to electronic aircraft displays* (ARP4032). Warrendale, PA: SAE.

Viveash, J. P., and Laycock, J. (1983). Computation of the resultant chromaticity coordinates and luminance of combined and filtered sources in display design. *Displays, 4*, 17-23.

Walraven, J. (1985). The colours are not on the display: a survey of non-veridical perceptions that may turn up on a colour display. *Displays, 6*, 35-42.

Weale, R. A. (1982). *A Biography of the eye*. London: H. K. Lewis.

Weale, R. A. (1988). Age and the transmittance of the human crystalline lens. *Journal of Physiology, 395*, 577-587.

Werner, J. S., Cicerone, C. M., Kliegl, R., and Della Rosa, D. (1984). Spectral efficiency of blackness induction. *Journal of the Optical Society of America, 1*, 981-986.

Werner, J. S., and Steele, V. G. (1988). Sensitivity of human foveal color mechanisms throughout the life span. *Journal of the Optical Society of America, A, 5*, 2122-2130.

Westheimer, G. (1964). Pupil size and visual resolution. *Vision Research, 4*, 39-45.

Wheatstone, C. (1838). Some remarkable phenomena of binocular vision. Reprinted in W.N. Dember (Ed.). (1964). *Visual perception: The Nineteenth century* (pp. 114-129). New York: Wiley.

Yuille, A. L., and Ullman, S. (1990). Computational theories of low-level vision. In D. N. Osheron, S. M. Kosslyn, and J. M. Hollerbach (Eds.), *Visual cognition and action* (pp. 5-39). Cambridge, MA: MIT Press.

Zeki, S. (1992). The visual image in mind and brain. *Scientific American, (Special Issue: Mind and Brain), 267(3)*, 68-76.

CHAPTER 4. AUDITORY PERCEPTION AND SPEECH COMMUNICATION

Kim M. Cardosi and John S. Werner

"...do not think an airline pilot or an air traffic controller never misses a transmission, because he does. The difference is that when an old pro misses a call, he does not think twice about asking to have it repeated."

Brenlove, 1987, p. 23

TABLE OF CONTENTS

CHAPTER 4. AUDITORY PERCEPTION AND SPEECH COMMUNICATION

LIST OF FIGURES AND TABLES

CHAPTER 4. AUDITORY PERCEPTION AND SPEECH COMMUNICATION

Controllers are constantly bombarded with auditory information. Pilots checking in, messages from other controllers, side tones, background noise, and other audible sounds combine to create a complex set of auditory stimuli that the controller must filter through and decipher. Fortunately, some of this is done without the need for conscious attention; for example, a clear transmission from a pilot is not confusable with low-level background noise behind the controller. However, the controller faces many other auditory tasks that are not as simple. It can be difficult, for example, to identify an aircraft checking in when the transmission comes in the middle of a message from another controller. Because the controller's job necessarily involves processing a great deal of auditory information, there are many questions about auditory displays that need to be addressed. This chapter will provide an overview of the limitations of the human auditory system, with particular focus on speech perception. It will also examine pilot-controller voice communications. [Auditory displays are discussed in Chapter 7, Computer-Human Interface (CHI) Considerations.]

4.1 AUDITORY PERCEPTION

Evaluating displays involves examining auditory alarms, alerts, and warnings for operational suitability and acceptability. To assist in these evaluations, this section provides a brief background on the mechanics of hearing. It focuses on factors, such as aging and ambient noise, that affect our ability to hear. Any auditory presentation of information (i.e., auditory display) proposed for ATC environments needs to compensate for these effects.

Perception of speech tends to be taken for granted, but it is actually a highly complex information-processing capability. Because of this natural complexity, miscommunication and misperception of speech are responsible for a high proportion of the human errors that occur in the aviation environment.

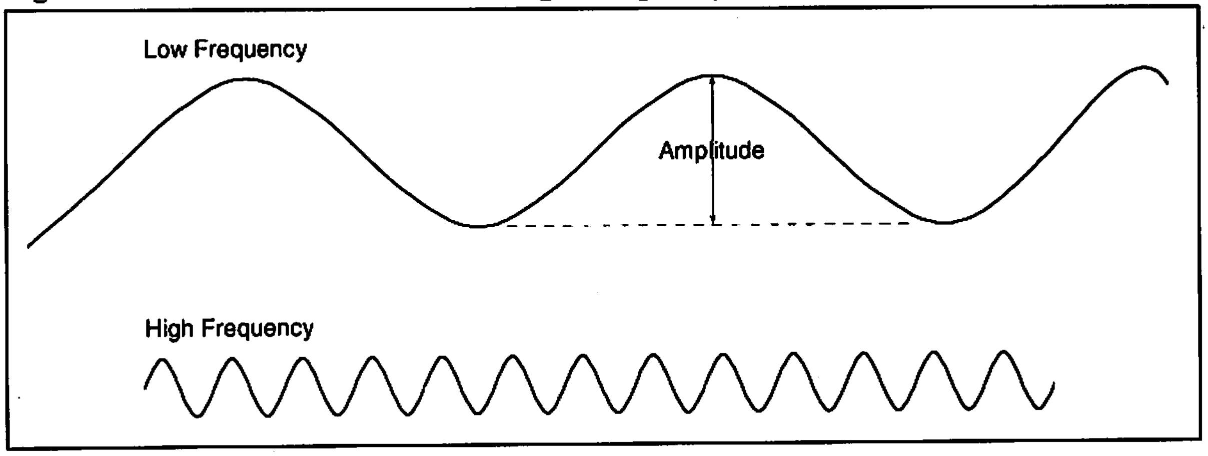
Communications systems designed for use in ATC operations should be evaluated for their ability to improve intelligibility and reduce the incidence of communication problems.

4.1.1 What are the physical properties of sound?

Let us start by examining how sounds are generated. When you pluck the string of a guitar, it vibrates back and forth compressing a small surrounding region of air. When the vibrating string moves away, it pushes air in the opposite direction, creating a region of decompression. As the string vibrates back and forth, it creates momentary increases and decreases in air pressure, or sound waves. These alternating increases and decreases travel through the air at a speed of approximately 740 miles per hour (Mach I, the speed of sound). Eventually they arrive at our ear, where our eardrum (the tympanic membrane) vibrates in synchrony with the pulsations of air pressure.

The simplest pattern of such pressure pulsations is generated for a "pure" tone, or sine wave. One important characteristic of the sine wave is its **frequency**. Frequency is the number of high to low variations in pressure, called **cycles**, that occur within a unit of time. The units we use to describe sound frequency are cycles per second, or **Hertz (Hz)**. Waveforms of low and high frequency tones are illustrated in Figure 4-1.

Figure 4-1. Waveforms of Low and High Frequency Tones.



Another important characteristic of pure tones is the degree of change from maximum to minimum pressure, which we call the **amplitude** or **intensity**, also illustrated in Figure 4-1. Sound amplitude is usually measured in dynes per square centimeter, which is a measure of force per unit area. The human auditory system is sensitive to an enormous range of variations in amplitude of a sound wave–from about 1 to 10 billion. Thus, intensity is more conveniently specified by a logarithmic scale using units called **decibels (dB)**. Equipment that is easy to use (i.e., a variety of sound level meters) is available that measures sound intensity and gives a readout in dBs. Table 4-1 shows some representative sounds on the dB scale.

Table 4-1. The Decibel Scale. From *Perception* by R. Sekuler and R. Blake, 1985, p. 298. Copyright 1985 by Knopf. Adapted by permission of McGraw-Hill.

Sound Pressure (dynes/cm²)	Sound Pressure Level (dB)	Example	Comment
0.0002	0	Threshold of Hearing	
0.00063	10	Normal Breathing	
0.002	20	Leaves Rustling	
0.0063	30	Empty Office	
0.02	40	Residential Neighborhood at Night	
0.063	50	Quiet Restaurant	
0.2	60	Two-person Conversation	
0.63	70	Busy Traffic	
2	80	Noisy Auto	
6.3	90	City Bus	
20	100	Subway Train	Prolonged Exposure Can Impair Hearing
200	120	Propeller Plane at Takeoff	
630	130	Machine-gun Fire, Close Range	
2000	140	Jet at Takeoff	Threshold of Pain
20,000	160	Wind Tunnel	
20,000,000	220	Cannon, Close Range (12" Cannon, 4 m in front of and below muzzle)	

Conversion: dynes/cm^2 = .0002[log^{-1} (dB/20)]

Sine-wave tones are considered pure because we can describe any waveform as a combination of a set of sine waves each of which has a specific frequency and amplitude. A sound comprising more than a single sine wave is termed a **complex sound.** Most of the sounds we hear are complex sounds. The bottom panel of Figure 4-1 shows how two sine waves of different frequencies can be combined to form a complex sound.

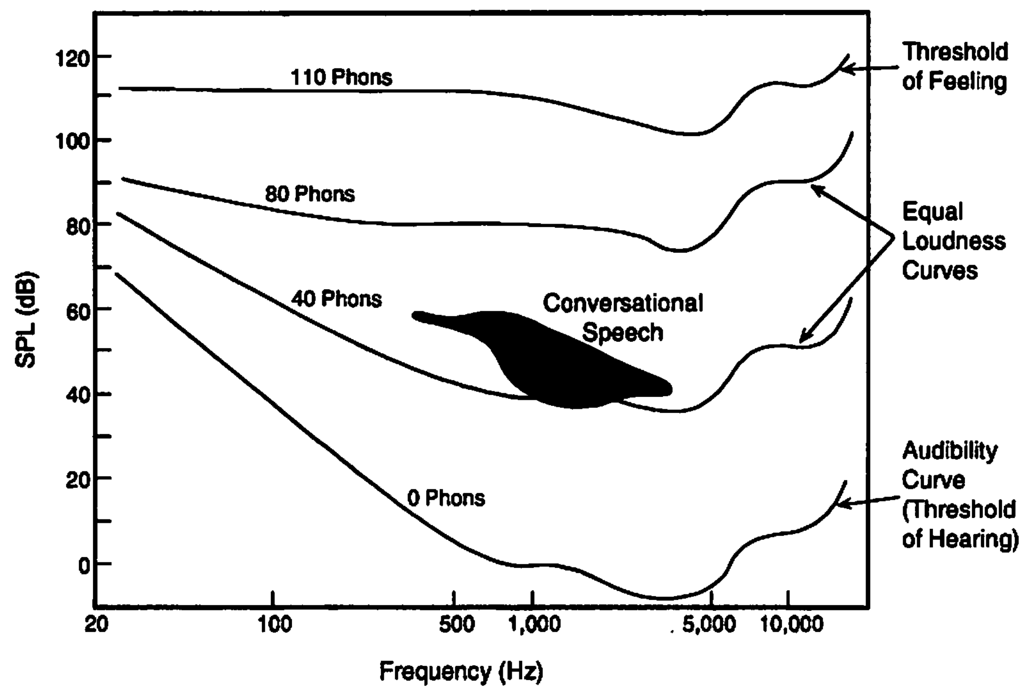
Typically, the pitch we hear in a complex sound corresponds to the pitch of the lowest frequency component of that sound. This component is called the **fundamental frequency.** Frequency components higher than the fundamental are called **harmonics**, and these harmonics affect the quality or the **timbre** of the sound. Two musical instruments, say a trumpet and piano, playing the same note will generate the same fundamental. However, their higher frequency components, or harmonics will differ. These harmonics produce the characteristic differences in quality between different instruments. If we were to remove the harmonics, leaving only the fundamental, a trumpet and a piano playing the same note would sound identical.

How do the Physical Properties of Sound Relate to What we Hear?

First, consider the range of frequencies over which we are sensitive. The lower curve in Figure 4-2 shows how absolute threshold varies with sound frequency for a young adult.

The frequency range over which sounds can be detected is from about 20 to 20,000 Hz. As you can see, we are most sensitive to sounds between 500 and 5,000 Hz. These are also the frequencies found in human speech. Our sensitivity declines sharply, i.e., the threshold increases, for higher and lower frequencies. What this means is that sounds of different frequencies require different amounts of energy so they will be loud enough to be heard.

Figure 4-2. Variation of Absolute Threshold with Sound Frequency for a Young Adult. From "Loudness, its Definition, Measurement and Calculation" by H. Fletcher and W. A. Munson, 1933, *Journal of the Acoustical Society of America, 5*, p. 91. Copyright 1933 by the Acoustical Society of America. Reprinted by permission.



It is interesting to consider our absolute sensitivity under optimal conditions. At about 2,500 Hz, we are so sensitive that we can detect a sound that moves the eardrum less than the diameter of a hydrogen molecule (Békésy and Rosenblith, 1951). In fact, if we were any more sensitive, we would hear air molecules hitting our eardrums and blood moving through our head.

To make the frequency scale a little more intuitive, consider that the range on a piano is from about 27.5 Hz to about 4,186 Hz. Middle C is 262 Hz. As sound frequency is increased from 20 to 20,000 Hz, we perceive an increase in **pitch.** It is important to note though that our perception of pitch does not increase in exact correspondence to increases in frequency. While pitch depends on frequency, it also depends on intensity. When we increase the intensity of a low frequency sound, its pitch decreases. When we increase the intensity of a high frequency sound, its pitch increases.

As we increase the amplitude, or physical intensity, of a particular frequency, its **loudness** increases. Loudness is a perceptual attribute referring to our subjective experience of the intensity of a sound; however, it is not a physical property of the sound. To measure the relative loudness of a sound, researchers typically present a tone of a particular frequency at a fixed intensity and then ask subjects to increase or decrease the intensity of another tone until it matches the loudness of the standard. This is repeated for many different frequencies to yield an **equiloudness contour**.

Figure 4-2 shows equiloudness contours for standards of 40 and 80 dB above threshold. This graph tells us, for example, that a 1,000 Hz tone at 40 dB will sound as loud as a 100 Hz tone at 60 dB. Note that the shape of the contour changes with increasing intensity. That is, the increase in the loudness of a sound with increasing intensity occurs at different rates for different frequencies. Thus, we are much more sensitive to intermediate frequencies of sound than to extremes in frequency. However, as Figure 4-2 illustrates, the difference in our sensitivity to various frequencies decreases with higher intensities.

Sensitivity to loudness depends on the sound frequency in a way that changes with the level of sound intensity. You have probably experienced this phenomenon when listening to music. Listen to the same piece of music at high and low volumes. Attend to how the bass and treble become much more noticeable at the higher volume. Some high-fidelity systems compensate for this change by providing a loudness control that can boost the bass and treble at low volume. The fact that the loudness of a tone depends not only on its intensity but also on its frequency further illustrates the need to objectively measure the characteristics of aural warnings and messages.

4.1.2 How do we judge where a sound is coming from?

The separated locations of our ears allows us to judge the source of a sound. We use incoming sound from a single source to localize sounds in space in two different ways. The intensity of high frequency sounds will be less in the left ear than the right because your head blocks the sounds before they reach your left ear. This **intensity difference** only exists for sounds above 1,200 Hz. At lower frequencies, sound can travel around your head without any significant reduction in intensity.

Whenever a sound travels farther to reach one ear or the other, a **time difference** exists between the arrival of the sound at each ear. Thus, if the sound source is closer to one ear, the pulsations in air pressure will hit that ear first and the other a bit later. We can use a time difference as small as 10 microseconds between our two ears to localize a sound source (Durlach and Colburn, 1978), but this information is only useful for low frequency sounds. Thus, localization of high frequency sounds depends primarily on interaural intensity differences, but low frequency sounds are localized by interaural time differences.

4.1.3 What affects our ability to hear?

Many factors affect our ability to hear. The largest effects are those due to age and exposure. We suffer some hearing deficits solely as a function of age. Our ability to hear a specific sound will also depend on what sounds we have been exposed to in the past and the amount of damage to our hearing that this exposure has produced. Furthermore, it is difficult to listen for a particular sound, such as an individual voice, in a noisy environment due to the background noise "masking" the voice. All of these factors are discussed in detail in the following sections.

Aging

The frequency range for an individual observer is commonly measured by audiologists and is known as an **audiogram.** Figure 4-2 showed that the frequency sensitivity of a young adult ranged from about 20 to 20,000 Hz. This range diminishes with increasing age, however, so that few people over age 30 can hear above approximately 15,000 Hz. By age 50 the high frequency limit is about 12,000 Hz and by age 70 it is about 6,000 Hz (Davis and Silverman, 1960). This loss with increasing age is known as **presbycusis**, and is usually greater in men than in women.

The exact cause of presbycusis is not known. As with all phenomena of aging, there are large individual differences in the magnitude of high frequency hearing loss. One possibility is that changes in vasculature with increasing age limit the blood supply to sensitive neural processes in the ear. Another possibility is that there is some cumulative pathology that occurs with age. For example, cigarette smokers have a greater age-related loss in sensitivity than nonsmokers (Zelman, 1973), and this may be due to the interfering effects of nicotine on blood circulation. There are other possibilities, but perhaps the most important to consider is the cumulative effect of sound exposure.

Effects of Exposure

Sudden loud noises have been known to cause hearing loss. This is a common problem for military personnel exposed to gun shots. Even a small firecracker can cause a permanent loss in hearing under some conditions (Ward and Glorig, 1961).

Exposure to continuous sound is common in modern industrial societies. Even when the sounds are not sufficiently intense to cause immediate damage, continuous exposure may produce loss of hearing, especially for high frequencies. Unprotected workers on assembly lines or airports have hearing losses that are correlated with the amount of time on the job (Taylor, 1965). Similar studies have shown deleterious effects of listening to loud music at rock concerts.

The potentially damaging effects of sound exposure on hearing depend on both the intensity and duration of the sounds. Thus, cumulative exposure to sound over the life span might be related to presbycusis.

Adaptation and Habituation

Our ability to detect sounds is not static but rather changes as a sound is repeatedly presented. This can be due to **adaptation**, a temporary change in sensitivity of the auditory system following exposure to sounds. However, physiological changes in the ability to detect sounds need not occur for us to "tune out" sounds around us. When a stimulus is repeatedly presented, there is a tendency to decrease responsiveness over time. For example, when sitting in a room we may notice a fan when it is first turned on, but over time the noise of the fan is not noticeable at all. This is called **habituation**, a decrease in response (noticing the sound) that cannot be attributed to adaptation. To distinguish between adaptation and habituation, the same sound (e.g., the fan) might suddenly be reduced in intensity (or turned off). If the response is due to habituation, you will notice when the fan goes off, even though you stopped noticing that it was on.

The importance of habituation is clear when specific sounds are repeatedly presented. There is a natural tendency to tune out what is repeated and pay attention to what is new. Tuning out what is repeated, and presumably irrelevant, keeps the sensory channels open to process new information. This is an important fact to keep in mind when designing warning systems. If a warning or alert appears too often, and/or has many false alarms, it will eventually be ignored even when it actually contains important information.

Ambient Noise (Masking)

Warning messages that sound fine in the laboratory may be nearly inaudible in a facility. This is because the sounds in the environment can make it more difficult to hear the warning or message. This effect is called **masking**. The masking of pure tones has been extensively studied (Wegel and Lane, 1924; or see Boff and Lincoln, 1988, Volume I, p. 594). These studies have determined the amount by which signal tones would have to be raised in order to be heard in the presence of different masking tones. The masking effect is strongest (i.e., the signal tone is the most difficult to hear) when the frequency of the mask is similar to the frequency of the tone. The masking effect is also strong for tones of frequencies higher than the mask; there is a weak masking effect on tones of frequencies lower than the mask. Masking effects are complex. Of course, the air traffic environment is not made up of simple tones. However, when deciding what the frequency components of a new warning should be, it is imperative to consider the frequency components of the acoustic environment. **The component frequencies of all warnings should be evaluated for confusability and masking effects.** Auditory masking is a complex phenomenon, but it is easily modeled, once the frequencies contained in the environment are known. Operationally, the simplest, yet critical, test is to measure the effectiveness of auditory signals (tones, warnings, or messages) in the environment that they will be used.

Under some conditions, having two ears reduces the effects of masking. To demonstrate this effect, sounds are played separately to the two ears by use of headphones. Suppose that a tone is delivered to the right ear and it becomes inaudible when masking noise is delivered to that same ear. Now, if the same masking noise (without the tone) is played to the other ear, the tone will become audible again. It is as though the sound coming into both ears can be separated from the tone that is presented to only one ear. This is known as **binaural unmasking.**

Binaural unmasking is probably one factor that helps an individual to focus on one set of sounds in the presence of others. This is a familiar experience for air traffic controllers, in which you can listen to one conversation while tuning out conversations in the background. If your name or sector happens to be mentioned in another conversation, however, you will probably automatically switch your attention in the direction of your name. This underscores our ability to monitor incoming information that we are not actively processing.

4.2 SPEECH PERCEPTION

4.2.1 What are the complexities of speech perception?

The fact that we rarely have trouble understanding what is said to us does not detract from the fact that speech perception is a complex task. Many facets of speech perception make it difficult. For example, if you look at speech as you would any other auditory signal (i.e., in terms of frequency, amplitude, etc.), you would find that it is impossible to tell where one word ends and the next one begins. We, as speakers, do not pause between every word, or even between every sentence. This problem of **parsing** poses difficulties for machines attempting to decode human speech and people trying to learn a foreign language. Both machines and students of a language need to know where a word begins and ends so that they can "look it up." It's almost a processing "Catch 22" that you need to be able to recognize a spoken word in order to know where it begins and ends.

Another factor that adds to the complexity of speech perception is that the exact sounds of individual letters (within a word) can change with speech rate and the context within which they are spoken. A specific vowel sound can be spoken slightly differently in different words, even though they may sound the same. There is also a tremendous amount of **variability** between the same speech sounds as spoken by different speakers. This variability further adds to the problems encountered by speech recognition machines and makes it impossible for a computer to recognize any word spoken by any speaker. (A discussion of speech recognition systems appears in Section 4.2.3.)

There is so much variability among speech sounds that, in many cases, it is only context that allows us to differentiate one from another. This type of variability further increases if non-native English speakers are included. Being a non-native speaker affects not only how we produce speech sounds (e.g., we are likely to have an accent), but it also affects how we **hear** those speech sounds. The most familiar example of this is the ra/la distinction. This distinction is used in English, but is not used in many Eastern languages, such as Japanese. If we were to say "ra" and "la" to native Japanese listeners, the two sounds would sound exactly the same. They cannot distinguish one from the other even though they can distinguish the acoustic cues that differentiate these sounds for native English speakers, when they are presented outside of a speech context (in a laboratory) (Miyawaki, Strange, Verbrugge, Liberman, Jenkins, and Fujimura, 1975).

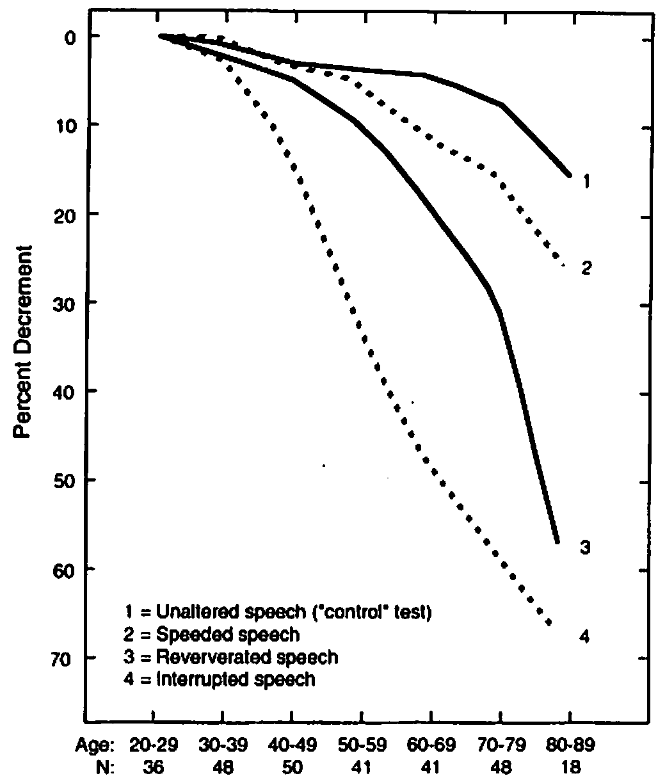
4.2.2 What affects our ability to hear and understand speech?

Several factors influence the intelligibility of speech. One is the **speech-to-noise ratio.** In a noisy environment, some of the critical speech information can be masked. Specifically, the speech sounds that will be lost through masking are the ones that are at and above the frequencies contained in the ambient noise. Generally, as the noise level increases, our ability to understand what is being said decreases.

Other factors that have an effect on our ability to understand speech are the **rate of speech** and the **age of the listener.** When a person speaks quickly in a noisy environment, much more information is lost than when a person speaks slowly in a noisy environment or speaks quickly in a quiet environment. The effects of age on speech perception are two-fold. First, there is a loss of sensitivity, particularly at the higher frequencies, that makes it more difficult to hear certain speech sounds. Second, after about age 50 our ability to differentiate between certain sounds suffers; this makes it even more difficult to understand speech in a noisy environment.

Figure 4-3 shows speech intelligibility as a function of age for normal and degraded speech. When speech is interrupted, speeded, or heard within a background of other voices, our ability to understand it suffers. The degree to which it suffers also depends, in part, on the age of the listener. Until approximately age 50, this type of degraded speech may be annoying, but it does not seriously affect our ability to understand what was said. Around age 70, there is a sharp decline in our ability to understand speech in these degraded conditions (Bergman, Blumenfeld, Cascardo, Dash, Levitt, and Margulies, 1976).

Figure 4-3. Speech Intelligibility as a Function of Age for Normal and Degraded Speech. From "Sampling and longitudinal studies. Age-related decrement in hearing for speech" by Bergman et al., 1976, *Journal of Gerontology, 31*, p. 534. Copyright 1985 by The Gerontological Society of America. Adapted by permission.



Other factors will add to the problems of a pilot (or controller) trying to understand a fast-talking controller (or pilot) in a noisy environment. Devices that transmit speech sounds, such as telephones, radios, and headphones, selectively attenuate certain frequencies. Aircraft radios usually attenuate frequencies above 3,000 Hz. Since these frequencies are within the speech domain (particularly for female speakers), some speech information is likely to be lost. Fortunately, the lost information is not likely to be critical. The best earphones available transmit everything from 25 to 15,000 Hz. These earphones are not very useful for air traffic purposes, since they cannot transmit what the radios have already attenuated. (See sections 9.5.2 and 9.5.3 in Chapter 9, Workstation and Facility Design and Evaluation for recommendations regarding earphones and loudspeakers.)

While many factors (e.g., noise, age, and transmitting devices) can degrade our ability to understand speech, few things can destroy it. One thing that can, however, is referred to in scientific circles as "**delayed auditory feedback**;" the most common form of this is known as an **echo**. It is very disruptive for a speaker to have to listen to his or her own speech slightly delayed. Similarly, if we present another person's speech in one ear and the same speech slightly delayed (more than 30-40 msec.) in the other ear, it makes the listener distressed and unable to understand the message. Delays below 30 msec. are not as disruptive to comprehension, but are annoying and distracting, and could be disruptive to controllers.

This turned out to be an important issue for air traffic controllers when one of the proposed voice switching and control systems (VSCS) had a delay between the "R" controller's voice and the transmission of his/her voice in the "D" controller's headset. Thus, the "D" controller heard the "R" controller's voice in real time and again, slightly delayed, through the headset. A study conducted with air traffic controllers showed that even a five msec. delay was regarded by some of the controllers as undesirable and that a delay as short as three msec. could be detected by five percent of the controllers (Nadler, Mengert, Sussman, Grossberg, Salomon, and Walker, unpublished manuscript). Fortunately, this is an artificial situation (in that it is induced by equipment) and can usually be avoided with better equipment.

Looking at all of these factors, it is amazing that we understand speech as well as we do. Speech as an auditory signal is incredibly complex and is often embedded in noise. Yet, under most circumstances, the system works very well and failures to understand spoken messages are the exception rather than the rule. Unless the signal is inadequate (e.g., as with a microphone that cuts in and out), we usually do a good job of comprehending what is said. Armed with our knowledge of the language and led (or misled) by context and expectations, we are usually able to decipher the signal and understand the message.

4.2.3 What are the most common errors in speech perception?

The most common errors we make in understanding what was said to us are due to expectations, acoustic confusions, and filling in the blanks in what we heard with what we thought we heard.

The powers of expectation are perhaps the most helpful and, at the same time, harmful single factor in speech perception. On the one hand, it is easier to hear a message that is expected than one that is unexpected. Most of the time, knowing what to expect helps us to decipher ambiguous messages. However, **expectation** is a double-edged sword. This same situation can also set us up for errors when we expect to hear one message and actually receive another. This is particularly true when what we *expect* to hear is what we *want* to hear. Both pilots and controllers need to guard against hearing what they expect to hear (this is one reason why catching readback errors is so difficult), and must take the time to question important discrepancies between what they heard and what they (think they) should have heard.

Some speech sounds are naturally more confusable than others. **Acoustic confusions** are even more likely to occur in noisy environments or wherever the speech-to-noise ratio is poor. Avoiding the use of similar sounding letters and numbers reduces possible confusion between them. This is done in air traffic communications with the use of "niner" (so it is not as confusable with five) and with the phonetic alphabet. "Sierra" and "Foxtrot" are much less confusable than "S" and "F." Without the phonetic alphabet, a lot of time and energy would be spent in trying to differentiate between highly confusable letters such as: "M" and "N," "T" and "P," and "S" and "F." Emphasizing the differences in similar sounding call signs (e.g., Universal four **fifty**-two and Universal four **sixty**-two) also helps to reduce the chances of confusion between them. This is particularly important when issuing a clearance that is desired by every pilot in the vicinity.

Hearing what we want to hear, guessing at an insignificant part of a spoken message, and reconstructing after the fact is commonplace. However, we also reconstruct parts of messages *unintentionally*—and we do so with utmost confidence that we *heard* what we actually reconstructed. Its no wonder pilots have confidently told controllers to "check the tapes" and then were embarrassed by the results. Studies, in which people listen to sentences and then repeat what they heard, have shown that if part of a word in a sentence is replaced with a noise, such as a cough or a tone, the listeners fill in the missing syllables and are confident that they heard the missing syllables. They are not able to locate the noise in the sentence, even though they are told to expect the noise somewhere in the sentence (Warren, 1970; Warren and Obusek, 1971).

4.2.4 What can be done to improve intelligibility?

The most basic safeguards toward preserving the **intelligibility** of speech are clear articulation over the proper equipment. Communication devices such as microphones and headsets should relay as much of the speech information as possible with the least amount of degradation (e.g., a delay or clipping). Other factors that aid intelligibility are proper speaking techniques and standardized phraseology. Proper use of the microphone and clear, deliberate articulation can improve a speaker's intelligibility. Limiting the possible vocabulary to a relatively small set of distinct phrases also helps by narrowing down the possibilities of what was said. This is particularly important in designing a set of aural alerts. The fewer the number of different messages, the easier it will be to distinguish between them.

4.2.5 What are the differences between natural and synthetic speech?

Despite manufacturer's best attempts, **synthetic** speech, that is, machine-generated speech (produced by a voice synthesizer) is not as intelligible as speech produced by native-speakers. In natural speech, there are variations in pitch, intensity, and other factors, that are absent in synthetic speech. Because of the obvious differences between synthetic and natural speech, synthetic speech is often preferred to natural speech for alerting functions in environments such as the cockpit and ATC, where many different voices are already present.

While synthetic speech presents some advantages in terms of its distinctiveness, it also has some minor drawbacks. The intelligibility of synthetic speech appears to suffer, more so than natural speech, with increased speech rate (Slowiaczek and Nusbaum, 1985) and ambient noise (Pisoni and Koen, 1982). **For effective communication, speech should be at least 6 dB louder than the surrounding noise** (Licklider and Miller, 1951). For example, if the ambient noise is 50 dB, the speech should be 56 dB. **For important messages produced by synthetic speech, the recommendation is 8 dB above ambient noise** (Simpson, 1980).

Synthetic speech requires some practice for listeners to be able to clearly understand and be comfortable with it. Even after we have adjusted to the system's "accent," however, synthetic speech still imposes increased processing demands on the listener. Synthetic speech is harder to retain in short-term memory than is natural speech (Luce, Feustel, and Pisoni, 1983). For example, if you heard two lists of words, one presented in synthetic speech and the other presented in natural speech, you would be able to recall more of the words from the list spoken in natural speech than from the synthetic list. Designers and evaluators of systems using synthetic speech need to keep in mind these increased demands on memory imposed by synthetic speech.

4.2.6 What is the preferred speech rate?

Our ability to understand both spoken and synthetic (i.e., computer-generated) speech will vary with **speech rate**. If it sounds to us as though the person (or computer) is speaking too slowly or too quickly, our ability to understand what was said can suffer. One study suggests that people listening to isolated spoken sentences prefer a speech rate of 120 words per minute (Licklider and Miller, 1951). The rate for a normal conversation (where context and other cues help fill what is missed) is about 140 words per minute, while a rate of 200 words per minute is not unusual for TV newsreaders.

Preferred speech rate is different, however, for people in high workload conditions who want to hear the message accurately, but quickly. In a study using pilots listening to synthetic speech messages, a speech rate of 150 words per minute appeared to be optimal in terms of intelligibility and pilot preference (Simpson and Marchionda-Frost, 1984). These pilots (who had been trained on the synthesizer's "accent") said that the slower rate of 126 words per minute took too much time away from their primary task of flying.

4.3 PILOT-CONTROLLER COMMUNICATIONS

4.3.1 How have pilot-controller communications been studied?

A special case of speech perception that is of interest to air traffic controllers and pilots is the communication that takes place between them. The sheer volume of communications between pilots and controllers makes human error inevitable. The opportunity for miscommunication is constant, and the consequences can range from annoying to dangerous. At the very least, miscommunications result in increased frequency congestion and increased controller workload, as more communications are necessary to correct the problem. Depending on the nature of the error, miscommunications have the potential for narrowing the margin of safety to an unacceptable level. Information obtained by sampling pilot-controller voice communications is useful in a variety of ways. It provides necessary insights into the frequency of occurrence of specific practices that are known to affect the accuracy of communications.

Studies in pilot-controller communications typically use one of three methods: analysis of incident reports, simulation studies on specific communication issues, and voice tape analysis.

Incident Reports. The largest informal reporting system in the U.S. is the Aviation Safety Reporting System (ASRS) run by NASA-Ames Research Center. This data base comprises voluntary and confidential reports that are submitted on any factor believed to be important to aviation safety. Although anyone can submit an ASRS report, most reports are submitted by airline pilots (who receive limited immunity for filing). The ASRS database is often searched for a subset of reports on a specific topic. The results are available as NASA ASRS publications. A partial listing of these publications is given in Table 4-2. (NASA ASRS will provide a complete listing of available reports upon request.) As can be seen in Table 4-2, ASRS publications cover various topics of interest to ATC, such as the readback/hearback problem (Monan, 1986). These studies are useful for pointing to trouble spots in the system, examining trends, and suggesting remedies for specific problems. For example, pre-departure clearance (PDC) is now being used at many facilities across the country. With an increase in the number of facilities using PDC, there is a concomitant increase in the number of ASRS reports on this issue. One of the factors that these reports point to is the lack of standardization in the format of the clearances. For example, the same information can appear in different positions on the print-out, depending on the issuing facility. These and other human factors issues are being examined as a result of the focus provided by the analysis of ASRS reports.

Table 4-2. NASA ASRS Publications, Partial List.

These documents may be requested through NASA ASRS, P.O. Box 189, Moffett Field, CA. When ordering documents, please cite the number beside the title.

Problems in Briefing of Relief by Air Traffic Controllers. QR #12.

Information Transfer Between Air Traffic Control and Aircraft: Communications Problems in Flight Operations. TP 1875.

Addressee Errors in ATC Communications: The Call Sign Problem. CR 166462.

Non-Airborne Conflicts: The Causes and Effects of Runway Transgressions. CR 177372

Human Factors Associated with Runway Transgressions. Special Paper.

ATC Control and Communications Problems: An Overview of Recent ASRS Data. Special Paper.

Human Factors in ATC Operations: Anticipatory Clearances. Special Paper.

A unique study of incident reports was conducted in conjunction with the **USAir Altitude Awareness Program** (MiTech, Carlow, FAA, 1992). This study examined 131 altitude deviation incident reports that were submitted by pilots between September 1990 and November 1991 and 474 reports submitted by air traffic controllers between May and November 1991. Reports submitted by pilots described the following pilot behaviors:

- Responding to the wrong call sign.
- Setting the wrong altitude in (or forgetting to set) the altitude alerter.
- Transposing or misunderstanding numbers.

Reports from controllers described the following issues:

- Pilots taking another aircraft's clearance (and the controller not catching the readback).
- Confusion between numbers given in the clearance.
- Confusion between altitudes of 10,000 (one-zero-thousand) and 11,000 (one-one-thousand), that is, hearing 10,000 when 11,000 was said and hearing 11,000 when 10,000 was given.
- Traffic advisories taken as altitude clearances.
- Language problems with foreign pilots.

Issuing multiple instructions (e.g., changes in altitude, speed, heading, etc.) in the same transmission was associated with 49 percent of the altitude deviations and 48 percent of the potential altitude deviations.

Simulation Studies. Only a few simulation studies have examined pilot-controller communications, but their number is increasing. In a simulation study, pilots either "fly" predetermined routes in a simulator or perform other pilot duties. One such study looked at the effect of controller message length on pilot readbacks. This study found that one transmission containing four commands (clearances) resulted in more inaccurate and partial readbacks than two transmissions containing two commands each (Morrow, 1993).

Voice Tape Analysis. Most studies that examine voice tapes from ATC facilities focus on specific aspects of pilot-ATC communications. For example, when Conflict Resolution Advisory (CRA) was being developed, the software engineers needed to know how much time should be expected to elapse between the time the controller issued a maneuver required for traffic avoidance and the time the aircraft maneuvered. This time has many components and includes pilot, controller, and system (transmission) response times. The pilot must first respond to the controller's transmission, and the controller may have to repeat the clearance (e.g., if the pilot does not respond or responds with a "say again"). In order to determine what this response time is, 46 hours of voice tapes containing controller to pilot communications from three Air Route Traffic Control Centers (ARTCCs) were analyzed (Cardosi, 1993b)[1]. In these 46 hours of tapes, 80 communications from controllers to pilots were found to contain time-critical messages, such as maneuvers required for traffic avoidance, or maneuvers followed by words expressing urgency (e.g.,"now" or "immediately"). The mean duration of the controller's initial call for these transmissions was 4.8 seconds. The pilots' verbal response times, as measured from the end of the controller's transmission to the beginning of the pilot's acknowledgement, ranged from one to 31 seconds with a mean (i.e., average) of three seconds (standard deviation = 5). The 90th percentile was 13 seconds. This means that we would expect most (90 percent) pilot responses to be initiated within 13 seconds.

On 5 percent of the transmissions, the controller received no response from the pilot on the first attempt to contact and had to try again. On 16 percent of the calls, the controller had to repeat or clarify part or all of the transmission once contact was established. This second call lasted an average of 3.2 seconds and the pilot's final acknowledgement averaged 1.7 seconds. The average response time, as measured from the end of the controller's transmission to the end of the pilot's initial transmission (even if it was only a "say again") was six seconds. The total time required for successful transmission of a time-critical message was measured from the beginning of the controller's transmission to the end of the pilot's correct acknowledgement (and included "say agains" and other requests for repeats). This total time ranged from four to 40 seconds and averaged 10 seconds. Ninety percent of the transmissions were successfully completed within 17 seconds.

Interestingly, times required to complete similar, but not time-critical transmissions, such as turns issued by controllers for reasons other than traffic avoidance, were very similar. The time required for successful transmission of such calls ranged from four to 52 seconds with a mean of 10 seconds.

Incidence Data

While ASRS reports and special issue reports can indicate an area that needs attention, they cannot give any information about the prevalence of problems. Studies that purport to examine the rate of occurrence of specific problems must use a broad sample of communications. Such a sample should include many hours of communications and should sample different facilities of a specific type (e.g., TRACONs). A few studies, with more in progress, have examined actual pilot-controller communications and recorded the incidence (i.e., percentage of occurrences) of specific problems, such as readback and hearback errors.

In one such study, 42 hours of voice tapes from four TRACONs were analyzed (Morrow, Lee, and Rodvold, 1993). This study found that less than one percent of the controller transmissions resulted in an incorrect readback. Incorrect readbacks were more common following longer controller messages than following shorter messages. Deviations from standard communication practices (e.g., dropped call signs, partial or missing readbacks) were more common than readback errors, occurring in 3 to 13 percent of the acknowledgements (depending on the TRACON sampled).

Another study examined pilot-controller communication practices in the en route environment (Cardosi, 1993a). There were 5,032 controller-to-pilot transmissions on the 48 hours of voice tapes analyzed. This included 3,576 clearances (e.g., instructions to maneuver or change radio frequencies, routing changes) and 1,456 requests for information, salutations, controller acknowledgements, and so forth.

The great majority of clearances contained only one or two pieces of information and were acknowledged with a full or partial readback. Less than one percent of the full readbacks contained an error, while two percent of the partial readbacks contained an error. An analysis of communication errors shows two effects of increased message length. First, the shorter the controller's transmission, the more likely the pilots were to respond with a full readback. Second, the longer the transmission, the more likely was a readback error. The readback error rate doubled (from .7 percent to 1.4 percent) as clearances increased in complexity from three elements to four. Still, the overall error rate was quite low, until clearances containing five or more elements were examined. While this category accounted for only 4 percent of the clearances examined, it accounted for 26 percent of the errors found. The most common type of readback error involved frequency changes. Such errors accounted for 37 percent of the 27 readback errors found in the analysis. The second most common type of error involved crossing restrictions; this accounted for 18 percent of the readback errors.

Pilots gave their complete call sign (i.e., airline name and flight number or last three alphanumerics for a general aviation aircraft) in only 58 percent of these erroneous readbacks. No call sign was given in 27 percent of these readbacks. In 29 instances (.8 percent of the clearances) pilots responded to transmissions with different call signs than the controllers used. What was surprising about these incidents was that only 45 percent of these call sign confusions were corrected.

In 51 instances (1.4 percent of the clearances) pilots requested that a controller repeat all or part of the transmission. For clearances containing five or more elements, the rate of pilot requests for repeats ranged from 1 percent to 2.5 percent. For clearances containing one to four elements, the rate was almost four percent. There was a 1 to 3 percent miscommunication rate (errors combined with requests for repeats) for clearances containing one to four pieces of information and a 8 percent rate for transmissions containing five or more elements. Clearly, the more information contained in a transmission, the higher the probability that a controller will need to repeat all or part of that message.

In only three instances the controllers did not notice an error in the pilot's readback. This represents 11 percent of the observed readback errors and less than one-tenth of one percent of the total number of clearances.

Several factors of interest (e.g., speech rate, non-standard phraseology) were examined as coincident to the communication errors. No relation was found between any single factor and communication errors, nor was there any evidence that any of these factors caused the error. However, there were too few errors found to allow for an in-depth study of communication errors.

One of the most striking findings of this analysis was how few errors were found. An error rate of less than one percent is a tribute to the pilots and controllers operating in the National Airspace System. Still, pilots and controllers need to be aware that catching readback errors is a difficult task, particularly when combined with other duties that need to be performed simultaneously. Pilots need to be encouraged to ask for clarification, rather than expect the controller to catch readback errors. Controllers also need to be aware that two shorter transmissions may be more effectively transmitted to an unsuspecting pilot than one longer one. (Different results would be expected when pilots are prepared for a longer transmission, as in clearance delivery.) Such increased awareness can further reduce the probability of communication problems and further increase the margin of safety.

4.4 REFERENCES AND SUGGESTIONS FOR FURTHER READING

Békésy, G. and Rosenblith, W. A. (1951). The mechanical properties of the ear. In S.S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1075-1115). New York: Wiley.

Bergman, M., Blumenfeld, V. G., Cascardo, D., Dash, B., Levitt, H., and Margulies, M. K. (1976). Age-related decrement in hearing for speech. Sampling and longitudinal studies. *Journal of Gerontology, 31*, 533-538.

Boff, K. R. and Lincoln, J. E. (1988). *Engineering data compendium: Human perception and performance.* Wright Patterson AFB, Ohio: Armstrong Aerospace Medical Research Laboratory.

Brenlove, M.S. (1987). *The air traffic system: A commonsense guide.* Ames, IA: Iowa State University Press.

Cardosi, K. (1993a). *An analysis of en route controller-pilot voice communications.* DOT/FAA/RD-93/11. Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Cardosi, K. (1993b). Time required for transmission of time-critical ATC messages in an en route environment. *International Journal of Aviation Psychology, 3*, 303-314.

Davis, R. G, and Silverman, S. R. (1960). *Hearing and deafness.* New York: Holt, Rinehart and Winston.

Durlach, I., and Colburn, H. S. (1978). Binaural phenomena. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception* (pp. 265-466). New York: Academic Press.

Licklider, J. C. R., and Miller, G. A. (1951). The perception of speech. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1040-1074). New York: Wiley.

Luce, P. A., Feustel, T. C., and Pisoni, D. B. (1983). Capacity demands in short-term memory for synthetic and natural speech. *Human Factors, 25*, 17-32.

MiTech, Inc., Carlow International, Inc., and Federal Aviation Administration Research and Development Service (October, 1992). *Altitude deviation study: A Descriptive analysis of pilot and controller incidents. Final Report.* Washington, DC: Department of Transportation/Federal Aviation Administration.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. (1975). An effect of linguistic experience: The Discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*, 331-340.

Monan, W. (1986). *Human factors in aviation operations: The Hearback problem.* (NASA Contractor Report 177398). Moffett Field, CA: NASA Ames Research Center.

Morrow, D., Rodvold, M. (1993). The influence of ATC message length and timing on pilot communication. *Proceedings of the Seventh International Symposium on Aviation Psychology*, (Vol. I, pp. 324-327). Columbus, OH: Ohio State University.

Morrow, D., Lee, A., and Rodvold, M. (1993). Analysis of problems in routine controller-pilot communication. *International Journal of Aviation Psychology, 3*, 285-302.

Nadler, E., Mengert, P., Sussman, E. D., Grossberg, M., Salomon, A., and Walker, K. (1992). *Effect of binaural delays in the communications channel linking radar and data controllers*, unpublished manuscript.

Pisoni, D. B., and Koen, E. (1982). Intelligibility of natural and synthetic speech at several different signal-to-noise ratios. *Journal of the Acoustical Society of America, 71*, UU1.

Sekuler, R., and Blake, R. (1985). *Perception*. New York: Knopf.

Simpson, C. A. (1980). *Synthesized voice approach callouts for air transport operations.* (NASA Contractor Report NASA CR-3300). Moffett Field, CA: NASA-Ames Research Center.

Simpson, C. A., and Marchionda-Frost, K. (1984). Synthesized speech rate and pitch effects on intelligibility of warning messages for pilots. *Human Factors, 26*, 509-517.

Slowiaczek, L. M., and Nusbaum, H. C. (1985). Effects of speech and pitch contour on the perception of synthetic speech. *Human Factors, 27*, 701-712.

Taylor, C. A. (1965). *The Physics of musical sounds.* New York: Elsevier.

Ward, W. D., and Glorig, A. (1961). A case of fire-cracker induced hearing loss. *Laryngoscope, 71*, 1590-1596.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*, 392-393.

Warren, R. M., and Obusek, C. J. (1971). Speech perception and phonemic restorations. *Perception and Psychophysics, 9(38)*, 358-362.

Wegel, R. L., and Lane, C. E. (1924). The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Physical Review, 23*, 266-285.

Zelman, S. (1973). Correlation of smoking history with hearing loss. *Journal of the American Medical Association, 223*, 920.

4.5 ENDNOTES

1. Because CRA was designed for use in ARTCCs only (and not in the terminal environment), only communications from ARTCCs were included in the analysis.

CHAPTER 5. HUMAN INFORMATION PROCESSING

Elizabeth D. Murphy and Kim M. Cardosi

"Many human-machine systems do not work as well as they could because they impose requirements on the human user that are incompatible with the way a person attends, perceives, thinks, remembers, decides, and responds, that is, the way a person processes information."

— Wickens (1992a)

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 5. HUMAN INFORMATION PROCESSING

One purpose of this chapter is to provide an introduction to the basics of human information processing and decision making. These are important topics because the controller's job centers on planning, processing information, and making decisions under time pressure. ATC systems must be designed to support and complement these critical processes.

5.1 BASICS OF HUMAN INFORMATION PROCESSING

As a basis for building and maintaining an awareness of the traffic situation and for issuing control instructions, the controller processes many different kinds of information from many sources: flight data information, weather information, communications from pilots and other ATC personnel, information on equipment status, and various other kinds of information too numerous to list. On the basis of his or her assessment of the situation, the controller makes and implements decisions. He or she decides if any control action is necessary, decides what action to take, formulates instructions, and issues instructions to aircraft, often under very tight time constraints. All of these are information-processing activities.

Controller information processing and decision making are key to daily success in meeting ATC goals, but they are also associated with a cost: the time it takes to process information and make decisions. Further, these activities contribute to controller workload, and they are subject to error. If automated decision aiding is to save time and effort for the controller and reduce the likelihood of error, it must be designed to support the controller's natural approach to situation assessment and selection of control actions.

This chapter points out the strong link between what we know about human information processing and how we should design systems. It is reasonable to assume that automated systems are designed for compatibility with the human information processing system, but this is not always the case. **One of the roles of the acquisition team is to**

evaluate design products for operational suitability. If a system is incompatible with its users' thought processes, it is operationally unsuitable.

To aid the team in planning evaluations, this chapter includes evaluation tips derived from what we know about the ways in which people process information. These tips are given in boldface type.

5.1.1 What do we mean by the phrase "information processing"?

To begin with, it may be valuable to make a distinction between data and information. **Data** (always plural) are the raw materials (e.g., numbers; symbols; physical energies, such as light or sound waves) that become information only after they have been processed analytically. **Information** is more valuable than are isolated items of data because information conveys more meaning. A discussion of formal information theory can be found in Wickens (1992). Here we consider the issue from a less formal perspective.

The numbers "2" and "4" are data. The fact that 2 plus 2 equals 4 is information. An aircraft's altitude (e.g., FL290) is the product of a transformation, that is, a computation. This transformation makes altitude an information element to be processed in the context of other information about the overall ATC situation (e.g., the aircraft's position, direction of flight, proximity to other aircraft). Transforming data into information takes time and mental effort if the computations have to be performed by the controller. **A design should minimize time spent on mental transformation of data. Information should be presented to the controller in a usable form when it is needed.**

People extract data and information from the environment via the senses, centrally process that input in numerous ways (e.g., mental arithmetic, inductive or deductive reasoning, extrapolation), and produce some action (e.g., verbal communications, control action). An "action" may be observable or unobservable. An unobservable action may be a decision to remember to issue a clearance at a later time. Observable actions include keying the mike and speaking. It takes times to progress from sensing or perceiving information to making a decision and acting. For example, the controller notices a conflict alert, confirms that there is a potential conflict, formulates a plan and, perhaps, a backup plan to avert the conflict, and then issues instructions to at least one of the aircraft involved. To some extent, human memory records this process of moving from sensing to acting and gains knowledge over time, through training and experience.

System design must allow sufficient time for the controller to perceive information; integrate new information with known information; manipulate information; project potential outcomes of possible courses of action; and formulate decisions, instructions, or other output, including intentions to be carried out later. The time required for these information-processing activities will vary with operational conditions and with the controller's level of experience.

5.1.2 How fast should the computer respond to support the controller's information processing?

The human brain can be thought of as a network of billions of cells called neurons. Connections between neurons allow rapid sequential and parallel processing of internal and external sensations (e.g., light striking the retina, sound waves striking the eardrum, thoughts, memories, hunger pangs). These neural networks process a great deal of such information "automatically," that is, with little or no conscious attention. Right now, for example, you know whether or not lighting is adequate; your favorite song is playing; it's almost lunchtime; or the room is too hot, even though you may not have been thinking consciously about any of these topics. Your brain is constantly processing data from both the internal and external environments, even when your conscious attention is directed elsewhere.

Some physical skills become "automatic" over time, with practice. These skills include handwriting, typing, playing the piano, and riding a bicycle. Skills depend on an extremely short temporal loop between action and feedback. If feedback is delayed, action becomes jerky and effortful. In the same way, the controller's skilled, automatic interaction with data entry devices can be adversely affected by lengthy computer response times.

Computer response times must be fast enough that the controller is not kept waiting for information. The computer should respond within: 150 milliseconds (msec.) for individual steps in data entry, pointing, and selection tasks; one second for simple, frequently performed tasks; and display of critical information; two to four seconds for more complex processing; and eight to twelve seconds for infrequent, highly complex tasks (Shneiderman, 1992). These response times may vary, but large variations in either direction may be disruptive. The user is frustrated by delays that are much longer than expected and suspicious of delays that are much shorter than expected. An indication of the time required for processing, such as a clock showing the countdown of the delay or a display of "percent completed" helps the user understand the nature of the delay and anticipate the end of the delay.

While a fast computer response time is essential, so that the controller is not slowed down by the computer, it is also important to guard against errors that are caused by the controller proceeding "automatically" with one response when another response is called for (e.g., quickly and accurately entering the wrong type of data). Design should be supportive of smooth, fast, "automatic" responses when appropriate, but be able to alert the controller when conscious attention is needed.

Information-processing tasks include perceiving and understanding the speech of others, communicating by means of language, planning, solving problems, making decisions, exercising judgment, and responding creatively to unfamiliar situations. People are flexible information processors who are sensitive to changing conditions and situations. They are resourceful in using both quantitative and qualitative information and in integrating information received from various sources. It is these unique information-processing abilities, honed by training and experience, that make the controller an invaluable component of the ATC system.

System design should support the controller's information-processing tasks, particularly the integration of information from multiple sources, and should provide integrated information where possible.

5.1.3 How much time is consumed by information-processing activities?

Pressing a button when a light comes on is an example of a simple action. In general, the time consumed by such simple actions is measured in milliseconds (ms) or thousandths of a second. Five hundred (500) milliseconds equal one-half second. More complex tasks are measured in seconds. Table 5-1 gives time ranges required for completion of some information-processing task constituents.

Table 5-1. Time Required for Information-Processing
Material marked with asterisk from Boucek, Pfaff, White, and Smith, (1985) p.70.
Material marked with triangle from Campbell and Westheimer (1960).

Information Processing Task Component	Time Required
Making a single, directed eye movement (saccade)	180 to 250 ms
Bringing an object into focus (accommodation)◂	360 ms
Noticing a warning message that is presented aurally or visually and directly in front of you*	2-3 seconds
Deciding what to do* (action selection)	5-6 seconds
Initiating a response*	2-3 seconds

Such time data for task components define the limits on what the controller can be expected to do in a given amount of time. For example, because the time required for two saccades and two accommodations could add up to 1,220ms or 1 second and 22ms, the design should not require more than two eye movements per second (Moray, 1986). Furthermore, information that the controller must read should never blink or flash rapidly. Note that this is not as much of a problem when the controller must only recognize what information is flashing (e.g., a known data block). Blinking and other forms of highlighting are discussed in Chapter 7.

5.1.4 What are the elements of a complex information-processing response?

A complex response, such as one to an automated warning system, may be broken down into four components: detection time, time to identify and interpret the message, decision time, and time to initiate (or complete) the appropriate response. When a warning signal or message appears, for example, the first component of the required response is to detect the presence of that signal (or incoming message), that is, to notice that it is there. The second component of the response is the interpretation of the message. The operator needs to identify the message. For example, when blinking text appears on a controller's scope, the first part of the response is to notice that some portion of the display is blinking.

The second step is to identify which part of the radar scope is blinking (e.g. an "H" and sector number, a set of datablocks, or a group of letters), so that the blinking may be interpreted (e.g., as a handoff, conflict alert, or an aircraft emergency). While this stage may sound simplistic, the task becomes more difficult as the number of possible alarms and messages (e.g., the number of things that can possibly blink and their meanings) increases. After deciding which message it is, the next response component required is to decide what physical action, if any, (e.g., issue a clearance to an aircraft, accept a handoff, coordinate with another controller, suppress the blinking, etc.) is required. Then, and only then, can a physical response be initiated.

Although experiments that examine controller response times to advisory systems are rare, results of a series of flight simulation studies indicate that with a system that requires immediate action, it will take a pilot approximately two to three seconds to detect that a message is there, five to six seconds to decide what to do about it, and one to two seconds to initiate a response (Boucek, White, Smith, and Kraus, 1982; Boucek, Po-Chedley, Berson, Hanson, Leffler, and White, 1981; Boucek, Erickson, Berson, Hanson, Leffler, Po-Chedley, 1980; see also Berson, Po-Chedley, Boucek, Hanson, Leffler, and Wasson, 1981). This leads to a total of eight to eleven seconds that should be allotted for a pilot to initiate a response. The time required to complete the response will, of course, depend upon the task. Actual task completion times for specific ATC tasks will need to be determined through performance measurement, as discussed in Chapter 10, Human Factors Testing and Evaluation.

5.1.5 What factors affect the time required to process information?

Many factors are known to affect human performance, and hence, response time. Some of these factors are characteristics of the visual or auditory display. Others are characteristics of the operator, such as previous experience, skill, and fatigue. (See Chapter 8, Workload and Performance Measurement in the ATC Environment, for a discussion of these issues.) Still others are characteristics of the operational or testing environment, such as workload, consequences of errors, and instructions to the participants. Each of these factors needs to be considered from the test design to the interpretation of the results, and they must be controlled as much as possible during a test.

Display factors. Detection of visual signals are affected by such factors as location in the visual field, and presentation format (e.g., blinking versus steady text, font size, and brightness). (See Chapter 3, Visual Perception, and Chapter 7, Computer-Human Interface (CHI) Considerations, for additional information on visual perception and visual displays.) Human response time will be faster if the signal is presented in the center of the visual field, as opposed to out on the periphery. If it is presented in the periphery, but flickering or blinking, detection time will be faster than if it is in the periphery but steady. (This is one reason why a flickering part of a display can be distracting.)

Blinking is an effective means of attracting the controller's attention, if it is not overused. The blink rate should be between .1 and 5 Hz; however, a blink rate of 2 to 3 Hz is preferred (Gilmore, Gertman, and Blackman, 1989). Since blinking text is difficult to read, the "on" time of the blink should be longer than the "off" time. An alternative to blinking text is to present the text as steady with a blinking symbol next to it.

Intensity is also an important factor. Within limits, a higher intensity stimulus will attract attention more efficiently than will a less intense stimulus. In the visual domain, intensity translates into brightness, although other factors, such as contrast, are also critical. For auditory displays (e.g., a tone or spoken warning message), intensity translates into loudness, with frequency as a critical variable. The frequencies that are contained in the background noise must be considered in deciding which frequencies should be contained in the alert. The relative intensity of a message (tone or voice) must always be measured in the environment in which it will be used.

Meaningfulness. Personally meaningful stimuli, such as one's own name, and culturally meaningful stimuli, such as the color red or a fire engine siren (both of which are associated with danger), will attract attention more efficiently than other stimuli of equal intensity. One exception to this, however, is if one of these "meaningful" signals is presented repeatedly without accompanying important information (as with false alarms). In this case, it is not difficult to learn to ignore a signal that previously attracted attention efficiently.

Ease of Interpretation. Symbols should be intuitive. For example, an arrow next to an aircraft ID pointing up, can easily be interpreted as an indication that the aircraft is climbing (as in a plan view display) or as a suggestion to climb that aircraft (as in an advisory display). However, all aspects of the symbol (e.g., color, exact shape, and the context in which it will be presented) must be taken into account when trying to predict how the symbol will be interpreted. For example, one of the first TCAS prototypes used an arrow to convey to the pilot the direction in which the pilot should fly. However, the arrow first used by one manufacturer of this display was **red** to convey the urgency of the alert. Even after training, some pilots felt that there could be instances in which pilots would be unsure as to whether a red arrow pointing up meant that they should climb or that the traffic was above them. The arrow was changed to a red or amber arc on the IVSI (Instantaneous Vertical Speed Indicator) with the instructions to the pilot to keep the IVSI needle out of the lit (red or amber) band. This provided a more consistent coding between the urgency of the alert and the required action.

Expectations and Context. Expectations and context are powerful shapers of perception. We are susceptible—particularly under high workload—to seeing what we expect to see and hearing what we expect to hear. Responses to a stimulus that occurs very frequently, or one that we expect to occur, will be faster than to one that occurs once every month. However, expectations may also lead to inaccurate responses, when what is expected is not what occurs. In many situations, particularly ambiguous ones, we erroneously "see" what we expect to see and we "hear" what we expect to hear. As illustrated in Figure 5-1, expectation is often guided by context. We read the second letter in row (a) as "B" and the second number in row (b) as "13" because those responses make sense in each context. On the basis of learning and past experience, we expect "B" to follow "A" and "13" to follow "12." Note, however, that the physical shapes of the "B" and the "13" are identical. If it were taken out of context, some people might read that physical form as "B," while others might read it as "13." There would be no right answer.

Figure 5-1. Effects of Context and Experience

An excellent example of the powers of expectation is seen in the videotapes that Boeing made of their original TCAS simulation studies. Here, the pilots had the traffic information display available to them and often tried to predict what TCAS was going to do. In one case with a crew of two experienced pilots, the pilot flying looked at the traffic alert (TA) display and said, "I think we'll have to go above these two guys" (meaning other aircraft). This set up the expectation for both crewmembers for a "climb" advisory. The crew started to climb when they received their first TCAS message, "Don't climb." The pilot flying told the pilot not flying to call ATC and tell them what action they were taking. Without reservation, the pilot not flying called ATC and said that, in response to a TCAS alert, they were climbing to avoid traffic. He also requested a block altitude. He then told the pilot flying that they were cleared to climb. Meanwhile, as the climb was being executed, "Descend" was repeated in the background over 25 times. Eventually, the pilot not flying said, "I think it's telling us to go down." The next thing that is heard on the tape is "[expletive], it changed. What a mess." Crash. (Boucek, personal communication.)

Anyone could have made a similar mistake. It is human nature to assess a situation and form expectations. In support of the pilot's expectation, and perhaps because of it, he didn't hear the first syllable, which was "don't" - he heard the action word "climb." The idea was then cemented. It takes much more information to change an original thought than it does to induce a different original thought.

Even when we do notice the difference between the expected and the actual message, there is a price to pay; it takes much longer to process the correct message when another one is expected than it does when the correct one is expected or when there are no expectations.

Practice. If the response is a highly-practiced one, then response times will be quicker and less variable than they will if the task isn't performed often. Eventually, highly practiced tasks may become "automatic"; that is, they will require much less effort and allow attention to be devoted to other tasks while simultaneously performing the "automated" one. Usually, this type of **automatic information processing** works in our favor. It allows us to talk and tune the radio while driving a car, instead of needing all of our attentional resources to drive safely. Under some circumstances, however, automatic processing can induce errors. When we try to perform a familiar task on equipment that is slightly different from what we are used to, performance will suffer. For example, after many years of driving a car with the horn in the middle of the steering wheel, you may find yourself pounding on the airbag in your new car wondering why the horn won't work. Similarly, if we try to use familiar equipment in a different way, errors are likely. For example, if a controller who was accustomed to a certain keyboard was now given the same keyboard but told that the same function keys now perform <u>different</u> functions (i.e., only the mapping of the function keys to the functions had changed), errors would be inevitable. When previous experience interferes with learning something similar, it is called **"negative transfer"**. Generally, the more subtle the differences between the "old" (equipment or procedures) and the "new" (equipment or procedures), the easier it will be to miss or forget the differences and the more prone to error we will be on the "new". Since designers may not be familiar with all of the ways in which controllers will use the system, it is up to the air traffic and human factors specialists to guard against such design-induced errors resulting from negative transfer.

User Confidence. Trust in the system may, or may not, develop with exposure to the system. Response time will increase with the time required to evaluate the validity of the advisory. Confidence in the system or a willingness to follow it automatically (regardless of whether or not this is desirable) will result in shorter response times. User confidence is a particularly critical factor in a controller's use of an automated decision aid. Because the controller is accustomed to making all strategic decisions (as opposed to allowing a machine to make the decisions) and because he or she is ultimately responsible for his or her actions, controllers can be expected to be wary of automated decision aids. This will result in longer response times than those observed with pilots' responses to executive systems.

Number of Response Alternatives. The decision component of response time will be affected by the number of response alternatives. In Ground Proximity Warning System (GPWS) operations, for example, once a pilot decides to respond, there is only one possible response—to climb. In TCAS II, there are two response alternatives: to climb, or to descend. With TCAS III, there are at least four alternatives: climb, descend, turn right, and turn left. Studies have shown that response times increase with the number of response alternatives. (See Boff and Lincoln, 1988, p. 1862, for a review.)

An Example. The factors that influence response time and response accuracy often interact. Without specific controls, it is usually impossible to identify how much of the effect (e.g., increase in response times) is due to specific factors. The following example illustrates how instructions to the test participants and workload level can affect response time on the same measure.

Conflict Resolution Advisory (CRA) software was designed to generated a solution to a potential conflict in an en route environment. At the onset of conflict alert, CRA provided a line of text in the tabular list (e.g., UAL 123 R 20, indicating a turn of 20 degrees to the right for UAL 123). As with any time-critical warning system, the algorithm must take into account the time required for the operator to use the system. In this case, the lag between the time the CRA message appears on the controller's scope and the time the aircraft begins to maneuver needed to be known, since this affects the number and type of potential resolutions. The time between the beginning of the controller's message to the end of the pilot's correct acknowledgement was determined by analyzing voice tapes of ATC communications (Cardosi and Boole, 1991). The mean (average) of this transmission and pilot response time was 11 seconds. The data collected by analyzing voice tapes provided only half of the equation.

Studies were still needed to measure the time required by controllers to use the new display.

A series of simulation studies was conducted to assess the controller response time to CRA. The first study used the prototype software, and the second study used the enhanced software. Controller response time was defined as the total time required to notice that an advisory was present, read and comprehend the text message, and decide that the resolution was usable. This decision task is critical and time-consuming since controllers have much more information available to them as decision-makers than the software does (e.g., about weather and restricted airspace).

In most studies, researchers are primarily interested in measuring the average response time. However, in this case, the validity of the suggested maneuver depends, in part, on the pilot and controller responding within the time parameter. In such cases, the response times at the 90th or 95th percentile (i.e., the response time at which 90 or 95 percent of the response times are at or below it) are more appropriate than is the average. Using the 99th percentile would be even more conservative, but would be too costly in that it would further reduce the number of potential resolutions.

Two approaches were used in assessing controller response time to the CRA display. In each case, the response time was measured from the onset of the conflict alert to the first indication that the resolution was usable. In the first study, controllers were asked to use the display as they would in actual operations. They were to use it at their discretion. Only those instances in which the display was used would be examined for response times. While this was an operationally valid approach, it resulted in too few response times for a valid study of the important delay parameter. Understandably, the controllers were reluctant to remove their eyes from the conflict on the scope to read the CRA message on the tabular list, particularly since a usable resolution may *or may not* have been present. The response times that were obtained using this approach averaged 18 seconds (Cardosi, Warner, Boole, Mengert, and DiSario, 1992).

A different approach to assessing controller response time to the CRA display was used in Part II of this first study. Controllers were instructed to perform their duties in controlling the simulated traffic as usual, but when conflict alert was activated, they were to read the CRA suggested maneuver and decide whether or not it was usable. This approach, which *required* the controllers to look at the CRA display, whether they wanted to or not, yielded quicker response times, averaging 13 seconds.

A second study, using the latter approach and the enhanced software, resulted in higher response times averaging 18 seconds (Cardosi, Burki-Cohen, Boole, Mengert, and DiSario, 1992). Why were the response times in the second test so much longer than those of the first test? Analysis of these test results revealed that this last group of controllers reported experiencing higher workload during the test, encountered a higher number of instances of conflict alert than did the controllers in the previous test, and had a higher percentage (28 percent as opposed to 11 percent for the first study) of conflict alerts where CRA was not able to display a resolution. Thus, even though the controllers were instructed to look at the display as it appeared, the extremely dense traffic situation, combined with the reduced value of the display, placed the added task of evaluating the CRA message low on the controller's list of prioritized duties.

It is impossible to calculate how much of the increase in human response times is attributable to each of these factors, since the study was not designed to do so. It is highly probable, however, that each of these factors had an effect and would be expected to do so in actual operations. A controller could be expected to look to a display that is useful and reliable much more quickly than one that is only occasionally useful. Higher workload will also add to the time needed to complete any additional task, such as look at a message on the tabular list. Finally, if the task is low on a list of priorities, either because of instructions to the controller to perform the task as time permits, or because of other factors, response time will be greater than it will if the task is a critical one.

Based on these data, what number could be used to account for pilot and controller response time? Rather than simply adding the 95th percentile of transmission and pilot response time to the 95th percentile controller response time (RT), we randomly combined each controller RT with each pilot RT for this sum. The test of the prototype software (the first test) resulted in a recommendation that 40 seconds be considered as the upper limit between the onset of the CRA display and the time the pilot makes an input into the aircraft's controls. (The assumption was made that a pilot would initiate an input into the aircraft's controls by the end of the verbal acknowledgement.) In the second test (of the enhanced software), the same calculations led to a total of 49 seconds. If such a test determines that the controller response time is too long to be operationally acceptable, then changes to the design must be considered either to allow the controller to process the information more quickly and/or to promote trust in the system. This trust must, however, be well-deserved by a system that is rarely in error.

Controllers can play an important role in educating designers about realistic time requirements for control tasks, especially when tasks overlap as they do in the field. If a design rushes or paces the controller, it should be considered operationally unsuitable. **The design must allow sufficient time for the controller to process the necessary information.** Information should be displayed long enough for it to be perceived and internalized. This means that, for example, blink coding should not be used to present information that the controller has to read and understand.

5.2 HUMAN INFORMATION-PROCESSING CAPABILITIES AND LIMITATIONS

High-level information-processing capabilities include perception, attention, language, memory, situation assessment, judgment, problem solving, and decision making. Perception is an information-processing capability because our perceptions are interpretations of what we see, hear, touch, and smell. The same physical stimuli can be perceived differently by different people or differently by the same people when presented in different contexts.

Chapters in this handbook focus on two major modes of perception for controllers: visual perception (Chapter 3) and auditory perception (Chapter 4). Information-processing issues related to speech, language, and communication are also addressed in Chapter 4.

Because human information-processing capabilities are limited in various ways, ATC system designs must be evaluated to ensure that the controller is not expected to play a superhuman role.

5.2.1 What kinds of ATC activities are made possible by the basic capabilities of the human as an information processor? How can system design best support these ATC activities?

Human information-processing capabilities support controller performance in the following areas:

- Goal setting and planning
- Paying attention
- Remembering (and forgetting)
- Perceiving and assessing the situation
- Making judgments and decisions.

This is not an exhaustive list of information-processing activities, but these are among the major cognitive or thought-based activities that guide and determine the controller's observable, job-related behavior.

Goal Setting and Planning. A key human capability is the ability to set goals and plan for the future. Planning is a complex, central-processing activity that controllers use to develop an overall plan for airspace management and subplans for managing the kinds of situations they typically encounter. Planning is designed to achieve ATC goals and to keep workload at manageable levels.

Over time, controllers develop a set of plans, sub-plans, and back-up plans for their airspace, depending on such factors as time of year, time of day, and specific conditions. For example, planning for the 3:45 rush will differ with weather conditions.

A skill developed by expert controllers is to avoid commitment to one plan if the available information is insufficient or if the situation can be expected to change quickly because of heavy traffic. In such a situation, expert controllers hold more than one plan active in working memory rather than committing prematurely to a particular plan. It is important to specify the information needed for ATC goal setting and planning. **Design products must be assessed for their ability to provide the information needed for planning. This includes information needed to reduce uncertainties in the ATC situation.**

Paying Attention. In a situation fraught with uncertainty and dependent on human alertness to maintain safety, the ability to attend to sources of information is critical. **Attention directs us to some information in our environment at the expense of other information.** Attention is a mechanism essential for reducing uncertainty and for maintaining situational awareness. A key human capability is our flexibility in switching attention rapidly from one source of information to another and from one task to another. Switching attention in this way is known as *timesharing*. When we seem to be performing two activities at the same time, we are actually timesharing, that is, alternating our attention rapidly between the two activities (McCormick & Sanders, 1982).

ATC task demands often require controllers to establish priorities for switching their attention between aircraft or groups of aircraft, while also timesharing their specialized capabilities for paying attention. For example, processing an airborne request for a change of altitude due to weather or turbulence will typically take precedence over formulating a clearance for a proposed departure. The controller must timeshare attention while processing the airborne request, however, because of the need to consider the total context of the ATC situation, including the status of other sectors. Communicating and coordinating with other controllers places high demands on attention and timesharing.

During the morning rush, the controller may be scanning the display (visual information processing), listening to air-ground communications (auditory information processing), thinking about how to manage aircraft about to enter the sector (planning), and entering data into the computer (transforming decisions into actions), all seemingly at the same time. When controllers are engaged in issuing instructions to aircraft or coordinating with other ATC personnel, they are likely to be timesharing attention with other visual, auditory, cognitive, and manual activities. We assume, in fact, that everything the controller does requires some level of information processing, whether conscious (controlled) or subconscious (automatic). **System design needs to accommodate the time required and provide aids to performing multiple, concurrent tasks.**

A warning to designers, however, is that controllers cannot be expected to timeshare the way computers timeshare. That is, controllers must be given more time to timeshare fewer messages than the computer must be given. While some combinations of tasks are easier to perform than other combinations (e.g., an auditory and a visual task would be easier than two auditory tasks), performing multiple tasks simultaneously is always done at a cost.

One study that investigated the cost of multiple tasks sat people in front of a display box and instructed them to press a button whenever a light came on (Johnston and Heinz, 1978). The light came on at random intervals. Subjects simultaneously listened to a tape of excerpts from *Reader's Digest* articles. Their task was to listen to the tape and press the button when the light came on. The participants also had to answer simple true/false questions about the passage at the end of each trial. These questions were asked to ensure that the subjects attended to the tape and didn't neglect to attend to the button-pressing task. Adding the task of listening to a message raised the time required to respond to the light from 320 msec. to 355 msec. Thus, there was a small, but statistically significant, rise in response time for a very simple task (i.e., a button press) when another simple and unrelated task (i.e., listening) was added to it. As the experimenters made the listening task more difficult (e.g., attend to one of two stories), response time rose with the difficulty of the task. For example, it took an average of 387 msec to press the button in response to the light as subjects tried to pay attention to one of two very different messages (i.e., on different topics with one spoken by a man and one spoken by a woman), and an average of 429 msec to respond to the light as subjects tried to attend to one of two very similar messages (i.e., with same sex speakers and similar content).

These experiments demonstrate three things. First, the time required to conduct even the simplest task will increase as other, even simple and unrelated tasks, are added to it. Second, the more difficult the added task is, the higher the attentional cost due to the additional burden on the attentional mechanism. Third, this attentional cost can be measured. On the average for these subjects, it took 320 msec. to simply press the button when the light came on without any information being broadcast to the ears. If a stimulus (e.g., a warning light or text message) appears directly in front of a person, response time to it will be faster than if eye movements are required to fixate, or focus on, the information. Similarly, if the stimulus appears within the person's visual field, but in the periphery rather than at the fixation point, response time will be lower than when an eye movement is required, but higher than when the target appears at the fixation point. While we usually move our eyes when we shift attention, this is not always necessary. We can shift our mental focus, or internal attention. Even when shifting internal attention does not involve eye movements, it does take time. The time required to shift internal attention increases with the distance from the fixation point, and internal attention travels at a velocity of about 1 degree per 8 msec. (Tsal, 1983).

It is important to evaluate designs for the effects of timesharing required. This includes the need to switch from one display to another and to perform new tasks in addition to other duties.

> "In general, the causal factors that resulted in greater loss of separation were those which would involve reduced situational awareness by the controller."
>
> Rodgers and Nye, 1992, p. 2

Remembering. Capabilities in memory allow the expert controller to organize and retrieve large amounts of highly detailed information, to timeshare multiple tasks, and to make decisions quickly on the basis of past experience. The limits of short-term, working memory and the need to conserve "space" there make it incumbent upon designers to limit demands on that capability. Overloading short-term memory risks loss of **situational awareness.** Providing aids to controller memory is a basic objective of user-centered ATC system design. **The design should limit demands on short-term, working memory and provide appropriate memory joggers (e.g., prompts and cues).**

The limits of short-term memory are discussed in detail in Section 5.4.2 of this chapter.

Perceiving and Assessing the Situation. Achieving situational awareness is a process that involves the capabilities of attention, perception, and memory. The controller must pay attention to the relevant sources of information to make perception possible; integrate information from various sources, including personal knowledge in long-term memory; detect differences between the current situation and the expected situation; and be able to predict the consequences of possible control actions, including doing nothing except allowing the situation to develop.

From the time the controller receives a relief briefing until s/he gives a relief briefing to the next controller coming on duty, s/he is thought to be building and maintaining a "picture" of the ATC situation (e.g., Whitfield and Jackson, 1982). All of the information presented to the controller contributes to that picture, which is another term for situational awareness. The quality of the controller's grasp of the situation is, itself, an element in the ATC situation because it determines the quality of the controller's decisions. Design must focus on providing information in a form that can be readily assimilated into the controller's on-going understanding of the situation and that will support the most correct, complete situational awareness possible.

Judgment and Decision Making. In the face of uncertain information or unpredictable outcomes, judgment is a key capability of expert controllers. When several options are possible, judgment, based on training and experience, guides the choice of an appropriate option. Weighing the options and making choices are central information-processing activities that lead to ATC decision making.

A key point is that design products exist to complement and extend human abilities. In order to do so, these products must be compatible with information-processing strategies and job strategies. By the same token, they should not replace human involvement or overcompensate for human limitations. Going too far in the direction of replacing the role of human consciousness runs the risk of degrading the person's ability to build and maintain situational awareness.

For example, a potential highly-automated ATC system of the future might present the controller with solutions to every potential separation violation between aircraft as potential violations are detected. If this system worked perfectly, (so that the machine-generated solutions did not need to be carefully evaluated by the controller), we could expect several effects on controller performance. First, situational awareness would probably deteriorate, since maintaining "the picture" would not be as critical a task as it was without the system (since no separation violations or mid-air collisions are possible as long as you do what the machine says to do). Second, the controller's ability to plan might be affected, since strategic planning would not be as important, and conflict avoidance would become an obsolete skill. Imagine that after several months of perfect performance, the system begins to malfunction; the system no longer provides resolutions to every conflict alert, and the resolutions that are provided are no longer perfect and must be evaluated by the controller. How much would the controllers' skills have deteriorated through disuse? This is not known, nor can it be predicted with available data. While this hypothetical example may seem far-fetched, it raises several issues that must be of concern to developers of highly automated aids for controllers. These issues and others are discussed in Chapter 6, Issues in ATC Automation.

5.2.2 What are some ways in which system design can support information processing in ATC?

Information processing can be thought of as cognitive behavior, directed toward achieving rational objectives and goals. Further useful distinctions can be made within this area if we borrow a distinction that has been made between three levels of observable behavior: skill-based, rule-based, and knowledge-based behavior (Rasmussen, 1983). These categories apply to information-processing as well as to observable behavior.

Skill-based behavior occurs in highly-familiar situations. Skills are highly practiced and are performed automatically, with little or no conscious attention. The performance of an expert typist, pianist, or athlete is highly skill-based. Similarly, a large portion of the expert controller's observable behavior is skill-based: visual search, cursor (slew) positioning, and command entry.

From an information-processing perspective, there are two arenas for the controller to master: the ATC domain and the interface with the computer. Through training and practice over time, controllers develop skills in both arenas. A design objective should be to make the controller's window onto the ATC situation (i.e., the computer-human interface) as transparent as possible—as intuitive and natural as possible—so it does not distract controllers from the "real" job.

Skills develop over time, and become essentially automatic with intensive practice, on highly-repetitive tasks. If skills are not maintained through continuous practice, performance will become "rusty." A concern about highly-automated ATC systems is that some method or mechanism needs to be provided to maintain controllers' operational skills in an environment where skills can decline through disuse.

Rule-based behavior occurs when well-learned procedures are executed under conscious attention. Situation assessment leads to a recognition of which procedures to apply to particular familiar situations. Procedures are learned and maintained in long-term memory. Execution of procedures may involve some skill-based behavior.

Clearly, ATC performance is highly rule-based. It is governed by detailed procedures, letters of agreement between facilities, and so forth. In some cases, pre-defined back-up procedures exist. For example, the Severe Weather Avoidance Program (SWAP) provides alternative procedures for re-routing aircraft under various weather conditions and procedures for operating under the Direct Access Radar Channel (DARC) system, which serves as a backup to the current Air Route Traffic Control Center (ARTCC) computers.

To some extent, design of ATC procedures depends on the allocation of functions to the controller and the computer. Procedural design also depends on the design of the dialogue or interaction between the controller and the computer. The link between requirements and procedures is one that must be tracked and documented as a basis for development of procedures that are consistent with the objectives for the new system.

Knowledge-based behavior is produced when existing procedures are inadequate to handle unfamiliar, rare events. What is required here is creative problem solving and decision making based on knowledge of the problem domain and past experience. At a minimum, problem solving under conditions of high uncertainty involves the following kinds of information-processing activities:

- Information gathering
- Situation assessment
- Generation of alternative plans and courses of action
- Evaluation of the alternatives
- Selection of a plan for implementation.

This kind of problem solving is time consuming and highly demanding of the problem solver's analytic resources.

In an ATC context, the "knowledge" referred to is everything the controller knows about ATC and the separation of aircraft. "Going back to basics," for example, is a knowledge-based strategy that may be employed in an unusual situation not covered by published procedures. Knowledge-based problem solving becomes necessary when there is no similar or identical problem "pattern" stored in the controller's memory. Whereas recognition-based problem solving is essentially a skill developed from experience, knowledge-based problem solving requires time and effort, with no assurance that a solution will be found within time constraints.

As discussed in the following section (5.2.3), system design can support skill-, rule-, and knowledge-based processing in various additional ways.

5.2.3 How can an understanding of information processing help to specify and evaluate new ATC systems?

The controller's information-processing activities are crucial to the productivity and performance of the overall ATC system. These activities provide the basis for situation assessment and selection of control actions. A general understanding of these activities can help define operational requirements and specifying classes of job aids to support operational performance.

Building a model of the controller's cognitive operations can help in formulating and resolving questions. The following questions should be asked about each combination of tasks that the controller is expected to perform:

- **What information is needed?**
- **When is it needed?**
- **In what form is it needed?**
- **How should it be sequenced?**

Once this information is acquired, we can ask the following questions of the proposed design product (especially information displays):

- **What information is provided?**
- **Is all necessary information provided?**

- **Is the information provided at the right time?**
- **Is the information sequenced appropriately?**

A model or schematic representation of the controller's information-processing activities can provide the basis for defining the controller's information-processing tasks and identifying the information needed to support those tasks. Task definition is one of the key activities in the acquisition process.

When analyzing ATC tasks, the "skill, rule, knowledge" distinction, discussed in the previous question, can help in determining the kinds of information needed for each kind of task. Skill-based performance is supported by information that permits the controller to update internal images or cognitive maps of both the workstation environment and the ATC situation. Smooth, integrated visual search, for example, is supported by consistent placement of text labels and symbols in data blocks. Skilled navigation between visual displays is supported by consistent location of related or identical fields from one display to another (Woods, 1984).

Rule-based performance is supported by information that leads to speedy, accurate recognition of a situation that requires intervention, be it a problem of equipment status or a developing conflict situation. Recognition of the situation then triggers the stored rules (or plans) for dealing with it. **Linking ATC procedures with information requirements will help to ensure that displays can be interpreted quickly and accurately.**

When a particular situation is not recognized as familiar, the controller needs information that will help in identifying the nature of the problem (if there is one), identifying the immediate goal (e.g., aircraft separation), and planning control actions to achieve the goal (e.g., a change or changes in altitude, speed, or route). Information to support planning would help the controller evaluate alternative plans by, for example, predicting the effects of particular control actions on the evolving ATC situation.

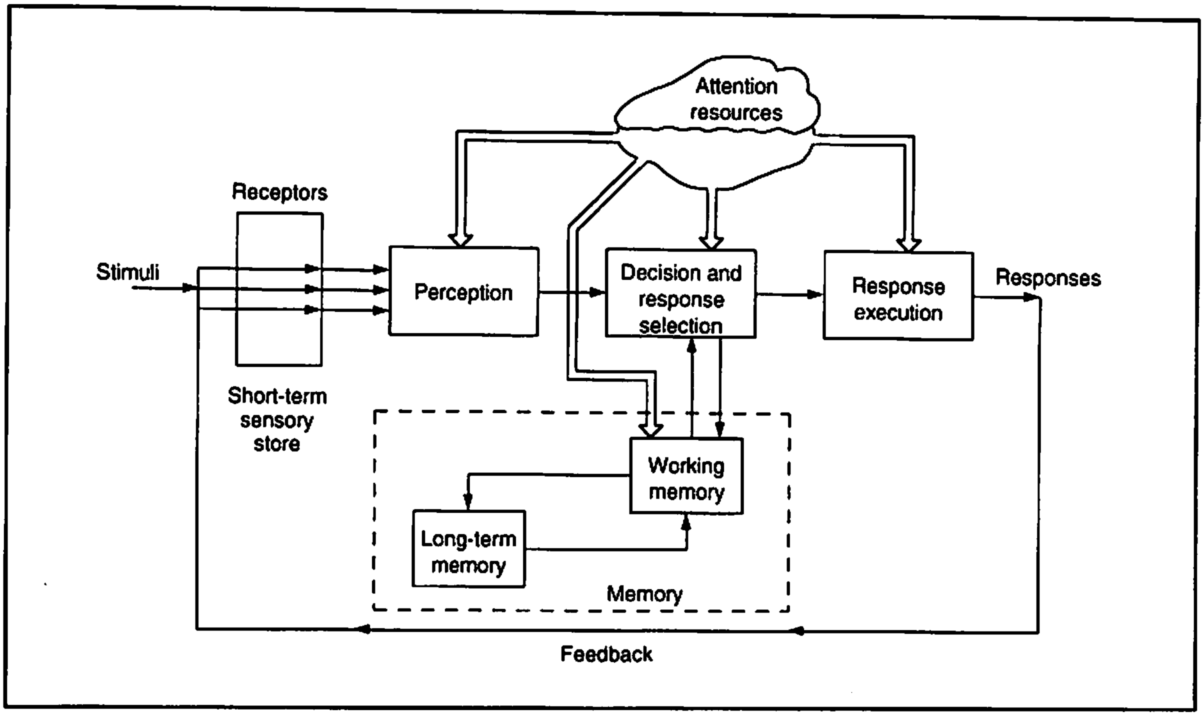
Unaided prediction of results from many interacting factors is difficult, even for highly trained personnel, because the amount of raw data, the number of relationships, and the transformations required to determine outcomes quickly overwhelm the limits of working memory. This is where *predictive displays* can be most helpful. For example, using the Future Situation Display, the controller will be able to try out various control actions and examine their effects in a simulated version of the "live" traffic situation. Such a capability should be designed to complement the controller's mental projection of future situations. A "what if" function that allowed controllers to try out solutions to problems could aid in strategic planning. Such a function could enhance (rather than degrade) a controller's own planning skills. The positive value of predictive capabilities can be compromised, however, if they place additional memory demands or other burdens on the controller. **The positive and negative effects of any design product, including predictive displays, should be assessed systematically.**

Having a model of ATC information processing can also help in formulating evaluation plans. It is important, for example, to evaluate how a design product will support the controller's performance of the information-processing tasks modeled in task descriptions.

5.2.4 What are the purposes and components of an information-processing model?

If we could exactly model human information processing, we would have a precise map of all the neural connections in the brain and be able to depict all the possible states of the brain. Given the current state of our knowledge, this is not possible. The purpose of an approximate modeling activity, however, is to help us think at higher levels about whatever it is we are trying to understand, in this case, information processing in air traffic control. The components of a contemporary model of human information processing are illustrated in Figure 5-2.

Figure 5-2. A Model of Information Processing
From "Engineering Psychology and Human Performance" by Christopher D. Wickens. Copyright 1984 by Scott, Foresman, and Company. Reprinted by permission of HarperCollins Publishers.



Sensory Storage. As illustrated in Figure 5-1, incoming physical energy is processed first in short-term sensory storage. For example, light stimulates or triggers receptor cells in the retina, and this pattern of neural energy is transmitted to the brain. As indicated in the figure, sensory processing is assumed to be essentially automatic. Research indicates that items are held in the sensory store for up to one second for visual input and up to 3.5 seconds for auditory input.

Information in the sensory store is encoded or represented in terms of physical features. For example, the reader's visual sensory store encodes this page of text as a pattern of dark lines on a white background. In the current ATC system, the air traffic controller's visual sensory store encodes the light energy from the plan view display (PVD) as a pattern of bright green lines and symbols on a duller green background. (Different levels of brightness are used for information-coding purposes.) For the reader, further processing is required to interpret the patterns on this page as letters or words. For the controller scanning the PVD, further processing produces meaningful targets, data blocks, weather cells, and other information.

Pattern Matching. Sensations held in sensory storage receive further processing in the second stage of human information processing: perception or pattern matching. In this stage, which normally involves little or no conscious processing, we make sense of our sensations, within the context of our own training and experience. Here, meaning emerges as patterns of lines, shapes, or sounds are recognized (or perceived), for example, as words, familiar faces, or specific messages. Pattern matching is an efficient but extremely complex process:

> "This pattern recognition process involves mapping the physical codes of the sensory store into semantic or meaningful codes from memory. This mapping is very complex in that many different physical codes may all map to a single memory code (e.g., a, A, the sound "a"), and a single physical code may map to different memory codes" (Wickens and Flach, 1988, p. 113).

Through this further skill-based processing, the controller is essentially "reading" the ATC situation, using inputs from various sources to build and maintain a dynamic mental image or "picture" on which to base decisions. Controllers also use this picture as the basis of projecting future situations, for example: what will happen if I slow down an aircraft? Will some other problem develop? What will happen if I take no immediate action? What will the situation look like in two minutes, in five minutes?

Just as readers have learned to "decode" text, controllers have learned to decode the PVD. Incoming physical energy conveys patterns that are decoded as radar returns, other meaningful symbols and codings, verbal communications from pilots and other ATC personnel, and tactile feedback from input devices. These meaningful items are then grouped together into larger units, such as clusters of departures, arrivals, and overflights; pilot reports (PIREPs); pointouts from other controllers; and so forth.

Over time, these decoding and grouping activities become second nature to the controller.

The controller is also surrounded by the facility environment, which impinges on the senses in the form of background noise levels, different levels of illumination (from the workstation to facility lighting), glare and reflections, temperature, and air quality. All of these inputs—and many others—impose demands on the controller's information processing capabilities. (Workstation and facility design are discussed in Chapter 9, with recommendations provided on levels of environmental factors.)

An important design goal is to limit sources of irrelevant sensation or distraction from the working environment.

In the visual or auditory pattern-recognition process, if an exactly matching pattern is not found, the possibility of error arises (as in mistaken identity or confusion between two similar callsigns), or memory is updated to account for the differences detected. Pattern recognition can occur without conscious attention, as when we drive along in "automatic mode" while thinking about other things. The fact that we are processing the environment automatically becomes apparent when, for example, our attention is captured by something new, such as a sign warning of road construction ahead.

In ATC, for example, specific patterns of light are perceived "automatically" as targets representing controlled aircraft, data blocks, aircraft history trails, and weather cells. Specific patterns of sound are perceived as messages from various pilots, which the controller must often sort out against a noisy background. Sound patterns are also interpreted as cues to pilots' experience, national origin, and level of concern, which the controller may consider in making control decisions.

Under routine conditions, all of these perceptions occur spontaneously, without a great deal of conscious attention. In an unusual situation, when the sensory pattern does not match a known or expected pattern, a greater investment of conscious attention is required. Errors can occur when a sensory pattern is mistaken for an expected pattern, which may differ in only small details, such as two very similar callsigns, for example, UA246 and UA264. **It is important for displays to support accurate pattern matching. The probability of confusing symbols or alphanumerics should be minimized.** These problems can be avoided by following the guidelines for displays described in Chapter 7, Computer-Human Interface (CHI) Considerations.

Decision and Response Selection. In the decision and response-selection stage, we decide what to do about whatever it is that we have recognized. Here we exercise judgment in weighing the potential costs and benefits of several available options:

- Store the information for later use.
- Integrate the information with other known information.
- Respond physically or verbally.

In ATC, a decision not to issue or to delay issuing control instructions is often the result of this active process of considering options. For example, in obscured weather conditions, a controller delays a clearance for a taxiing aircraft to cross a parallel runway until the aircraft's position on short final is accurately determined.

Air traffic control requires a great deal of integration of information from various sources. For example, in determining whether or not aircraft are potentially in conflict, the ARTCC/HOST controller may extract and integrate the following data and information from the PVD and the Flight Progress Strip:

- Plan View Display: Primary Target, Target/Track Position Symbols, Track History, Velocity Vector, Target Halo, Reported Altitude, Assigned Altitude, Time, Aircraft ID, Ground Speed, Destination, and Scratch Pad Information

- Flight Progress Strip: Flight ID, Aircraft Type, Assigned/Requested Altitude, Route Information, Posted Fix, Next Posted Fix, Estimated Ground Speed, True Airspeed

Plans for the Initial Sector Suite System (ISSS) called for many different visual displays that a controller might consult in problem solving and decision making. Additional visual and auditory displays are associated with the new ATC communications system, the Voice Switching and Control System (VSCS). Input devices also present visual, auditory, and tactile cues and feedback to the controller.

Information-processing tasks of information integration and projection of possible future outcomes are at the heart of ATC as it is currently practiced. They are performed successfully to the extent that accurate information is displayed to the controller and interpreted quickly and accurately by the controller's information-processing system. Training and experience affect accurate interpretation, as do physical constraints such as fatigue and/or sensory capacity. The selection and presentation of information, by design, can support or impair the speed and accuracy of the controller's interpretation.

Response Execution. If the controller has made a decision to initiate a response to the recognized pattern of physical energy, this intention is translated into a precise sequence of manual actions and verbalizations during the response execution stage. As pictured in Figure 5-1, the feedback loop communicates any responses to the controller's sensory store for interpretation by the pattern-matching process. In this way, through self-monitoring (which can be faulty), we all detect or fail to detect errors in carrying out our responses (e.g., errors in data entry or verbal communications). **One**

characteristic of an operationally-suitable system is that it helps controllers detect and correct their own errors. An example of this is the confirmation question that is presented after the "delete" key is pressed on a computer (e.g., "Are you sure you want to delete the message?").

The efficiency and effectiveness of human information processing is affected by many factors, which are discussed later in this chapter. **The key point here is that design of displays, controls, and workstations for controllers must be grounded in knowledge of human capabilities and limitations and must be supportive of the controller's natural, phased approach to information processing.** For example, because we know that it is difficult for people to extract meaning from a cluttered background, a major design goal should be a demonstrable gain in the controller's ability to distinguish meaningful sensory input (e.g., radio communications, radar returns) from non-meaningful input or "noise" (e.g., static on the frequency, visual display clutter).

Other major design goals should be to provide support for pattern matching; support for mental integration and projection of ATC information; support for "trying out" possible control actions; support for response selection; and feedback for self-monitoring. Designing the system software to capture errors is a way of supporting controller productivity. If these design goals are to be achieved, they must be stated clearly as requirements in the system specification. Design products must then be evaluated against these requirements.

5.3 ATTENTION

The controller's attention is a key factor in detecting and resolving ATC problems. Attention is a familiar word and concept, but the underlying processes are complex and, to some extent, not well understood. This section discusses the role of attention in human information processing, with a focus on its role in ATC. Evaluation tips suggest ways in which design can compensate for natural limits on attention.

5.3.1 What is the role of attention in human information processing?

Attention may be considered as a limited resource which is drawn upon to process large amounts of information in support of our performance of daily activities, including our job tasks. Two models of attention currently in use assume that human information processing is fueled by a limited supply of attention. One model, the single-channel model, assumes that there is one pool of attention, while the second model, the multiple-resources model, assumes that there are several different pools of attention, each specializing in distinct mental operations.

In the single-channel model, the resources of attention are pooled in one "reservoir." In this model, perceptual processing is relatively automatic, but attention can be paid to only one item of information at a time (Kahneman, 1973; Wickens, 1984). The single-channel model includes an information-processing "bottleneck," which slows the processing of items after they have been perceived. The single-channel model, however, is unable to explain why some tasks can be time-shared (performed together) with no impact on the performance of either task.

A second basic model is the multiple resources model of attention, as outlined in Table 5-2. This model represents the view that people have several different attentional capacities, which can be drawn upon in parallel by the demands of different task types (Wickens, 1984; 1992). According to this model, as two tasks require more sharing of a person's resources, the more one task will interfere with the other's performance. **A recommendation based on this model is that ATC tasks that require timesharing should draw upon different resource pools in order to make the most efficient use of controller attention.**

Table 5-2. Multiple Resource Model
Adapted from "Information Processing" by C. D. Wickens and J. M. Flach, 1988. In *Human Factors in Aviation* by D. Nagel and E. Wiener (Eds.), p. 146. Copyright by Academic Press, 1988. Adapted by permission.

Model Dimensions	Examples
Processing modalities	
Auditory vs. visual	ATC communications and auditory alerts vs. instrument scanning or out-of-cockpit monitoring
Voice vs. manual control	ATC communications vs. flight control or keyboard entry
Processing codes	
Verbal vs. spatial	Processing navigational coordinates, radio frequencies or understanding conversations vs. tracking or maintaining spatial orientation
Processing stages	
Perceptual and cognitive (working memory) vs. response	Instrument scanning, rehearsing, listening, calculating, and predicting vs. speaking, switch activation, or manual control

Current thinking suggests that people use both single-channel (sequential) processing and multiple-channel (parallel) processing (Wickens, 1992b).

More practiced and familiar tasks require less attention than new or unfamiliar tasks. Highly practiced tasks can require so little attention that they can be thought of as being performed automatically. Thus, consistent practice that results in expertise reduces the resource demands experienced by the novice.

One price of this "automaticity," however, can be substitution errors (Reason, 1990). If, for example, the controller is handling both UA246 and UA264, along with numerous other aircraft, a flight level change intended for UA264 may be issued inadvertently to UA246 and not be caught by the controller, whose conscious attention may be directed toward a developing conflict situation.

As illustrated in the previous example, errors committed in "automatic" mode can go unnoticed. Further, there may be no recall or inaccurate recall of the action taken, simply because of the highly practiced, automatic nature of the response. For example, an approach controller may not recall confirming an altitude change from the wrong aircraft.

Because much of what the controller does is "automatic" (e.g., responding to an aircraft's initial check-in), and because the controller's attention is often directed in many places at once, **it is important that the design direct the controller's attention to areas where it is needed by means of alerts, coding, and emphasis techniques. Overuse of any one technique will severely reduce its effectiveness.**

5.3.2 What are the limits of attention?

Analyses of controller errors come back, again and again, to the need to pay attention in performing ATC duties (e.g., Office of Aviation Medicine, 1985) The controller is literally surrounded by demands for attention, yet, attention is a limited commodity. First, we will examine some of the capabilities and limitations of attention. Later, we will discuss ways in which design can help the controller assess demands and allocate attention when and where it is needed.

Selective Attention. Sometimes, we choose to pay attention to certain sources of information and not others. We do not, however, always select the most relevant sources of information at the most appropriate times. This is reflected in our less-than-optimal ability to sample sources of information in complex work environments. For example, the pilot who checks airspeed but ignores the altimeter shows less-than-optimal sampling behavior that can have disastrous consequences.

A major ATC function, monitoring for conflicts, is subject to the limits of selective attention. People are notoriously poor monitors on classical "vigilance" tasks, which require detection of new targets. After 30 minutes on a pure vigilance task, performance declines rapidly (e.g., Mackworth, 1948). Recent research in an automated environment indicates that the ability to monitor a highly-automated system declines significantly after about 20 minutes (Parasuraman, Molloy, and Singh, 1993). Since monitoring in the current ATC system is not a strict vigilance task, alertness can be maintained for up to one hour (Thackray and Touchstone, 1980). But, in new systems, if controller's job is simply to monitor the automation, attention is likely to degrade after 30 minutes, even for highly-trained and highly-motivated controllers. **Beware of any allocation strategy that places the controller purely in a monitoring role for extended periods of time (30 minutes or more).**

Focused Attention. Another issue in attention is how easily we can focus on one demand and ignore the distraction of other demands. The difficulty of focusing attention comes from the fact that items of information that originate from many different sources compete for our attention despite our efforts to ignore them. The competing information may take the form of extraneous conversation or display clutter. The limits of focused attention are reached when attention is unintentionally "captured" by informational demands that we meant to shut out.

Positional differences help us to attend to some information and ignore other information. For example, it is easy to read text presented in front of us and ignore the scene out the window behind us (unless the scene is particularly interesting). We are physically incapable of focussing on something close and directly in front of us and, at the same time, focus on something in the periphery or in the distance. Gross physical differences also help us selectively attend to information. This is true for both visual and for auditory information. For example, it is easy to zero in on highlighted information, if the information is a different color or is sufficiently brighter than the other text.

In focusing on auditory information, it is easier to focus on one of two voices if the two voices are widely separated across a range of pitches, for example, if the speakers are a male and a female. If both speakers are male, or if both speakers are female, the similarities of pitch make it more difficult to focus on either speaker. If two messages deal with the same content material, it is more difficult to avoid being distracted by the unwanted message. Thus, pitch, gender, and meaning have been found to help focus attention. **To ease focusing on competing auditory messages, differences in pitch, gender, and meaning should be made highly recognizable through the design of voice communications systems.**

Under time pressure, extremely high workload, or other stressors, attention can become too focused on one aspect of the overall situation. This is known as *cognitive lock-up* or *cognitive tunnel vision.* In this condition, the controller simply will not see information displayed on the PVD unless it is in the area on which s/he is focused. **The design should help the controller to focus attention when appropriate, to avoid distraction, and to counteract or compensate for cognitive tunnel vision.**

Divided Attention. If we need to pay attention to more than one kind of information at the same time, we attempt to divide our attention. For example, the controller may be trying to focus visual attention on the information in one data block, while at the same time trying to focus auditory attention on air-ground communications from the pilot of a different aircraft. Divided attention allows the controller to process different kinds of information at the same time.

Controllers develop strategies for dealing with numerous requests that require divided attention. For example, using a strategy based on criticality judgments, a controller may try to segregate information related to an aircraft emergency from other, less critical ATC requests. Thus, ability to judge criticality is a key ATC skill underlying appropriate division of attention. This skill can be supported by design. **Design should help the controller assess the relative criticality of situations that are competing for ATC attention. Is the controller alerted adequately to critical situations with enough lead time to formulate and execute appropriate responses?**

The danger of divided attention is in its demands for resources. The limits of divided attention are reached when we cannot divide our attention between all the sources of information that we want to process. Since attention is a limited resource, it is important to keep some of it in reserve to deal with unexpected events.

Several interesting experiments demonstrate the limits of divided attention and our limited ability to process information while performing complex tasks. In one such study, researchers showed videotapes of games to observers and had them perform simple tasks (Neisser and Becklen, 1975). On one videotape, three men bounced a basketball back and forth to each other. The observers' task was to count the bounces. Then, the researchers showed a tape of two people playing a handslapping game. The observers' task here was to count the number of hits. If either task was performed alone, counting accuracy was near perfect. When the two tapes were superimposed, it was still quite easy to count either the number of ball bounces or the number of hand slaps. Trying to count both at the same time, however, was quite difficult. It was so difficult that the observers failed to notice the "odd" events of the ball disappearing or the men being replaced by women.

This is one example of the filtering of information. We can attend to and process complex information quite efficiently. However, if the task is attentionally taxing, we may not process all of the information available to us.

Design to Support Attention. If we make the reasonable assumption that well-trained, expert controllers do not purposely allow their attention to slip, there is not much value in blaming controller errors on failure to pay attention. Doing so is like blaming someone for aging as the years go by. Lapses of attention are to be expected under circumstances that overwhelm human limitations in attention. If any blame is to be placed, it must be placed on the design of systems that overload or underload human information-processing capabilities (Wickens, 1992b).

Although lapses of attention are natural outcomes, they can be defended against by design. While it is useful to alert controllers to error-inducing situations (such as similar call signs and readback/hearback errors with clearances containing altitudes of 10,000 and 11,000), simply reminding controllers of the need to pay attention is futile. Controllers know that they need to pay attention, but they are human and subject to human limitations. Features of display design can, however, help controllers maintain attention and support effective display-scanning strategies.

Visual coding techniques, such as color, brightness, or blinking, can be used to attract the controller's attention to unusual situations or potential problems. Visual coding alone, however, may not be sufficient, overuse of visual coding will be self-defeating. If the controller is focused on a particular situation, s/he simply will not detect unusual visual codes. Even a blinking object (such as a hijack indicator) or a conflict alert (CA) in the data block may not attract attention under certain circumstances. In some instances, the addition of an auditory alert may be necessary. With future technology, it may be possible to place visual alerts within the controller's focus of attention no matter where the source of the problem is located on the display.

5.4 MEMORY AND FORGETTING

The controller's memory is another key factor in accomplishing the ATC mission. Because memory is so critical a capability, this section tries to point out ways in which design can help to compensate for memory limitations.

5.4.1 What is the role of memory in information processing?

Memory plays many roles in information processing. As we have seen, the sensory store, which is a kind of memory, preserves sensations from the environment for brief periods of time, long enough for the attentional filter to select certain sensations for further processing. Memory plays an enormous role in pattern matching. (In this sense of the term, faces are patterns, configurations of aircraft on the PVD form patterns, and spoken sounds form patterns that we recognize as meaningful.) Our expectations, that is, our stored patterns, based on past experience, are held in memory and to a large degree determine what we perceive. Surprise is a strong lack of agreement between the expected pattern and what actually occurs.

Expert decision makers know what to do in a given situation because their perception of the situation (as it is represented to them by the display system) triggers rules or plans, which are additional types of patterns stored in long-term memory. Thus, all the stages of information processing draw directly upon memory. The controller's memory is a critical element in the ATC system.

5.4.2 What do we mean by "working memory" as compared to long-term memory?

Working memory (or short-term memory) has been described as the mind's scratchpad. It has also been equated with consciousness (Martindale, 1981). The idea is that working memory holds active information, both the products of perception being attended to and whatever elements of long-term memory are currently active.

In working memory, we manipulate these elements, performing pattern matching, mental arithmetic, mental projection of situations, and many other forms of active processing. For example, working memory supports the controller's building and updating of the current and projected assessments of the ATC situation. Maintaining situational awareness is a key function of working memory. According to some research, assessing a complex situation in working memory requires perception and integration of information as well as projection of consequences into the near future (Endsley, 1991).

To reduce the amount of information integration that the controller has to do in his/her mind (working memory), information from multiple sources should be appropriately integrated, to the extent possible, before being presented to the controller. Sub-systems that provide different kinds of information should be integrated with the total system so that all the relevant information can be presented to the controller in one

place. Of course, if task performance requires separate pieces of information, the controller should not be required to separate out what is needed from an integrated display. Information should always be presented in a directly-usable form.

Working memory is constrained by a limited capacity to keep more than a few clusters of information active, as discussed further in Section 5.4.5. Knowing when to clear or update working memory is a necessary skill in ATC, where the controller must constantly be purging information that is no longer necessary and updating awareness of the current and projected traffic situations. Judicious forgetting is critical to the efficiency of working memory. However, something we want to remember can be "lost" from memory almost as easily as extraneous information, because of a breakdown in the retrieval system.

In contrast, long-term memory is the storehouse of knowledge, the mind's database. There is no known limit on the capacity of long-term memory. Knowledge stored in long-term memory includes everything we know: for example, our personal histories, cultural history, mathematical principles and procedures, ATC principles and procedures, science, engineering, gardening, sports and how to play them, and the names of TV shows. Knowledge can be either declarative or procedural. Knowing that aircraft come into a sector at certain altitudes and have to leave the sector at other altitudes is declarative knowledge. Knowing how to implement those changes is procedural knowledge.

Another useful distinction can be made between two dimensions of long-term memory: semantic memory and episodic memory. Semantic memory is our knowledge for meaningful information, which is not tagged in memory for a time or place when we learned it. For example, we know that George Washington was the first President of the United States, but we probably do not have a specific memory for when or where we learned this. Episodic memory is our memory for events in their sequence of occurrence, usually associated with some form of tagging for time, place, other people present, and so forth.

The controller getting off shift, for example, can recount from episodic memory the aircraft involved and the series of control actions taken for any troublesome situation that occurred. Aircraft that were not involved are remembered less clearly. Episodic memories of highly specific, critical ATC situations, such as a near mid-air collision, can persist for years, just as do our episodic memories for critical personal experiences.

5.4.3 How is knowledge organized in long-term memory? How can design make use of the controller's natural organization of ATC knowledge?

Although there is no definitive answer, evidence indicates that knowledge is organized in conceptual structures, which are often called mental models. The air traffic controller's "picture" is an example of a mental model (Carroll and Olson, 1988; Mogford, 1990; Whitfield, 1979; Whitfield and Jackson, 1982).

Knowledge structures allow us to access specific items of knowledge efficiently. For example, if asked to list all the "red fruits" we can think of, we go to our memory structure for fruits and start to look for those that can be red: apple, cherry, raspberry, and so forth. If asked to name the best times and places to fish for sea trout, the anglers among us go to their fishing memories built up over past experience, and find sea trout among the linkages as a basis for responding to the question. Recall breaks down when we are unable to access the appropriate knowledge structure.

Research has shown that controllers organize their knowledge of ATC into conceptual structures (Harwood, Mogford, Murphy, and Roske-Hofstrand, 1991). As might be expected, developmentals seem to structure their ATC knowledge in ways that resemble the conceptual structures of their instructors, although further research is needed to confirm this inference. User-interface design should be consistent with the ways in which effective controllers naturally organize ATC information. User-interface design should also facilitate development of appropriate and useful conceptual structures by trainees.

Thus, in design evaluations, controllers who evaluate the new system should look for anything that makes them perform more slowly or less accurately than usual. This does not mean, however, that evaluators should sit down "cold" at a new workstation and be expected to attain instant mastery. It is important for evaluators to be trained to a consistent level of expertise in the use of new design features before performance measures are taken. Valid comparisons between old and new systems, or between alternative versions of new systems, cannot be made if levels of expertise differ across evaluators.

Evaluators can and should differ in years of ATC experience, but they should be thoroughly trained in the features to be evaluated so that performance problems cannot be attributed to lack of training.

5.4.4 What are the limits of working memory?

Conventional wisdom has it that working memory is limited to seven units of information, plus or minus two units: "Everybody knows that there is a finite span of immediate memory and that for a lot of different kinds of test materials this span is about seven items in length" (Miller, 1956). However, exactly what constitutes an "item" is debatable. If a phone number is considered to consist of seven digits or "items," it would fill up working memory. Typically, however, people group the first three digits together and the last four digits together, reducing the memory load. This leaves some spare capacity. Also, not all "items" are created equal in terms of the memory demand that they impose. The complexity and similarity of items, as well as how the information is organized can affect how easy it is to remember.

To account for the fact that experts seem to be able to hold more information in working memory than novices can, Miller developed the concept of "chunking." The idea is that, for example, where novices to ATC may see 12 data blocks on the PVD and find it difficult to make sense of the traffic, the experienced controller sees three groups of four aircraft sequenced in trail to various destinations. By structuring or chunking their knowledge, experts increase the apparent capacities of their working memories.

Since Miller's classic paper, the "magical number seven" has been revised downward for several reasons, which are important for their ATC design implications. It turns out that, when rehearsal is prevented by the demands of another task, the capacity of working memory is more like three chunks than seven (Broadbent, 1975; Newell and Simon, 1972; Simon, 1976). This means that, when rehearsal is possible, some of the items are learned and go into long-term memory, from which they can be retrieved. When rehearsal is prevented, this long-term backup is unavailable.

Time pressure or other stressors can reduce the capacity of working memory to three or four chunks, even for the expert. This is known as cognitive narrowing or cognitive tunnel vision. It accounts for the fact that, under pressure, it is difficult to think of alternatives or to generate new ideas. **Because controllers typically deal with more than one task demand at a time, it is likely that an estimate of three chunks is more realistic than five to nine for the capacity of dynamic working memory** (Moray, 1986). Given that controllers routinely work under some level of time pressure, the lower estimate is more reasonable than the higher estimate. Designs based on the higher estimate will tend to overload the controller's working memory. The capacity of controllers' working memories, thus, varies depending on (at least) experience, task demands, and time pressure.

Given the limitations of short-term memory, what can be done to guard against errors attributable to forgetting? First, the equipment and procedures that controllers use must be designed to be compatible with these limitations. Second, there are ways to help avoid some of the predictable pitfalls of memory. Perhaps the most effective way is to remove the burden of having to remember (e.g., by writing the information down and posting it, or by making the information easily accessible). Another option is to ensure an active involvement in processing the information. For example, a controller is more likely to remember an action that he/she performed than one that he/she watched being performed. Ways of increasing your involvement with information include repeating it to yourself (rehearsal), writing it down (the very process of writing it down will make it easier to remember), or picturing the implications or consequences of an action. Stein and Bailey (1994) have compiled tips for controllers (written by controllers) on how to avoid making common errors as the result of forgetting. Specific tips that controllers offered included: good planning, writing notes on strips, and requesting that pilots report when leaving specific altitudes.

5.4.5 What is the effect of context on pattern recognition?

Context refers to whatever surrounds the item or event of interest. We have already seen how context can affect the way information is interpreted. In Figure 5-1 (page 179), whether the item is seen as a "B" or the number "13" depends entirely on the context. When presented alone, the item could be interpreted either way: there is no right or wrong answer. Similarly, context guides our expectations and understanding of more complex patterns, such as the meanings of words. Consider how the meaning of the word "fly" changes in the following contexts: "Fly Brand X Airlines" and "a fly in the ointment".

Context not only affects the way information is interpreted, it can even influence what information is attended to, what information is ignored (either consciously or subconsciously), and how accurately the information is perceived. An appropriate context can aid information processing. An inappropriate one, i.e., one that is similar, but different in subtle, but important ways, can lead to errors. Just as subtle differences in context can lead to negative transfer in performance, so too, can an inappropriate context lead to errors in information processing.

Many studies show that an appropriate context aids our ability to identify visual stimuli. For example, lines are easier to identify when they are presented in the context of an object, such as a box, than when they are presented alone (e.g., Weisstein and Harris, 1974). Letters are easier to identify when they are presented in a word than when they are presented alone (Reicher, 1969).

Other studies demonstrate how context can trick us into perceiving something different from what is actually presented. In one such study, observers were shown pictures of a loaf of bread, a mailbox, and a drum (Palmer, 1975). The bread and the mailbox were physically very similar. The subject's task was to decide which of the three pictures they saw. The subjects saw the pictures for such a short period of time that they could not be sure which picture they saw. Sometimes, before seeing one of these pictures, subjects were presented with a scene such as a kitchen scene (i.e., a picture of a kitchen counter with utensils and food). When subjects saw a scene that was appropriate for the target picture (such as seeing the kitchen scene before seeing the loaf of bread), accuracy was significantly better than it was when they saw nothing before seeing the target. Performance suffered when subjects were "led down the garden path" with an inappropriate context and a target object that was physically similar to an appropriate object. For example, after seeing the kitchen scene, many subjects were sure they had seen the loaf of bread even if, in fact, they had been shown the mailbox.

Even the simplest forms of pattern recognition show the detrimental effect of a discrepancy between the expected and actual information. To investigate this effect, researchers played a pure tone between 600 and 1500 Hz that was just barely audible and told the listeners that this tone would be played again during one of two time intervals (Scharf, Quigley, Aoki, Peachey, and Neeves, 1987). No tone was played in the other interval. The listeners' task was to decide in which interval the tone was played. When the tone that they had to listen for (the target) was the same frequency as the one they heard first, listeners were 90 percent correct in identifying which interval contained the tone. When the frequency of the target was changed, performance suffered. For example, when a 600 Hz tone was expected and a 600 Hz tone was the target, performance was near perfect with 90 percent accuracy. When a 1000 Hz tone was expected and a

600 Hz tone was the target, performance was near chance with participants guessing which interval contained the tone with only 55 percent accuracy. The same was true when the target tone was 1500 Hz and the prime was 1000 Hz. Even a difference of only 75 Hz (with targets of 925 and 1075 Hz) resulted in a drop in accuracy from 90 percent to 64 percent. The closer auditory warnings are to what is expected (e.g., from training or other previous experience), the easier it will be to "hear," all other things being equal.

The powers of expectancy are even more obvious in higher level processing, such as speech perception. If you quickly read aloud, "the man went to a restaurant for dinner and ordered **"state and potatoes,"** chances are any listeners would hear "the man went to a restaurant for dinner and ordered **"steak and potatoes."** It is not surprising that there have been many Aviation Safety Reporting System (ASRS) reports of pilots accepting clearances not intended for them after requesting higher or lower altitudes. Again, we are most likely to make such mistakes when what we expect to hear is only slightly different from what should be heard (as with similar call signs).

In most cases, context helps or hurts us by setting the stage for expectations. When what we see or hear is compatible with what we expect, we process the information quickly and accurately. When it is incompatible, performance suffers. Our pattern recognition system is set into motion every time our senses perceive something. It is the first step toward processing complex information and problem solving. It is important to understand that pattern recognition cannot be considered in isolation. **When we want to know how much time is required to see or hear a particular stimulus (whether a simple line or tone or a complex message), we must consider the physical attributes of the stimulus, the context in which it will be presented, and the knowledge or expectations of the perceiver.**

5.4.6 What is the role of expectation in information processing?

Expectations are patterns built up over time in long-term memory. They represent our confidence that history will repeat itself. For example, most people are 100 percent confident that the sun will rise in the East and set in the West. This is what we expect to happen.

If United flight 272 enters a controller's airspace at 11:05 every morning, she expects to see that aircraft show up at that time. If an airshow is held at a particular airport every September, controllers know to expect a rush into and out of this airport at specific times that month. Based on patterns of this kind, controllers build up expectations for traffic activity in their sectors by time of day, time of week, time of year, and so forth. These become the routine, expected events that provide the support for recognition-based information processing. Anticipatory clearances, for example, are given on the basis of expectations that are held with a high degree of certainty.

Having expectations for sector activity, the controller can select a plan for managing the sector that has been used successfully time and time again. As the expected aircraft arrive in the sector, they are managed as planned. It is the "unexpected" aircraft or the "unexpected" event that places a burden on information processing, requiring re-planning and attention to pilots' compliance with instructions. If weather becomes a factor, for example, the re-planning and attentional demands may overload the controller's information-processing capabilities, unless an appropriate back-up plan has been pre-defined. To help controllers avoid the pitfalls of their own expectations, design should call attention to exceptions; provide re-planning aids; and assist the controller by monitoring aircraft compliance with instructions.

Exception reporting can be implemented in various ways. The controller can be alerted when an exception occurs, or the exception can be emphasized by coding techniques. Choice of the technique to use depends on several factors, such as the criticality of the exception. Levels of criticality may be indicated by gradations of the technique selected.

5.4.7 What are the mechanisms of remembering and forgetting? How can design compensate for forgetting? How can design support remembering of critical information and remembering of intended actions?

The positive side of forgetting is that it serves to provide storage space for new information. Some forgetting occurs simply as a function of time and the relative importance of the information. Information that has received a significant amount of processing is remembered better than items that receive only superficial processing (Craik and Lockhart, 1972).

For example, in any array of colored numbers, more time will be needed to count the blue numbers than to decide whether or not blue numbers are present in the display. Still more time will be needed to add the blue numbers.

This type of difference in the level of processing is referred to as "depth" of processing. The more, or "deeper," the information is processed, the easier it will be to remember. For example, if a controller has communicated with a particular aircraft several times, s/he will be more likely to remember "seeing" that aircraft and less likely to remember another aircraft that traversed the sector with little communication.

In our previous example of an array of colored numbers, the person who added the blue numbers would have more success in recalling them than the person who counted the same numbers. Information that is not specifically attended to is not likely to be remembered. The more attentional resources spent on processing the information, the more accurately the information will be remembered. This has implications for complex tasks in which it is important to remember certain pieces of information. We can maximize the chances of being able to remember information by requiring that the information be used or processed in some way. Information that is not actively attended to will not be easily recalled from memory when needed. Physically handling and marking flight strips, for example, are forms of deeper processing that may need to be compensated for in the era of electronic flight data.

The key element in forgetting is interference from intervening events and activities. It is important, for example, in the case of any accident to take reports from the participants as early as possible to capture the full context of what took place. As other events intervene, details that might be important are likely to be lost, remembered incorrectly, or filled in with logical inferences on the basis of what the person thinks must have happened or should have happened.

After some passage of time, which is filled with other events, memory goes into a kind of default mode, with strong schemas supplying details that may or may not correspond with what actually happened. This is one sense in which our memories are "reconstructive." We reconstruct our accounts of events in line with prototypical models of those events (i.e., based on what usually happens). Small, but crucial details of an actual event are likely to be forgotten because they are overwhelmed by the strength of the mental model.

We know from research on eyewitness testimony that eyewitnesses to the same event do not necessarily agree with each other on much of anything (Loftus et al, 1979). This is because of the nature of human memory and forgetting, not because of any dark motive on the eyewitnesses' part. Contrary to a common assumption, the strain of involvement in an accident actually impairs, rather than enhances, eyewitness memory.

Design products should be assessed for their ability to aid memory and to overcome the natural human tendency to fall back into default mode based on previous experience.

Recognition tasks are much easier than recall tasks. In a recognition task, the person learns a set of items, known as the memory set. An example of a memory set could be the call signs of all aircraft currently in your sector. The person is then presented with a set of items and asked to pick out those that occurred in the original list (e.g., "Which of the following aircraft are in your sector?"). In a recall task, the person learns the memory set and is then asked to reproduce the items from memory without any cues (e.g., "List all the aircraft in the sector."). The recognition task is accomplished more quickly and accurately than the recall task. **Design**

needs to support recognition and minimize the amount of information that the controller must recall from unaided memory.

5.5 PROBLEM SOLVING AND DECISION MAKING

The controller's perception, attention, and memory all play key roles in ATC problem solving and decision making. In all of these areas, system design should capitalize on human capabilities and compensate for human limitations. By problem solving, we refer to the process of developing a plan to deal with any actual or potential situation that threatens to impact goal accomplishment. Through problem solving, for example, the controller may develop two or three alternative plans for managing a potential conflict situation. By decision making, we refer to selecting from among two or more alternatives, based on evidence collected by the decision maker. Decision making takes place amid uncertainty and can be influenced by individual preferences. Through decision making, the controller chooses one plan to go with and classifies the others as backups.

Alternatively, decision making may occur through a pattern-matching process, whereby the configuration of the situation triggers an appropriate plan, based on the expert decision maker's past experience. Known as Recognition-Primed Decision Making (RPD), this form of skilled decision making is quick and relatively effortless (Klein, 1989). The decision maker recognizes the situation as identical to or similar to one or more situations encountered in the past, and this recognition is sufficient to call to mind a solution that worked in the past. Although RPD has not as yet been scientifically demonstrated in ATC environments, it seems likely to be characteristic of expert ATC decision making and, therefore, a candidate for design support.

To apply a distinction discussed earlier, decision making in unfamiliar situations may be characterized as "knowledge based," that is, requiring effort and creativity. In contrast, decision making in highly familiar situations may be characterized as "rule based," that is, requiring little effort beyond selection of the appropriate rule (solution) to apply. Highly experienced decision makers may exhibit something close to "skill-based" decision making in which decisions are made almost instinctively, with little or no conscious attention.

Both problem solving and effortful decision making are constrained by the limits of working memory. Time pressure is another important constraint.

5.5.1 What are the characteristics of problem solving and decision making? How should design be guided by knowledge of these characteristics?

The following are key characteristics of problem solving in operational settings:

- Requires situation assessment, which may be incomplete or inaccurate.

- Is constrained by level of effort, time required versus time available, limits of attention and working memory, and extent to which recovery from error is possible.

- Is guided by representation or description of problem, by context, and by expectations.

- Requires exercise of many abilities.

The following are key characteristics of decision making in operational environments:

- Time required increases with uncertainty about the choice or decision to be made (Card, Moran, and Newell, 1983).

- Is influenced by general world knowledge (Ashcraft, 1989) and the representational forms of knowledge stored in memory (e.g., mental images of physical objects; non-quantitative, abstract representations of concepts).

- In unfamiliar situations, involves the following steps of classical decision making (Fishhoff, 1986):

  - Identify all possible courses of action
  - Evaluate the consequences of each course of action

- Assess the likelihood that each consequence will actually occur
- Integrate all considerations, using a rational decision rule to select the optimal action.

* In highly familiar situations, draws upon expert recognition of the situation and past successes or failures.

Problem solving and decision making have several constraints in common.

Constraints. Problem solving and decision making can be more or less effortful depending on how familiar the problem or how well known the alternative choices. Someone unskilled in solving a particular problem will take more time and go down more blind alleys than an expert will. This happens because the expert has developed conceptual linkages between problem "patterns" and solutions that worked in the past (Klein, 1989).

Problem solving and decision making are constrained by the time available, which can be critically short in ATC situations. Any uncertainties in the situation will reduce the time available because time will be spent gathering information to increase certainty. The time required for information-processing activities further reduces the time available.

The limits of working memory further constrain problem solving and decision making. With a capacity of three to five chunks while working under time pressure, the controller can consider only the most relevant information and the most feasible alternatives. To conserve time and limit the demand on working memory, experts typically use shortcuts, known as heuristics, which are rules-of-thumb that have worked successfully in the past.

Compared to formal, specified rules (algorithms), **heuristics** are educated guesses or short-cuts that will work sometimes, under some circumstances. Although they work well when they do work, heuristics come with no guarantee of finding an appropriate problem resolution or of making a correct decision (Ashcraft, 1989). Heuristics can get us into trouble if we overuse them, or try to use them in place of necessary, but more complex strategies (Nisbett and Ross, 1980). A built-in problem for problem solvers is that we naturally try to limit the burden on our working memories. As a result, we are generally disinclined to develop or apply complex strategies that threaten to overload our limited WM capacity. Research findings indicate that people do not naturally use formal logic or apply probabilistic reasoning to everyday situations (Kahneman, Slovic, and Tversky, 1982). People are non-optimal decision makers, in the sense that they do not consider all alternatives and their consequences, but instead tend to rely on what has worked in the past. Since considering all alternatives is often not necessary, however, problem resolutions or decisions may be reached faster through the pattern-matching process discussed earlier.

People are satisfied with "good enough" solutions, which are often not as good as solutions achieved through formal analysis. In general, people will do whatever it takes to avoid getting mired down in analytic decision making. For example, professional researchers are limited in their ability to apply statistical knowledge correctly. Knowing about probability and statistics does not guarantee their proper use in a time-pressured decision-making situation. When people are asked to make probability estimates or judgments, they typically fall back on rules-of-thumb or heuristics. Over 25 such heuristics have been identified (Sage, 1981).

Decision making, therefore, "reduces to a gamble surrounded by uncertainty regarding what one will get and how one will like it" (Fischhoff, 1986, p. 61). People tend to be overly influenced by salient information, that is, information that "pops out" at them because it is brightly coded, centrally located, or emotionally charged. People also rely on the first inferences or possibilities that come to mind (Wickens, 1992). Highly publicized events tend to be more available even though their actual rate of occurrence may be fairly low. For example, people tend to overestimate the dangers of flying on domestic air carriers (particularly in comparison to automobile travel) because they remember the vivid reports of plane crashes. In reality, traveling by car is far more likely to end in accidental injury or worse.

We also tend to seek information that supports our beliefs and not to look for information that would disconfirm our beliefs about situations. Worse yet, we tend to be overconfident that we have then made the right decision (Wickens, 1992).

Problem solvers can afford to be more divergent or creative in their thinking if there is a high likelihood of recovering from errors in problem solving. A low likelihood of recovery severely constrains the feasibility of creative approaches. Estimating the likelihood of recovery is a matter of judgment, which is subject to error.

To compensate for constraints on problem solving and decision making, system design needs to provide support for these activities across the spectrum from novice to expert.

Problem Representation. Research shows that the way a problem is represented also affects the ease of solving it (e.g., de Keyser, 1986; Woods, 1991). Even an expert will have trouble with a familiar problem if it is presented differently than it has been before. A different representation can disrupt pattern matching.

In general, problems can be represented in terms of four elements (Refs TBS):

- A description of the initial situation when problem solving begins.

- A description of the goal state to be reached.

- A set of actions that can be taken, which serve to alter the current situation.

- Constraints that impose additional conditions on finding a successful path to a solution, beyond simply reaching the goal.

Applying these elements to a basic ATC problem, we might describe a general initial situation as the detection of a potential loss of separation; the goal state as separation; ATC procedures and possible control actions as the set of actions; and any flow or airspace restrictions in effect as constraints. Weather would be an additional path constraint.

A well-designed system should provide support for each of these elements of problem solving in ATC: support for detection of potential conflicts, support for achieving separation, support for selecting appropriate procedures and control actions, and support working within constraints (e.g., traffic management restrictions). To aid the controller in breaking out of an unproductive mind set induced by time pressure, design should also make available alternative views of the same problem.

Required Abilities. Problem solving calls upon many abilities, including perception, memory, classification, language, judgment, and timing of actions. Through the experience gained in trying different solution paths, intelligent problem solving involves the learning of methods and tactics that may or may not resolve similar situations in the future (Holyoak, 1990). Learning and remembering what not to do can shorten the time it takes to resolve a future problem.

Expertise develops through learning. Outcomes of actions taken or not taken guide learning. Unwanted outcomes signal a need for future revision of some part of the problem-solving process. In the ATC environment, where getting it right the first time may be critical, design and evaluation of automated systems should ensure that opportunities are "built in" for the controller to be aware of possible consequences of action or inaction, especially for rare events.

As an aid to pattern matching, an intelligent system is capable of storing situation descriptors along with operational inputs made to resolve the situation.

Search Process. Problem solving is typically described as a search for an appropriate route to a goal. It is something like finding your way through a maze. As you make your way through the problem space for the first time, you note the dead ends, make many wrong turns, but try to keep the goal firmly in mind, even as you become fairly disoriented. As you advance through the maze, you build a cognitive map of the solution, piecing together the productive paths. Your first trip through the maze is effortful, slow, and fraught with errors, but later trips become easier and faster as you master a path to the goal. The path you master may not be the shortest or fastest, but it is the path you come to know.

If we think of the maze as the highway system in a strange city, we can identify a set of aids provided to the problem solver: maps, highway signs with route numbers and street names, exit numbers, street signs in the city, and so forth. Imagine trying to find a particular location in a strange city without any of these aids. Or imagine the problems created by faulty aids, such as maps that no longer correspond to street names or a highway in Massachusetts once marked as both 95N and 128S. Importantly, automated aids must be integrated at a system level so that changes made in one aid are compatible with the other aids and with the information they presented.

In ATC, problem solving is bounded by procedures and constraints, such as flow restrictions. Procedures define acceptable paths to the goal, but part of the problem is to identify the applicable procedure.

5.5.2 What factors can affect information processing and decision making in the ATC environment?

We have considered the effects of many factors that can affect ATC information processing and decision making. These include the limitations of attention and memory, strategies to reduce mental workload, time pressure and stress, as well as contextual cues and expectations. Any of these factors can result in flawed situational awareness and the formation of an incorrect inference, prediction, or intention, which can guide the controller down a non-productive path.

Individual differences in personal characteristics play a large part in defining a person's "cognitive style," that is, his or her preferred approach to information processing or problem solving. Some people, for example, are more comfortable dealing with textual information, while others prefer visual images as the basis for mental manipulations or transformations of data.

Offering alternative ways of coding data and different ways of representing ATC situations may be approaches to improving productivity in ATC environments.

Fatigue is an additional factor that can have serious consequences for the quality of information processing. Fatigue resulting from shiftwork is a chronic problem in ATC environments. Shiftwork disrupts the controller's sleep-wake cycle and other bodily cycles or 24-hour circadian rhythms (Melton and Bartonowitz, 1986). Effects of shiftwork include not only fatigue, but also decreased alertness, diminished cognitive functioning, and slower response times (Monk and Folkard, 1983). Research currently in progress at CAMI is investigating the fatigue and performance effects of shiftwork on controllers in varying age groups (Della Rocco, 1991).

Age is, of course, another factor that affects information processing. Typically, after the mid-forties, memory capacities diminish, and the speed of information processing is slower than it is in younger people. However, while it may take an older person longer to solve a problem than it takes a younger person, it is important not to underestimate the benefits of experience in complex problem solving.

5.6 SUMMARY OF IMPLICATIONS FOR DESIGN AND EVALUATION

From a system design perspective, the major implication of what we know about human information processing is that systems should be designed for compatibility with the way people process information. That is, systems should be engineered for smooth interaction between their human and machine constituents.

One key point to take away from this chapter is that information processing takes time, and design should support the most efficient information processing possible. Information processing is also subject to error, and design should reduce the likelihood of error. Rather than forcing the controller to adjust to the computer's techniques for processing information, however, the controller should experience the computer-human interface as natural and easy to use.

Assessing design products for cognitive compatibility is a key evaluation objective. **Compatible design products will respect both the capabilities and limitations of the human information-processing system.** They will complement capabilities and help to compensate for limitations, while keeping the controller informed about the status and behavior of the automated system.

5.6.1 What is the relationship between human information processing and system design?

System design, particularly design of the computer-human interface (CHI), should be grounded in knowledge of human capabilities and limitations in information processing. CHI design includes a broad range of topics, including the design of displays, controls, and decision support. In each of these areas, successful design depends largely upon the depth of the designer's understanding of human perception and information processing. Chapter 7 of this handbook is devoted to a consideration of CHI issues. Here, we review some key design challenges.

Design of Displays and Controls. The primary purpose of perceptual displays in ATC is to provide the controller with the information needed to build and maintain an awareness of the ongoing and upcoming traffic situation. Those displays will be most effective in accomplishing this purpose if their design is based an understanding of ATC information requirements and the capabilities and limitations of human information processing. The basic principles of display design are derived from such an understanding.

Design of visual and auditory displays is discussed in Chapter 7, Computer-Human Interface (CHI) Considerations. The issues that arise when visual and auditory displays occur in combination are also discussed in Chapter 7. Design of interactive controls is discussed in Chapter 9, Workstation and Facility Design and Evaluation.

Design of Decision Support. A major challenge in providing decision support is to know how much is enough. If too little decision support is provided, the controller may be overwhelmed by the available raw data and the various forms in which information is presented. If too much decision support is provided, the controller may be taken too far out of the loop. In either case, situation awareness will be impacted.

Another challenge is to design decision aids so that they do not impose high demands on information-processing resources. If the controller needs to spend extra time and effort to consult the decision aid, or, if the aid must be monitored continuously, the aid may be seen as imposing an unacceptable burden on the controller. **Automated tools and capabilities designed to support decision making need to provide the controller a clear benefit in terms of processing time saved and improved quality of decisions made.**

A related challenge is to integrate ATC decision support systems so that the controller consults one, not several automated advisors. A proliferation of individual systems that provide the controller with specialized information is not the prescription for reducing controller workload or increasing controller productivity. Having to understand each source of information and having to integrate information from multiple sources are highly demanding cognitive tasks. **Costs in time and effort can be expected to outweigh the benefits unless items of information from multiple sources**

are integrated as appropriate and presented to the controller in a readily usable form.

5.6.2 Given what we know about the demands of information processing in ATC, how might automated aiding be useful?

If automated aiding is designed from a controller-centered perspective, it can be useful in supporting key information-processing tasks, such as building and maintaining an awareness of the ATC situation, remembering intended actions, and resolving conflicts or potential conflicts. The objective in developing an automated ATC advisor or "Controller's Associate" must be to coordinate and integrate controller expertise with the automated expertise in such a way that both, working together, produce better combined performance than is possible for either working alone (Woods and Roth, 1988). The controller must be in control of the interaction with the automated advisor, which must help the controller define the problem, identify and evaluate possible solutions, and address the right questions.

Development of aiding for improved situational awareness can begin with support for perception, pattern matching, and classification of situations. Perceptual aids might include improved tools for visualization, such as pictorial displays and perspective displays. **Perception of changes or unusual events in the situation should be supported by clear, salient indicators, not by small changes that are easy to miss.** Perceptual aids need to provide the controller with an adequate, dynamic representation of both traffic status and system status, on which to base decisions.

Improved situational awareness can also be supported by tools that aid integration of data and comprehension of information. These include integrated displays that put data in context and that support an immediate, holistic grasp of the right information at the right time (Mitchell and Saisi, 1987). Other aids might help the controller consider peripheral cues to distinguish between situations and overcome too narrow a focus. Support for building a strategic "big" picture and a tactical "little" picture might be useful aids to situational awareness.

Further aids to situational awareness are tools that help the controller project the current situation into the future. The "velocity vector" in today's system is an example of such an aid. Predictive displays would allow the controller to simulate the effects of possible control actions on a developing situation. Such predictions could help generate evidence to confirm or disconfirm the controller's own expectations. In this way, the controller's precision in working out "what if" questions could be greatly enhanced. Predictive displays, however, must be designed so that the controller does not lose touch with the current system or delay making necessary decisions (Findler, Bickmore, and Cromp, 1985).

Tools to support decision making and conflict resolution must do more than offer solutions. Use of the aid must provide an active role for the controller so that s/he will be able to respond flexibly to unanticipated situations that are beyond the capabilities of the machine expert. The controller needs to know the limits of the advisor and have access to the data or information used by the advisor in developing a solution. Without this knowledge and knowledge of the advisor's reasoning process, the controller will be hampered in evaluating the quality of the advice offered.

Automated aiding can reduce the loading on the controller's working memory and help in maintaining an optimal level of workload. It can help to overcome the limitations of habitual information-processing shortcuts by identifying and presenting all pertinent information. **If not designed properly, however, automated aiding can reduce situational awareness.**

If automated aiding is designed to completely take over some aspect of the controller's information-processing work, it risks placing the controller in a passive role, with the further risk of total system breakdown when unanticipated situations arise. What we can be sure of is that unusual situations will occur; pre-planned solutions will be inadequate; adaptation to special conditions will be required; and recovery from human or computer errors will be necessary (Woods and Roth, 1988). This means that automated aids and advisors must be evaluated carefully for their effect on the controller's situational awareness and active role as a flexible decision maker.

5.6.3 When tradeoffs must be made, what are minimal human factors requirements for information processing?

The information presented to the controller must be the right information, at the right time, in the right format, and in the right sequence. Information must be selected, integrated, and organized for compatibility with task requirements. It must be represented in a form that lends itself to rapid extraction from the display and appropriate integration within the controller's picture of the ATC situation. Design must allow sufficient time for the controller to perceive, integrate, project, and act upon ATC information. The amount of time required will vary with operational conditions and controller experience. All these requirements need to be evaluated in actual tests of the system in order to be sure they are satisfied. Early field evaluations are recommended (Harwood and Sanford, 1994).

Information must not overwhelm the limits of the controller's working memory, especially the reduced capacity of a novice under time pressure. How much information will result in overload? Since this depends on the individual (e.g., experience level, skill, fatigue) and circumstances (e.g., traffic volume and complexity), a conservative approach should be taken in the testing of this issue with a new system or any additions to the current system.

5.7 REFERENCES AND SUGGESTIONS FOR FURTHER READING

Ammerman, H. L., Bergen, L. J., Davies, D. K., Hostetler, C. M., Inman, E. E., and Jones, G. W. (1988). *FAA air traffic control operations concepts* (Vol. VI: ARTCC/HOST En Route Controllers, Change 1, DOT/FAA/AP-87-01). Washington, DC: Department of Transportation/Federal Aviation Administration, Advanced Automation Program.

Ashcraft, M. H. (1989). *Human memory and cognition.* Boston, MA: Scott, Foresman.

Baddeley, A. (1990). *Human memory: Theory and practice.* Boston: Allyn and Bacon.

Bainbridge, L. (1975). *Working memory in air traffic control* (Unpublished manuscript). London: University College, Department of Psychology.

Bentley, R., Hughes, J. A., Randall, D., Rodden, T., Sawyer, P., Shapiro, D., and Sommerville, I. (1992). Ethnographically-informed system design for air traffic control. *CSCW '92: Sharing perspectives* (*Proceedings of the Conference on Computer-Supported Cooperative Work*, pp. 123-129). New York: ACM Press.

Berson, B. L., Po-Chedley, D. A., Boucek, G., Hanson, D. C., Leffler, M. F., and Wasson, R.L. (1981). Aircraft alerting systems standardization study. (Volume II: *Aircraft alerting system design guidelines*, Report No. DOT/FAA/RD-81-II). Washington, DC: Department of Transportation/Federal Aviation Administration.

Bisseret, A. (1971). Analysis of mental processes involved in air traffic control. *Ergonomics, 14(5)*, 565-570.

Block, N. (1990). The computer model of the mind. In D. N. Osherson and E. E. Smith (Eds.), *Thinking: An invitation to cognitive science* (Vol. 3, pp. 247-289). Cambridge, MA: MIT Press.

Boff, K. R., and Lincoln, J. E. (Eds.). (1988). *Engineering data compendium: Human perception and performance.* Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Boucek, G. P., Erickson, J. B., Berson, B. L., Hanson, D. C., Leffler, M. F., and Po-Chedley, D. A. (1980). *Aircraft alerting systems standardization study* (Phase I final report, Report No. FAA/RD-80-68). Washington, DC: Department of Transportation/Federal Aviation Administration.

Boucek, G. P., Pfaff, T. A., White, R. W., and Smith, W. D. (1985). *Traffic Alert and Collision Avoidance System—Operational simulation* (DOT/FAA/PM-85/10). Washington, DC: Department of Transportation/Federal Aviation Administration, Program Engineering and Maintenance Service.

Boucek, G. P., Po-Chedley, D. A., Berson, B. L., Hanson, D. C., Leffler, M. F., and White, R. W. (1981). Aircraft alerting systems standardization study. (Volume 1: *Candidate system validation and time-critical display evaluation*, Report No. FAA-RD-81-381). Washington, DC: Department of Transportation/Federal Aviation Administration.

Boucek, G. P., White, R. W., Smith, W. D., and Kraus, J. M. (1982). *Traffic Alert and Collision Avoidance System - Developmental simulation* (Report No. DOT/FAA/RD-82/49). Washington, DC: Department of Transportation/Federal Aviation Administration.

Broadbent, D. E. (1975). The magical number seven after fifteen years. In A. Kennedy and A. Wilkes (Eds.), *Studies in long-term memory* (pp. 3-18). New York: Wiley.

Campbell, F. W., and Westheimer, G. (1960). Dynamics of accommodation responses of the human eye. *Journal of Physiology, 151*, 285-295.

Card, S. K., Moran, T. P., and Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.

Cardosi, K. (1993). Time required for transmission of time-critical air traffic control messages in an en route environment. *The International Journal of Aviation Psychology* 3 (4), 303-313.

Cardosi, K., and Boole, P. (1991). *Analysis of pilot response time to time-critical air traffic control calls* (Report No. DOT/FAA/RD-91/20). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Cardosi, K., Burki-Cohen, J., Boole, P., Mengert, P., and Disario, R. (1992). *Controller response to conflict resolution advisory* (Report No. DOT/FAA/NA-92/1). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Cardosi, K., Warner, M., Boole, P., Mengert, P., and DiSario, R. (1992). *Controller response to conflict resolution advisory prototype* (Report No. DOT/FAA/NA-92/1). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Carroll, J. M., and Olson, J. R. (1988). Mental models in human-computer interaction. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 45-65). New York: Elsevier/North Holland.

Carpenter, P. A., and Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr and K. Kotovsky (Eds.), *Complex information processing* (pp. 31-68). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., and Lockhart, R. S. (1972). Levels of processing: A Framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

De Keyser, V. (1986). Technical assistance to the operator in case of incident: Some lines of thought. In E. Hollnagel, G. Mancini, and D. D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 229-253). Berlin: Springer-Verlag.

Della Rocco, P. (1991). *Fatigue and performance: Shiftwork in controllers of varying age* (Dissertation Prospectus).

Diges, M. (1988). Stereotypes and memory of real traffic accidents. In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, Memory in everyday life, pp. 59-65). New York: Wiley.

Endsley, M. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors Society.

Endsley, M. (1990). *Situation awareness in dynamic human decision making: Theory and measurement* (NOR DOC 90-49). Los Angeles: Northrop Corporation.

Ericsson, K. A., and Staszewski, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr and K. Kotovsky (Eds.), *Complex information processing* (pp. 235-267). Hillsdale, NJ: Erlbaum.

Falzon, P. (Ed.). (1990). *Cognitive ergonomics: Understanding, learning and designing human-computer interaction.* New York: Harcourt Brace Jovanovich.

Findler, N. V., Bickmore, T. W., and Cromp, R. E. (1985). A General-purpose man-machine environment with specific reference to air traffic control. *International Journal of Man-Machine Studies, 23*, 587-603.

Finkelman, J. M., and Kirschner, C. (1980). An information-processing interpretation of air traffic control stress. *Human Factors, 22*, 547-567.

Fischhoff, B. (1986). Decision making in complex systems. In E. Hollnagel, G. Mancini, and D. D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 61-85). Berlin: Springer-Verlag.

Frost, G. (1972). Man-machine system dynamics. In H. P. VanCott and G. R. Kinkade (Eds.), *Human engineering guide to systems design* (pp. 227-309). Washington, DC: U. S. Government Printing Office.

Gardiner, M. M., and Christie, B. (1987). *Applying cognitive psychology to user-interface design*. New York: Wiley.

Gilmore, W. E., Gertman, D. I., and Blackman, H. S. (1989). *The User-computer interface in process control: A Human factors engineering handbook.* New York: Academic Press.

Glaser, R. (1989). Expertise and learning: How do we think about instructional processes now that we have discovered knowledge structures? In D. Klahr and K. Kotovsky (Eds.), *Complex information processing* (pp. 269-282). Hillsdale, NJ: Erlbaum.

Goodstein, L. P. (1990). Computer aids for decision-making in process control. In P. Falzon (Ed.), *Cognitive ergonomics: Understanding, learning and designing human-computer interaction* (pp. 231-244). New York: Harcourt Brace Jovanovich.

Greene, R. L. (1992). *Human memory: Paradigms and paradoxes.* Hillsdale, NJ: Erlbaum.

Harwood, K., Mogford, R., Murphy, E., and Roske-Hofstrand, R. (1991). Conceptual structures of air traffic controllers: Implications for training and display aiding. *36th Annual Air Traffic Control Association Conference Proceedings* (pp. 394-400). Arlington, VA: Air Traffic Control Association.

Harwood, K., and Sanford, B. (1994). Evaluation in context. In J. A. Wise, V. D. Hopkin, and D. J. Garland (Eds.), *Human factors certification of advanced aviation technologies* (Aviation Human Factors Series, pp. 245-262). New York: Springer-Verlag.

Harwood, K., Wickens, C., Kramer, A., Clay, D., and Liu, Y. (1986). Effects of display proximity and memory demands on the understanding of dynamic multidimensional information. *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 786-790). Santa Monica, CA: Human Factors Society.

Held, J. S. (1975). *Measurement of relative productivity of ATC automation enhancements* (MTR-7097). McLean, VA: MITRE.

Holyoak, K. J. (1990). Problem solving. In D. N. Osherson and E. E. Smith (Eds.), *Thinking: An introduction to cognitive science* (Vol. 3, pp. 117-146). Cambridge, MA: MIT Press.

Hopkin, V. D. (1982). *Human factors in air traffic control.* (AGARDograph No. 275, pp. 31-46). Neuilly-sur-Seine, France: NATO, AGARD.

Hopkin, V. D. (1980). Boredom. *The Controller, 19*(1), 6-9.

Hughes, J. A., Randall, D., and Shapiro, D. (1992). Faltering from ethnography to design. *CSCW '92: Sharing perspectives (Proceedings of the Conference on Computer-Supported Cooperative Work*, pp. 115-122). New York: ACM Press.

Johnson, W. A., and Heinz, S. P. (1978). Flexibility and capability demands of attention. *Journal of Experimental Psychology: General, 107*, 420-435.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., Slovic, P., Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kantowitz, B. H., and Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships.* New York: Wiley.

Keblawi, F. S. (1976). *Controller productivity in the upgraded third generation air traffic control system* (Part I: Automation in the pre-data link era [FAA-EM-76-3/MTR-7212]). McLean, VA: MITRE.

Klapp, S. T., and Netick, A. (1988). Multiple resources for processing and storage in short-term working memory. *Human Factors, 30*, 617-632.

Klein, G. A. (1993). A Recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, and E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 138-147). Norwood, NJ: Ablex.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in man-machine system research* (Vol. 5, pp. 47-92). Greenwich, CT: JAI Press.

Klein, G. A., Calderwood, R., and Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 576-580). Santa Monica, CA: Human Factors Society.

Klein, G. A., Calderwood, R., and MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man and Cybernetics, 19(3)*, 462-472.

Klein, G. A., Orasanu, J., Calderwood, R., and Zsambok, C. E. (Eds.). (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Koriat, A., and Ben-Zur, H. (1988). Remembering that I did it: Processes and deficits in output monitoring. In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, Memory in everyday life, pp. 203-208). New York: Wiley.

Lachman, R., Lachman, J. L., and Butterfield, E. C. (1979). *Cognitive psychology and information processing*. Hillsdale, NJ: Erlbaum.

Larkin, J. H. (1989). Display-based problem solving. In D. Klahr and K. Kotovsky (Eds.), *Complex information processing* (pp. 319-341). Hillsdale, NJ: Erlbaum.

Loftus, G. R., Dark, V. J., and Williams, D. (1979). Short-term memory factors in ground controller/pilot communication. *Human Factors, 21*, 169-181.

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30*, 583-598.

Mackworth, N. H. (1948). The Breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1*, 5-61.

Martindale, C. (1981). *Cognition and consciousness*. Homewood, IL: Dorsey Press.

McCormick, E. J., and Sanders, M. S. (1982). *Human factors in engineering and design* (5th ed.). New York: McGraw-Hill.

Melton, C. E., and Bartanowicz, R. S. (1986). *Biological rhythms and rotating shift work: Some considerations for air traffic controllers and managers* (DOT/FAA/AM-86/2). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Mogford, R. H. (1990). *The Characteristics and importance of the air traffic controller's mental model.* (Doctoral dissertation). Ottawa, ON, Canada: Carleton University.

Monk, T., and Folkard, S. (1983). Circadian rhythms and shiftwork. In R. Hockey (Ed.), *Stress and fatigue in human performance* (pp. 97-122). New York: Wiley.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. Boff, L. Kaufmann, and J. Thomas (Eds.), *Handbook of perception and human performance.* (Volume II. Cognitive processes and performance, pp. 40-1, 40-51). New York: Wiley.

Moray, N. (1986). Modelling cognitive activities: Human limitations in relation to computer aids. In E. Hollnagel, G. Mancini, and D. D. Woods (Eds.), *Intelligent decision support in process environments* (NATO Advanced Study Institute F21, pp. 273-291). Berlin: Springer-Verlag.

Moray, N. (1978). Strategic control of information processing. In G. Underwood (Ed.), *Strategies of information processing* (pp. 301-328). New York: Academic Press.

Neisser, U., and Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology, 7*, 480-494.

Nelson, D. L. (1988). How what we know interferes with what we remember. In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, Memory in everyday life, pp. 484-489). New York: Wiley.

Newell, A., and Simon, H. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R., and Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Norman, D. (1983). Design rules based on analyses of human error. *Communications of the ACM, 26* (April), 254-258.

Norman, D. A. (1986). New views of information processing: Implications for intelligent decision support systems. In E. Hollnagel, G. Mancini, and D. D. Woods (Eds.), *Intelligent decision support in process environments* (NATO ASI Series F21, pp. 123-136). Berlin: Springer-Verlag.

Norman, K. L. Models of the mind and machine: Information flow and control between humans and computers. In M. C. Yovits (Ed.), *Advances in computers* (Vol. 32, pp. 201-254). New York: Academic Press.

Office of Aviation Medicine. (1986). *ATC operational errors* (Staff study). Washington, DC: Department of Transportation/Federal Aviation Administration, Author.

Osherson, D. N. (1990). Judgment. In D. N. Osherson and E. E. Smith (Eds.), *Thinking: An invitation to cognitive science* (Vol. 3, pp. 55-87). Cambridge, MA: Mit Press.

Palmer, S. E. (1975). The Effects of contextual scenes on the identification of objects. *Memory and Cognition, 3*, 519-526.

Parasuraman, R., Molloy, R., and Singh, I. L. (1993). Performance consequences of automation-induced complacency. *International Journal of Aviation Psychology, 3*, 1-24.

Parsons, H. M. (1985). Automation and the individual: Comprehensive and comparative views. *Human Factors, 27*, 99-112.

Pitz, G. F., and Sachs, N. J. (1984). Judgment and decision: Theory and application. *Annual review of psychology* (Vol. 35, pp. 139-163). Palo Alto, CA: Annual Reviews.

Potter, M. C. (1990). Remembering. In D. N. Osherson and E. E. Smith (Eds.), *Thinking: An invitation to cognitive science* (Vol. 3, pp. 3-32). Cambridge, MA: MIT Press.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York: North Holland.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13*, 257-266.

Reason, J. (1979). Actions not as planned: The price of automatization. In G. Underwood and R. Stevens (Eds.), *Aspects of consciousness* (Vol. 1, pp. 67-89). London: Academic Press.

Reason, J. (1990). *Human error.* New York: Cambridge University Press.

Reason, J. (1986). Recurrent errors in process environments: Some implications for the design of intelligent decision support systems. In E. Hollnagel, G. Mancini, and D. D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 255-270). Berlin: Springer-Verlag.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*, 275-221.

Rodgers, M. D., and Nye, L. G. (1992). *Factors associated with the severity of operational errors at Air Route Traffic Control Centers* (Poster presented at the Human Factors Society Annual Meeting). Oklahoma City, OK: Department of Transportation/Federal Aviation Administration, Civil Aeromedical Institute.

Rucker, R. A., Parlow, R. D., Sterfels, P. R., and Mayer, D. S. (1973). *Controller productivity study* (FAA-EM-73-3). McLean, VA: MITRE.

Sage, A. P. (1981). Behavioral and organizational considerations in the design of information systems and processes for planning and decision support. *IEEE Transactions on Systems, Man and Cybernetics, SMC-11*, 640-678.

Sanderson, P. (1986). Skills, displays and decision support. In E. Hollnagel, G. Mancini, and D. D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 347-352). Berlin: Springer-Verlag.

Sarter, N. B., and Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *International Journal of Aviation Psychology, 1*, 45-57.

Sarter, N. B., and Woods, D. D. (1993). Pilot interaction with cockpit automation: Operational experiences with the Flight Management System. *International Journal of Aviation Psychology, 2*, 303-321.

Scharf, B., Quigley, S., Aoki, C., Peachey, N., and Reeves, A. (1987). Focused attention and frequency selectivity. *Perception & Psychophysics, 42*, 215-221.

Schooler, J. W., Clark, C. A., and Loftus, E. F. (1988). Knowing when memory is real. In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, Memory in everyday life, pp. 83-88). New York: Wiley.

Senders, J., and Moray, N. (1991). *Human error.* Hillsdale, NJ: Erlbaum.

Sheridan, T. B. (1981). Understanding of human error and aiding human diagnostic behavior in nuclear power plants. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failure* (pp. 19-35). New York: Plenum.

Shneiderman, B. (1992). *Designing the user interface* (2nd edition). New York: Addison-Wesley.

Simon, H. (1976). The information-storage system called human memory. In M. R. Rosenzweig and E. L. Bennett (Eds.), *Neural mechanisms of learning and memory* (pp. 79-96). Cambridge, MA: MIT Press.

Spoehr, K. T., and Lehmkuhle, S. W. (1982). *Visual information processing.* San Francisco: W. H. Freeman.

Stein, E. (1992). *Air traffic control visual scanning* (DOT/FAA/CT-TN92/16). Atlantic City International Airport, NJ: Department of Transportation/Federal Aviation Administration Technical Center.

Stein, E., and Bailey, J. (1994). *The controller memory guide: Concepts from the field* (report No. DOT/FAA/CT-TN94/28) Atlantic City International Airport,NJ:Department of Transportation/Federal Aviation Administration Technical Center.

Stewart, L. J. (1983). Conceptual models of information processing. In C. M. Mitchell, P. M. Van Balen, and K. Moe (Eds.), *Human factors considerations in system design* (NASA Conference Publication 2246, pp. 215-238). Greenbelt, MD: NASA/Goddard Space Flight Center.

Thackray, R. I. (1980). *Boredom and monotony as a consequence of automation: A consideration of the evidence relating boredom and monotony to stress* (FAA-AM-80-1). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

Treisman, A. (1964). The effect of irrelevant material on the efficiency of selective listening. *American Journal of Psychology, 77*, 533-546.

Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 523-530.

Tsang, P. S., and Vidulich, M. A. (1989). Cognitive demands of automation in aviation. In R. S. Jensen (Ed.), *Aviation psychology* (pp. 66-95). Brookfield, VT: Gower Technical.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Weisstein, N., and Harrison, C. S. (1974). Visual detection of line segments: An Object-superiority effect. *Science, 186*, 752-755.

Whitfield, D., and Jackson, A. (1982). The Air traffic controller's "picture" as an example of a mental model. In G. Johannsen and J. E. Rijnsdorp (Eds.), *Analysis, design, and evaluation of man-machine systems* (pp. 45-52). New York: Pergamon.

Wickens, C. D. (1993). Decision making. In K. Cardosi and M. S. Huntley (Eds.), *Human factors for flight deck certification personnel* (DOT/FAA/RD-93/5, pp. 133-163). Washington, DC: Department of Transportation/Federal Aviation Administration.

Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles E. Merrill.

Wickens, C. D. (1992a). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

Wickens, C. D. (1987). Ergonomic-driven displays. *Proceedings of the Mid-Central Ergonomics/Human Factors Conference.*

Wickens, C. D. (1992b). Predictive workload models and multiple-task performance. *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 12-16). Santa Monica, CA: Human Factors Society.

Wickens, C. D., and Flach, J. M. (1988). Information processing. In D. Nagel and E. Wiener (Eds.), *Human factors in aviation* (pp. 111-155). New York: Academic Press.

Wickens, C. D., and Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors, 30*, 599-616.

Wiener, E. L. (1985). Beyond the sterile cockpit. *Human Factors, 27*, 75-90.

Woods, D. D. (1991). The Cognitive engineering of problem representations. In G. R. S. Weir and J. L. Alty (Eds.), *Human-computer interaction in complex systems* (pp. 169-188). New York: Academic Press.

Woods, D. D. (1988). Coping with complexity: The psychology of human behavior in complex systems. In L. P. Goodstein, H. B. Anderson, and S. E. Olsen (Eds.), *Tasks, errors, and mental models* (pp. 128-148). London: Taylor & Francis.

Woods, D. D. (1986). Paradigms for intelligent decision support. In E. Hollnagel, G. Mancini, and D. D. Woods (Eds.), *Intelligent decision support in process environments.* Berlin: Springer-Verlag.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Systems, 21*, 229-244.

Woods, D. D., and Hollnagel, E. (1987). Mapping cognitive demands in complex problem-solving worlds. *International Journal of Man-Machine Studies, 26*, 257-275.

Woods, D. D., and Roth, E. M. (1988). Cognitive systems engineering. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 1-41). New York: North-Holland.

Yntema, D. (1963). Keeping track of several things at once. *Human Factors, 6*, 7-17.

APPENDIX 5A. CHECKLIST ITEMS
(Also see Chapter 7)

1. This design presents all information in a usable form; it requires little or no mental transformation of data from one unit of measurement into another (5.1.1).

2. The design allows the controller sufficient time to perceive information and to integrate new information with known information (5.1.1).

3. The design allows the controller sufficient time to project potential outcomes of optional control actions (5.1.1).

4. The computer responds quickly enough that the controller is not kept waiting for information (i.e., in less than one second for simple, frequent tasks; two to four seconds for more complex processing; and eight to twelve seconds for infrequent, highly-complex tasks) (5.1.2).

5. This design helps the controller to integrate information from multiple sources, if the information is not already integrated before it is presented to the controller (5.1.2, 5.4.2, 5.6.1).

6. Information that the controller must read and understand such as alarms and critical error messages never blinks or flashes rapidly (greater than 3 Hz) (5.1.3, 5.1.5).

7. Flashing text blinks at a rate of 2 to 3 Hz (5.1.5, 7.2.9).

8. Information that is blinking has an "on" period that is at least as long as the "off" period (5.1.5).

9. Symbols chosen for the display are intuitive (5.1.5, 7.2.9).

10. For a time-critical warning system (such as a conflict detection or resolution advisory), the controller response time assumed in the algorithm has been measured (5.1.5).

11. This design provides all the information needed for planning purposes (5.2.1).

12. This design supports timesharing of information-processing activities, that is, visual, auditory, and decision-making processes can be performed together without overloading the controller (5.2.1).

13. This design provides appropriate memory joggers (e.g., prompts, cues) (5.2.1, 5.4.4).

14. This design supports complete, accurate awareness of the ATC situation (5.2.1).

15. This design is not likely to overload the controller's working memory (5.2.1, 5.4.4).

16. The information that is selected and presented fully supports the controller in making judgment calls and decisions (5.2.1).

17. Information from sub-systems is integrated and presented in a way that minimizes the need to switch from one display to another (5.2.1, 5.4.2).

18. This design provides the controller with all the necessary information for a specific task when it is needed in the appropriate sequence (5.2.3).

19. The information provided helps the controller to recognize situations that require control action (5.2.3).

20. If predictive displays are provided, they assist the controller in projecting the combined effects of many situational factors (5.2.3).

21. If the design includes predictive displays, they do not place additional memory demands or other information-processing burdens on the controller (5.2.3).

22. Design of the work environment limits sources of distraction (5.2.4).

23. The design assists the controller in detecting and correcting errors in data entry (5.2.4, 6.3.3).

24. This design effectively directs the controller's attention by means of alerting, coding, and emphasis techniques (5.3.1).

25. This design provides an active, involved role for the controller (5.3.2, 5.6.2, 6.2.1, 6.3.3).

26. This design does not require the controller to perform purely monitoring tasks for more than 20 to 30 minutes at a time (5.3.2, 6.2.1).

27. This design alerts the controller to critical situations with enough lead time to formulate and execute appropriate responses (5.3.2).

28. This design allows the controller to keep some information processing resources in reserve for unexpected events (5.3.2, 5.4.4).

29. Visual and auditory coding techniques help the controller maintain productive scanning and problem-detection strategies (5.3.2).

30. This design requires little or no unaided recall of information (5.4.7).

31. This design calls attention to situations that depart from what the controller would normally expect (5.4.5, 5.4.6, 5.6.2).

32. This design provides adequate support for achieving aircraft separation and for detection of potential conflicts (5.5.1).

33. Automated aids are adequately integrated with each other (5.5.1).

34. Perceptual displays help the controller in building and maintaining situational awareness, i.e., in perceiving, integrating, and projecting information about the ATC situation (5.6.1, 5.4.2).

35. Decision aids don't need to be monitored continuously (5.6.1).

36. Decision aids benefit the controller (5.6.1).

37. This design supports the controller's development of strategies for dealing with short-term (tactical) and long-term (strategic) situations (5.6.2).

38. Changes in the situation or unusual events are clearly indicated and are not easy to miss (5.6.2).

39. The design allows sufficient time for the controller to perceive, integrate, project, and act upon ATC information (5.6.1, 5.6.3, 5.1.5).

CHAPTER 6. ISSUES IN ATC AUTOMATION

Elizabeth D. Murphy and Kim M. Cardosi

"Although controllers need assistance in dealing with increased air traffic, the form of this assistance must be carefully designed so that the controller's job remains one that human beings can perform successfully and with satisfaction."

FAA, 1990

TABLE OF CONTENTS

CHAPTER 6. ISSUES IN ATC AUTOMATION

LIST OF FIGURES AND TABLES

CHAPTER 6. ISSUES IN ATC AUTOMATION

The purpose of this chapter is to provide an introduction to human factors issues associated with automation. The term, "automation," refers to the use of computers to perform processing steps that were previously performed by controllers. Increasing automation of ATC will involve the use of faster and smarter computing machines than those currently in use. This chapter takes a look at automation from the controller's side of the human-computer interface.

Increased automation of the complex ATC system raises a set of critical human factors issues, which are explored in this chapter. Evaluating design products in light of these issues, can help to ensure that automated tools and capabilities are usable, operationally suitable, and supportive of the controller's job.

This chapter reviews the rationale for automation, discusses the benefits versus the potential drawbacks of ATC automation, and responds to questions on human factors issues in ATC automation.

6.1 TO AUTOMATE OR NOT TO AUTOMATE

As compared to results when process-control tasks are performed manually, automation makes possible greatly increased system productivity. Used properly, to assist but not replace the controller, ATC automation can complement human capabilities and compensate for human limitations. Across a continuum from a semi-manual operation to a fully-automated operation, several levels of automation are possible. At each successive level, the computer takes on more of the work formerly performed by the controller.

Finding a balance between the computer and the controller is a challenge to design engineers. If the human is being retained as a flexible decision maker, it is not always appropriate to automate a task or function just because it can be automated. Human capabilities, limitations, and needs must be taken into account when decisions are made about what and when to automate. The objective is to design and build a system in which the human and computer function together, using knowledge about the total system and its environment in planning, problem solving, and decision making (Hollnagel and Woods, 1983). To realize its potential, such a system should be built from a user-centered perspective.

Following are theoretically possible levels of automation in computer-controller interaction (Sheridan and Verplank, 1978):

Semi-manual:

- The controller does the whole job up to the point of turning it over to the computer to implement.

Semi-automated:

- The computer helps by determining the options.
- The computer helps determine options and suggests one, which the controller need not follow.
- The computer selects a course of action, and the controller may or may not act upon it.
- The computer selects a course of action and, if the controller approves, implements the action.
- The computer selects a course of action and informs the controller in plenty of time to abort or change it.
- The computer does the whole job but must tell the controller what it did.

Automated:

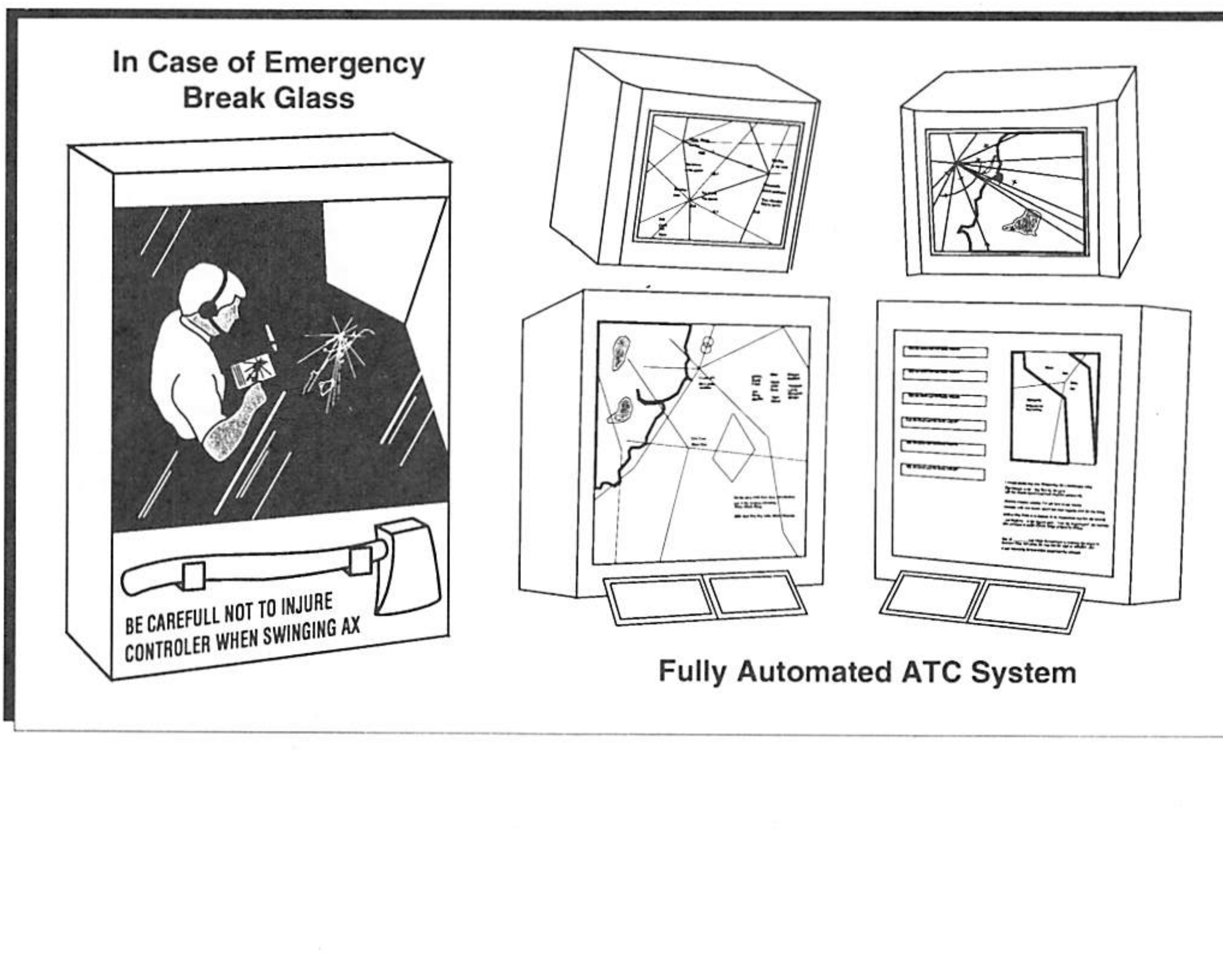
- The computer does the whole job and tells the Airspace Manager (AM) what it did only if the AM asks.

- The computer does the whole job and tells the AM what it did if it decides it should tell.

Fully automated:

- The computer does the whole job if it decides it should and tells the AM if it decides the AM needs to know.

While a completely automated air traffic control system may seem absurd, it is important to identify and understand the philosophy behind the allocation of functions between the controller and the computer for any automated ATC system.



6.1.1 Why do we need increased automation in ATC? Does increased automation guarantee a better system?

The most frequently mentioned reason for increasing ATC automation is the projected sharp rise in demand for ATC services from the aviation community. To meet these increasing demands for system capacity and operational efficiency, the argument goes, we need to reduce controller workload and increase controller productivity. The chosen means to this end is investment in automation in support of ATC operations and controller tasks.

Automation in and of itself does not necessarily reduce workload or increase productivity. Depending on how controller aids are designed, increased automation can produce just the opposite results: unacceptable increases in workload and reduced productivity. (See Chapter 8, Workload and Performance Measurement in the ATC Environment, where these concepts are defined and discussed in detail.)

"Software control of a system is not a panacea, nor is it a guarantee that the system will operate more safely or reliably than its predecessor nor that it will be more dependable or display any of the other attributes that we wish of a system."

Brown, 1993

People sometimes think that the best solution to a design problem is a software solution. Software (i.e., the encoded instructions to the computing machinery) can perform a wide range of routine-to-complex functions, but there is a limit to the extent to which computer control should be extended.

For example, software design flaws in a medical system have been blamed for patient deaths due to radiation overdoses delivered by a "state-of-the-art" system upgrade. This upgrade failed to provide sufficient safeguards, even though hardware-controlled safeguards had been present in the prior system (Brown, in press). Examples of software design flaws are rampant in aviation environments, as well (cf. Brown, in press; Weiner, 1987). When disaster occurs, however, accident analysts are likely to attribute the cause to operator error, rather than blaming design flaws (which are, in fact, human errors of a different order). **Attention to human factors issues is crucial if the potential benefits of increased ATC automation are to be realized.**

To realize the benefits of automation and to maintain workforce job satisfaction, automated ATC aids need to be designed from a user-centered perspective (Billings, 1991; Norman and Draper, 1986). The controller needs to be able to understand and trust any job aids provided as part of the system. The design of automated systems should be driven by the controller's information requirements.

6.1.2 What are the benefits and potential drawbacks of increased ATC automation?

Some of the benefits thought to occur from automation are listed in Table 6-1 (Gabriel, 1993).

These specific benefits derive from the general expectation for automation, that over time it will lower long-term costs and improve productivity.

Table 6-1. Expected Benefits of Automation

• Increased capacity (traffic throughput)	• Reduced staffing
• Improved human and system performance	• Improved management control
• Reduction in perceived workload	• Better integration of data from multiple sources
• Reduced training requirements	• Enhanced services
• Expanded capability to perform functions beyond human capabilities	• Lower task complexity
• Enhanced safety	

The problem with listing generic benefits of automation is that realizing the benefits depends largely on the usability, operational suitability, and acceptability of the specific automated capabilities provided to the workforce. Poorly designed automated systems that do not meet these criteria can have exactly the opposite effects of those listed.

Consider, for example, the expectation that automation will result in better integration of data from multiple sources. Realizing this benefit requires a large investment in determining the sources of data and actually integrating them. For the controller, appropriate data integration has the following benefits:

> "Many future forms of automation may possess characteristics hitherto thought of as uniquely human; these include intelligence, adaptability, flexibility, and innovation. An issue is how to apply these advances to air traffic control optimally to promote safety and efficiency."
>
> Hopkin, 1991, p. 8

- Only having to look at a single, localized display to acquire needed information.
- Being presented with compatible and comparable units and components for ease of working within various frames of reference.
- Avoiding unnecessary and time-consuming searches for information.
- Avoiding confusion between two or more similar data sets.

What often happens in the evolution of system design is a proliferation of unintegrated subsystems, each with its own unique user interface. The end result might resemble the real-world situation pictured in Figure 6-1, which depicts three different computer monitors, each with a unique keyboard, for use in a tower environment. This configuration has a high potential for negative transfer of skills between keyboards, leading to slower and potentially less accurate input to the system. Memory for one keyboard arrangement will interfere with memory for other keyboard arrangements, making substitution errors highly likely to occur. These errors can be called "human errors" on the controller's part, but they are actually the result to be expected from the inconsistent keyboard layouts. Rightfully, they are errors in system design.

Integrating separately developed components like these into the operational environment is one of the biggest human factors challenges facing the FAA. Although integration of individual subsystem components is typically addressed from a design-engineering perspective, lack of integration into existing operational and technological contexts is an obstacle to improved usability and productivity. Some of the other potential drawbacks of automation are listed in Table 6-2.

Figure 6-1. A Human-Factors Challenge: Need for Integration of Subsystems into the Operational Context

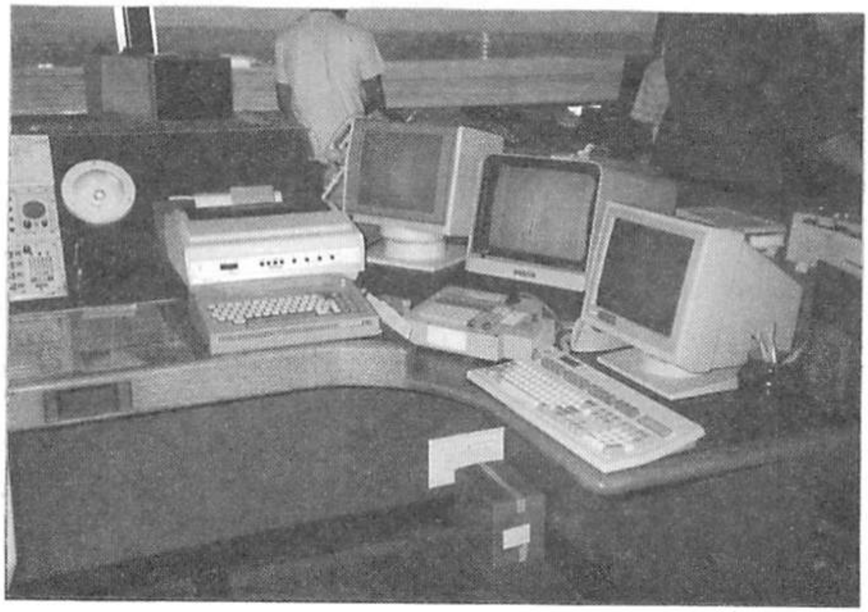


Table 6-2. Potential Drawbacks of Automation
Bainbridge, L. (1987). Ironies of Automation. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.), *New technology and human error* (pp. 271-283). New York: John Wiley.

• Loss of control skills and readiness to respond	• Unforeseen changes in human roles
• Unexpected negative interactions between human performance and computer performance	• Introduction of new forms of human error
• Inability of automated systems to resolve complex, critical problems	• Perception of full autonomy for what is really semi-autonomy
• Perception of automation as replacing operators	• Brittle computer performances (failure to degrade gradually)
• Overconfidence and lack of trust in the automation	• Increased boredom and loss of job satisfaction
• Changes in the sources and patterns of workload, with possible increases in integration operational workload	• Reduced efficiency and lack of productivity

These potential drawbacks of automation makes the designer's role especially crucial. Realizing the benefits of automation while avoiding the pitfalls is not an easy task. As discussed later in this chapter, requiring designers to take a user-centered approach to automation is the place to begin.

6.2 POTENTIAL EFFECTS OF ATC AUTOMATION

Increased automation in ATC can be expected to have many effects, some of which we can foresee, others of which will come as surprises. The enormous number of possible interactions between and among system components makes it impossible to predict every potential outcome of changes in equipment, software, workstations, and facilities. Because of changes in the automated tools and capabilities provided to the controller, it is likely that the controller's role in the system and methods of operation will change accordingly. In turn, it is likely that increased automation will affect the workload, job performance, and productivity of individual controllers and ATC teams. Many of these effects will be positive. Some are likely to be negative. Approaches to staffing and training will evolve in keeping with the changing needs brought about by increased automation.

6.2.1 How will increasing automation affect the controller's role and way of doing business?

The expectation is that, under increasing automation, the controller will gradually come to operate as an airspace manager. In this future scenario, the controller will monitor and supervise the automated system's operations and intervene when situations arise that are beyond the computer's capabilities. In the shorter term, with the introduction of the initial phases of the Advanced Automation System (AAS), the controller's role will remain largely what it is today. The transition to airspace manager will occur as automation becomes capable of devising reliable resolutions to conflict situations, and as more ground-air communication is automated via data link systems.

A critical task for the controller of the future will be evaluation of resolutions offered by an automated system (Hopkin, 1988; Weitzman, 1986b). **The controller will need**

> "A basic impact of automation on air traffic control can be to expand greatly the kinds of machine and human roles that are feasible, and the kinds of interaction between them."
>
> Hopkin, 1991, p. 10

to understand how the computer devises a resolution, that is, what data the computer accesses and by what rules alternative resolutions are selected and prioritized. In many cases, the computer will not have a complete picture of the situation because of the limited nature of the information it considers in generating a set of resolutions. For example, the computations may not be able to take into account weather, restricted airspace, or how inexperienced the pilot seems to be.

If the controller does not have a clear understanding of the information that the computer draws upon, and, to some extent, the algorithms used by the computer, then one of two errors is likely. Either the controller will think the system is more competent than it is and mistakenly place too much confidence in the system, or the controller will think the system is unreliable or malfunctioning (when it is not) and may not trust the system. To some extent, we have seen this in the controller's perception of TCAS (Traffic Alert and Collision Avoidance System). When aircraft first started operating with TCAS, it was not unusual for TCAS to generate an RA (a command to the pilot to maneuver the aircraft in a specified direction), in situations that did not generate a conflict alert for the controller or in situations in which the controller had everything under control (e.g., knew that an aircraft intended to level-off). Many controllers did not think TCAS operated properly or reliably because they did not understand the basis on which it would generate an RA.

While acting as airspace manager, the controller will always need to be aware of those aspects of the ATC situation that the automated resolution generator does not take into account. The controller will also need to be prepared to handle complex situations that exceed the computer's capabilities. The computer should be considered as a junior assistant, one that does its best with the information available to it, but one that operates within constraints and limitations.

IRONIES OF AUTOMATION

Many examples of the "ironies of automation" have been drawn from the nuclear power and chemical-processing environments as well as from flight-deck automation (Bainbridge, 1987). To the extent that ATC involves controlling the process of transitioning an aircraft from departure to arrival status, these same ironies are applicable to future, highly automated ATC environments.

"The more advanced a control system is, so the more crucial may be the contribution of the human operator" (p. 271)

"The designer who tries to eliminate the operator still leaves the operator to do the tasks which the designer cannot think how to automate" (p. 272).

"There is some concern that the present generation of automated systems, which are monitored by former manual operators, are riding on their skills, which later generations of operators cannot be expected to have" (p. 273).

"The automatic control system has been put in because it can do the job better than the operator, but yet the operator is being asked to monitor that it is working effectively" (p. 274).

A key human factors issue is the changing nature of the controller's situational awareness in future ATC systems. The reason for keeping the controller "in the loop" is to retain access to uniquely human abilities and expertise in judgment, problem solving, and decision making. If the computer is doing the situation monitoring and detecting conflict situations 20 minutes or more into the future, to what extent will the controller need to maintain and project the mental "picture" of the traffic situation? If aircraft are on direct (random) routes, will the controller be able to maintain such a picture? Will the "picture" become an image of the status of the automated system instead of an image of the traffic situation? If such changes take place, will the controller be prepared to step in when a complex situation arises that cannot be dealt with adequately by the automation or when the automation fails?

The weight of research evidence shows that **people are innately poor monitors in target-detection situations, i.e., when required to detect rarely occurring signals (or targets) in an unstimulating context** (e.g., Mackworth, 1950; Moray, 1986). As time spent in a purely monitoring mode increases, a person's ability to remain attentive to a "vigilance" task declines dramatically. As alertness degrades, and boredom or complacency sets in, a person's ability to detect and anticipate problems will be seriously impacted. For example, complacency can set in as early as 20 minutes into a flight simulation session when highly reliable automation controls system monitoring (Parasuraman, Molloy, and Singh, 1993).

"It is the nature of attention to wander, and no amount of exhortation can prevent this."

Hopkin, 1991, p. 12

In the ATC context, situation monitoring on the midnight shift comes closest to the conditions of the classic vigilance task. Not much is going on because traffic is extremely light. ATC monitoring, however, is not the controller's only task, and it is done for purposes other than detection of individual targets. Its primary purpose is to provide information needed by the controller to identify potential conflicts and to determine whether any control intervention is required. Even with this extended meaning of monitoring, however, the tendency toward boredom or complacency is a natural effect of low stimulation during periods of slow traffic.

This effect cannot be completely overcome by training or coffee, or corrected by admonishments to pay attention. It is simply a characteristic of the human cognitive system, which is always searching for stimulation. Lacking task-related stimulation, attention becomes distracted by random features of the environment, or the person becomes listless and groggy.

Another tendency of operators in highly automated systems is to become complacent and assume that the computer has everything under control (Weiner, 1987). This can be a serious problem if the operator becomes increasingly "out of touch" or "out of the loop," and later needs to step in and take action.

Some techniques exist to counter the natural decline in monitoring performance (Fisk and Schneider, 1981; Kantowitz and Sorkin, 1983; Wickens, 1984). Providing immediate feedback or "knowledge of results" to the monitor can improve performance. Introduction of false signals (e.g., false targets for detection) without operators' awareness of the deception has been helpful in maintaining alertness in other domains, but this is not an option for ATC. Enhancement of salient information by various forms of display manipulation reduces memory loading and improves monitoring performance. Results of a study on monitoring performance in an ATC-like task environment indicate that **detection of target events declines significantly after 60 minutes** (Thackray and Touchstone, 1982). In this study, there was no automated control of system monitoring as was the case in the study of automation-induced complacency, mentioned earlier (i.e., Parasuraman, Molloy, and Singh, 1993).

Monitoring of automated systems is associated with the term "supervisory control," which is used in human factors literature in reference to the operator's function as a monitor and supervisor of nearly autonomous computer functions (e.g., Mitchell, 1983; Mitchell and Saisi, 1987; Sheridan, 1988a and b). The major human factors issues in supervisory control focus on the operator's role and the long-term effects of that role on human performance:

- Demands for understanding system functioning and maintaining awareness of system status are likely to increase.
- Loss of alertness is likely, with a resulting delay in noticing and responding to system anomalies.
- Deterioration of operational skill and "sharpness" can be expected.
- Operators who thrive on dealing with sustained activity and challenging situations may be less satisfied with their jobs.
- Workforce selection criteria and training objectives will be affected.

- In the case of system degradation or failure, manual reversion may be difficult at best. Following system recovery, returning to automated operations may be problematic. In a highly automated system, the human operator may not be able to act as a backup because the number of operations being managed by the automation may far exceed the limits of human capabilities.

- Over time, operators are likely to become complacent and over-confident in highly reliable automation, making it more difficult to detect and respond to anomalous situations should they arise.

In general, increasing automation is associated with the introduction of new possibilities for human error. For example, acceptance of an inappropriate conflict resolution will be a new form of human error that can be expected to occur from time to time after the introduction of automated conflict resolution. (A particular conflict resolution may be perfectly appropriate within the limitations of the data considered by the resolution generator, but inappropriate within the broader context of the ATC situation, e.g., because the system has no information about weather or restricted airspace.) Recognition of the flaw(s) in a particular computer-generated resolution may occur only with hindsight, after it has been implemented and has resulted in a further problem or problems. Thus, **the limitations of the computer's information and advice must be clear to the controller.**

A key point is that **human error cannot be automated out of a system.** Some error forms will disappear, but others will appear to take their place. New systems must recognize this fact and be able to "trap" errors before they are propagated through the system. For any new system or subsystem, opportunities for human error must be identified, and an error-tolerant system must be designed to catch such errors while they can still be corrected.

Maintaining operational skills and efficiency is critical if the controller is expected to monitor computer functioning and assume manual control under abnormal conditions.

6.2.2 How does automation affect the controller's workload, job performance, and productivity?

ATC **workload** is traditionally defined in terms of the ATC situation: number of aircraft, complexity of aircraft mix, amount of communication and coordination needed, weather, and so forth. These situational variables describe essentially the demands of the situation, but workload is another matter from a human factors perspective. **Workload is a measure of the costs imposed by a task on the human system.**

Potential Effects on Workload. **The same demands can impose more or less workload on an individual controller depending on the controller's knowledge, skill, experience, and personal characteristics such as health and fatigue.** What may be high workload for a developmental air traffic controller is likely to be experienced as moderate to low workload by a Full Performance Level (FPL) controller who has 20 years of experience. This relative nature of workload will not change as the result of new automation.

When we look at workload, we are trying to get at the effort, particularly the mental effort, needed to manage the demands of the situation. Situational demands may be thought of as **objective workload**, while mental effort is the controller's **perceived workload**. Given the individual differences between controllers, it is the effects of automation on perceived workload that matters most. (See Chapter 8, Workload and Performance in the ATC Environment, for a full discussion of workload.)

The effects of increased automation on controller perceived workload cannot be predicted with any degree of accuracy. The FAA expects that increased automation will decrease some aspects of controller objective workload. A controller will presumably spend less time with an individual aircraft and perform fewer actions to transition that aircraft through a sector. The number of aircraft transitioning a sector per unit of time can be expected to increase if each aircraft can be handled in less time. "Number of aircraft" per time unit is, thus, both a measure of objective workload and a measure of productivity.

The actual effects of increased automation on perceived workload will depend to a large extent on the design of the user interface to the automated functions. The nature and extent of data entry, for example, will have varying effects on both objective and perceived workload. A user-centered approach to system design is needed to achieve acceptable levels of perceived workload.

Because perceived workload should be consistently moderate to maintain attention and alertness, an approach that lowers workload below moderate for long periods could be just as problematic as one that raises workload above moderate levels for long periods. During light traffic loads, perceived workload may decrease below the minimum needed to maintain attention and alertness. Following an extended period of underload, a sudden shift to higher workload can have negative effects on efficiency for tasks that require sustained attention (Huey and Wickens, 1993).

During periods of heavy traffic, workload may increase above moderate. Higher-than-moderate workload can be maintained for brief periods of time without adverse effects on performance, but adverse performance effects will occur immediately following a period of high workload. During that recovery period, there is a greater likelihood of human error.

The fact that people can sustain high workload for brief periods can give the impression that high workload should be sustainable for longer durations. **Requiring sustained high perceived workload over long periods can result in errors and be harmful to both the controller's physical and mental health.**

The perceived workload effects of increased automation depend on the scheme adopted for allocating functions to the computer and the controller, long before any software is written. Under a static scheme for deciding what the

"The combination of low control and high demand is the most destructive working condition for the human operator."

Hancock, 1991, p. 194

computer does and what the controller does, at what times, there is little or no flexibility for the controller to offload tasks to the computer. When the controller gets busy, for whatever reason, and has to start deciding what must be done now and what can wait, such a scheme does not permit the computer to pick up the slack. Similarly, when the controller is not very busy, the computer cannot give over any of its allocated functions to the controller.

Under a dynamic or adaptive approach to function and task allocation, however, the controller could give the computer more to do when s/he is busy and take over some tasks from the computer when perceived workload is light. Thus, **the perceived workload effects of increased automation depend, in large part, on decisions that are made early on regarding the level of flexibility in allocation of functions and tasks.** Pre-defining computer and human roles in ATC may not be necessary (Hopkin, 1991). Research indicates that a dynamic, adaptive approach to function allocation increases the productivity of the total human-computer system, whereas static function allocation yields lower levels of productivity (Rieger and Greenspan, 1983). Application of such an approach to ATC may be investigated at some point in the future.

A natural tendency on the part of controllers is to want the computer to handle all of the "easy" problems, reserving only the complex or unusual problems for controller resolution. Since the controller is the expert, this seems to make sense. However, coming on "cold" to complex situations is very different from solving difficult problems as they develop; it deprives the controller of familiarity with the situation that may be critical to successful problem solving.

Expertise develops and is maintained over years of dealing with all kinds of ATC situations. Through constant practice, the controller develops and modifies a mental "library" of problem situations and their resolutions. Research indicates that a characteristic of an expert as compared to a novice is an ability to recognize problem situations and to "know what to do" without going through a long process of generating and examining alternative solutions (Klein, 1989). It needs to be determined whether maintaining controller expertise (i.e., establishing, maintaining, and updating the mental "library") will require continued active involvement with all levels of ATC situations in the future.

Research is also needed to determine whether the future controller's ability to respond to a complex situation will require familiarity with the development of the situation over time. Coming on "cold" to a complex situation, with no warm-up on easier situations, may throw the controller into a state of high perceived workload mode that s/he is likely to be unprepared for. **Decisions about which easy or difficult tasks should be allocated to the controller or computer should be made with the benefit of data derived from systematic research, not solely on the basis of intuition.**

Potential Effects on Job Performance. Like workload effects, the possible job performance effects of increased automation are difficult to predict, since they depend in large part on the final configuration of roles for the controller and the computer in the system (i.e., the allocation of functions and tasks). In today's system, the chief yardstick of controller performance is success in handling aircraft in complex, high-volume traffic situations, but many other intangibles are involved. A controller who appears to be busy may not be performing as well as another controller who is proactively managing sector traffic according to a plan that keeps him or her from getting busy. If the job of separating aircraft becomes largely automated, the nature and definition of controller performance will change.

In general, increased automation of functions formerly performed by a human operator is associated with deterioration of skill and alertness, leading to a response lag when computers degrade or fail. In a "worst-case" scenario, the controller is likely to be deskilled, bored, and suddenly overloaded. If the design and the organizational climate encourage controller reliance on computer-generated resolutions, the controller's sharpness in evaluating resolutions is likely to decline over time, and acceptance of inappropriate resolutions is likely to increase, with possible catastrophic results.

Over time, what constitutes acceptable performance will undergo evolutionary changes. As the controller transitions into the role of airspace manager, criteria for evaluating human performance will need to change. Detection of system anomalies, rather than detection of conflict situations, may come to the fore of the controller's responsibilities. Thinking through the various workload and performance implications of new computer capabilities is a major responsibility. To the extent that these implications can be defined, field-site testing for operational suitability should be prepared to address them.

Potential Effects on Productivity. A common sense expectation is that providing the controller with an automated "assistant" will simplify what the controller needs to do in routine situations, thus increasing the controller's productivity. Predicting the effects of increased automation on productivity, however, depends on the working definition of productivity. Productivity can be defined in purely quantitative terms, such as system throughput (the number of aircraft transitioned through a sector per unit of time) or fuel savings.

A quantitative definition of productivity, however, does not consider important qualitative aspects of controller performance, such as the level to which safety was ensured or the orderliness of traffic flows maintained (Weitzman, 1986). The ATC mission is to ensure the safe, orderly, and expeditious flow of air traffic. This mission statement stresses quality (safety, orderliness, and expeditiousness) over quantity. A definition of productivity that stresses numbers of aircraft does not consider situational factors, which may conflict with moving aircraft through sectors (e.g., weather, restricted airspace going hot, emergencies, airline schedules, sector complexity, pilot experience). A quantitative approach to productivity ignores the uncertainties that controllers live with daily.

In terms of numbers of aircraft, productivity can be very high during one hour and very low the next hour, simply as a function of airline scheduling, weather, or any one of numerous other variables. Thus, the "average" number of aircraft moved through a sector can be a misleading estimate of controller productivity. Comparing the average numbers of aircraft moved through different sectors would be misleading unless situational variables were held constant.

Similarly, comparing old and new systems in terms of numbers of aircraft or fuel savings can be misleading unless the only change is in the automated capabilities provided to the controller. Differences in the controller's amount of practice or level of involvement, alone, could be responsible for apparent changes in controller productivity. **As compared to a passive role with few opportunities for practice, an active, involved role in a control process is known to be associated with better performance.**

In the early stages of a new system, automation can lower quantitative productivity while the workforce develops expertise with new tools and functions. If their logic is difficult for controllers to understand, expert systems and artificially intelligent decision-aiding tools will potentially slow down ATC operations while controllers try to determine, for example, the suitability of computer-generated conflict resolutions.

If controllers are even subtly encouraged by organizational pressures to accept computer-generated resolutions, with little more than a cursory evaluation, resolutions judged to be safe and reasonably good by the computer may worsen the problem or inject new problems into a situation. **The controller must have a clear understanding of both what the automation is doing and why, as a basis for developing confidence in the new tools and as a basis for intervening competently when necessary.**

Neither quantitative nor qualitative productivity will be enhanced if the automated tools have the *unanticipated* effect of increasing controller workload by requiring additional information processing and data entry. It cannot be assumed that any automated tool will, in itself, reduce controller workload. The tool must be designed to be usable, operationally suitable, and acceptable if it is to enhance quantitative and qualitative productivity (Harwood, 1994; Harwood and Sanford, 1994).

If selection and training costs are factors in quantitative productivity, it is possible that new systems may require more training and cognitive ability as compared to current systems. Further, safe procedural short cuts that the controller may take today to decrease delays may not be possible in an even more structured environment, with the computer enforcing various constraints, thus decreasing quantitative productivity (e.g., fuel savings).

Any lack of acceptance of the new system and subsequent non-use of new capabilities will also have adverse effects on quantitative productivity. Thus, changes in attitude, acceptance, and morale—changes not solely due to design differences—can be reflected in quantitative measures of productivity. Quantitative measures may vary inversely with qualitative measures, that is, numbers of aircraft processed may increase while safety decreases.

Qualitative measures of productivity are needed to assess the effects of increased automation on the safety, efficiency, and reliability of ATC services.

Just as the relationship between workload and performance should not solely be determined by a "common sense" approach, the effects of increased automation should not be predicted on the basis of intuition or expectations. Prior to operational test and evaluation, systematic modeling and simulation studies, as well as both formal and informal demonstrations, are needed to investigate the possible impact of increased automation on controllers' perceived workload, job performance, and quality assurance. Iterative field evaluation design and development can be highly effective in disclosing issues in these areas (Harwood and Sanford, 1994). Technical Interchange

Meetings (TIMs), attended by the design contractor, human factors specialist, and controllers, are helpful in reducing some of the risks associated with new product development.

6.2.3 How is increased automation likely to affect selection and training of controllers?

The criteria for selecting controller trainees are likely to change as a result of increased automation (Della Rocco, Manning, and Wing, 1991; Hopkin, 1991). The criteria for selecting Airspace Managers (AMs) of the future will not be exactly the same criteria used for selection of today's Air Traffic Control Specialists (ATCSs). The reason for these changes is that the profile of interests and abilities needed for the job is likely to change. Exactly what those criteria will become is uncertain.

Increased automation will affect the objectives and content of training and, possibly, the length of training required. For example, AMs of the future will need to have a fairly detailed understanding of automated functions, especially the automated conflict-detection-and-resolution capabilities. It will be crucial for AMs to understand the basis for and limitations of resolutions offered by the automation. The AM will need to be trained in new tasks and procedures, such as evaluating computer-generated resolutions within the context of the complete ATC situation.

On-going training and re-training may be needed to maintain manual ATC skills, knowledge, and abilities. Introduction of training in ATC Crew Resource Management (CRM) may be necessary to realize the full benefits of increased ATC automation.

Transition training becomes a major issue with increased automation. Questions like those following need to be considered (FAA, 1990):

- How should controllers be trained for the transition from the old to the new system?
- When should transition training begin and end?

- How can training help controllers make use of their skills from the old system (positive transfer) without carrying over habits that are no longer effective (negative transfer)?

The selection and training implications of requirements and design products need to be considered from the earliest stages of system concept exploration. Selection and training issues derive largely from the controller's role in the system, which is represented in the operational-requirements document. Identification and discussion of these issues is also a good way to validate the operational requirements. Support contractors can help derive selection and training implications from the operational requirements.

6.3 USER-CENTERED AUTOMATION

Few controllers who work in radar environments would want to return to the days prior to the introduction of radar into ATC. There is no question that the display of radar targets eases the burden on the controller of remembering aircraft positions and projecting those positions in relationship to each other. Similarly, despite the uncertainties associated with increased automation, few controllers would argue that additional automation is not needed to enhance their abilities to handle increased traffic demands. What is at issue is the specific design of new automated tools and capabilities.

6.3.1 What are the key differences between technology-centered and user-centered automation?

Automated aids can be designed from a technology-centered perspective or from a user-centered perspective. A technology-centered approach automates whatever functions it is possible to automate and leaves the human to do the rest. This places the operator in the role of custodian to the automation; the human becomes responsible for the "care and feeding" of the computer. In contrast, a user-centered approach provides the human with automated assistance that saves time and effort; the operator's task performance is *supported*, not *managed*, by computing machinery.

6.3.2 How can a user-centered approach to ATC automation be applied?

In a technology-centered approach, whatever automated functions that *can* be provided to the controller to assist in managing aircraft and increasing capacity *are* provided. This is different from a user-centered approach that aims to provide only those functions that the controller needs, based on information and task requirements. While the goals of the automation remain the same, i.e., to assist the controller in managing aircraft and improve system performance, the approach taken to automation will determine which functions are automated and how the controller will use these automated functions.

A user-centered approach is applied by systematically mapping user requirements to technical solutions (e.g., hardware and software that will satisfy requirements). A user-centered approach also requires attention to evaluating the usability, suitability, and acceptability of the design products. Such an evaluation benefits controllers by ensuring that the tools and capabilities provided by the computer system are in fact the ones needed by controllers to do their jobs.

Other ways in which user-centered automation can benefit controllers have been suggested in an analysis of aircraft automation and its effects on flight crews. The following recommendations are based on the conclusion of this report (Billings, 1991):

- Humans must remain in command of flight and air traffic operations.

  Automation can assist by providing a range of planning and control options.

- Human operators must remain involved in the task.

  Automation can assist by providing better integrated and more timely information.

- Human operators must be fully informed about the purposes and functioning of automated processes. At no time should the controller be wondering, "What is the automation doing or why is it doing that?," (as pilots sometimes have).

  Automation can be designed on the basis of a coherent model of its use, which can be explicitly communicated to users; automation should assist users by providing explanations of its intentions, recommendations, and actions.

- Human operators must have the information needed to anticipate and resolve problems. Automation can assist by monitoring trends, providing decision support, and making required information accessible when it is needed.

The specification of requirements and review of design products can go a long way toward ensuring that ATC automation is user-centered, not technology-centered.

6.3.3 What are user-centered objectives for ATC automation?

There are three higher-level objectives for ATC Automation: usability, operational suitability, and workforce acceptance (Harwood, 1994).

Usability is a function of measurable ease-of-use outcomes, such as the ease of navigating through a menu structure, ease of remembering data-entry requirements (e.g., command formats and sequences), and ease of locating specific items on a visual display. Overall usability depends on several interdependent factors, such as the reliability of system performance, the organization of the user interface, and maintainability in the field (Gould, 1988). Establishing usability goals and evaluating design products against these goals is a key system development activity in which controllers should participate. (A useful discussion of usability goals and usability testing can be found in the *Handbook of Human-Computer Interaction* (Whiteside, Bennett, and Holtzblatt, 1988. Usability testing is also described in Chapter 7 of this document, Computer-Human Interface [CHI] Considerations)

In order for a design to be **operational suitable,** it must support the controller's effective and efficient planning, maintenance of situational awareness, separation of aircraft, and performance of other ATC tasks. Support is provided to the controller by the design primarily in the form of information about the ATC situation and the status of ATC equipment and facilities. A design can be usable but operationally unsuitable if it does not meet the controller's requirements for appropriate and timely information. Methods have been developed for quantifying operational suitability (Phillips, 1987). Early indications of operational suitability can surface in systematic field evaluations (Harwood and Sanford, 1994). The ultimate tests of operational suitability occur later during rigorous site-implementation evaluations.

Workforce acceptance derives, in part, from a design's reliability, usability, and operational suitability. Acceptance also depends on the impact that new ATC technology has on controller job satisfaction (cf. Mackie and Wylie, 1988). It may be the case that sources of job satisfaction in the current system are disrupted or removed by the new technology (Hopkin, 1991 and 1992). Further, there may be fewer opportunities for individual recognition by one's peers and managers (Hopkin, 1991 and 1992).

As a basis for ensuring that the design is usable, operationally suitable, and acceptable to the workforce, it may be useful to specify human factors objectives for the new ATC automation and its operations. Following is a set of objectives for automation that should be considered during requirements and development:

- Transparency of underlying software operations so that the controller does not need to be aware of the inner workings of the computer, but perceives a smooth, responsive operation.

- Error-tolerance and recoverability.

- Consistency with controllers' expectations.
- Compatibility with human capabilities and limitations.
- Ease of reversion to lower levels of automation and of returning to higher levels of automation.
- Ease of handling abnormal situations and emergencies.
- Ease of use and learning.

Of course, for any specific design, these objectives will need to be further specified in terms of individual system functions and operational objectives.

Transparency of underlying operations. Internal software operations will not be apparent to the controller if designers have paid attention to meeting usability goals. For example, if it is difficult for the controller to maintain a sense of orientation within a user-interface menu structure, it is probably the case that programming convenience has taken priority over usability goals and good design practice.

Software operations will be transparent to the controller if ATC tasks can be performed naturally or intuitively without needing to pay conscious attention to underlying computational structures. Thus, transparency also relates to operational suitability. Similarly, a non-transparent (or opaque) user interface is likely to have adverse consequences for job satisfaction because it will be difficult to understand and use. It will foster negative attitudes, such as frustration and resentment. Such problems should be identified early in design development so that systems sent to the field for early field evaluation and later site testing will be more usable, operationally suitable, and acceptable than they would otherwise be.

Error tolerance and recoverability. The objectives of error tolerance and recoverability fall primarily under operational suitability, but they are related to the other goals as well. **Error-tolerant designs allow for the conceptual**

equivalence of different commands (e.g., Exit and Quit) and will accept any of the pre-defined equivalents. Further, error-tolerant designs anticipate possible user errors in data entry and include capabilities to catch errors before they propagate through the system (Hollnagel, 1993; Rouse and Morris, 1985). **As a matter of course, error-tolerant designs query the user at critical choice points (e.g., "Are you sure you want to delete this flight plan?").** Under such a regime, recovering from error is simple and enhances usability.

Consistency with controllers' expectations. Automation that does not assess a situation in the same way that a controller does or take the same actions that a controller would is likely to elicit skepticism or inappropriate actions from the controller. Controllers will be more likely to accept, trust, and use, automated functions that handle situations in the same ways that a controller would. To be most useful to the controller, the design of automated functions should take into account air traffic procedures and operations (e.g., airspace and traffic management restrictions, rules for assignment of flight levels, etc.), so that the automation does not violate common practices.

Compatibility with human capabilities and limitations. As previously discussed, people are, by nature, poor monitors of automation. Controllers cannot be put in the position of passively monitoring the automation and then be expected to be able to detect a failure, determine the problem, and take appropriate action. Automated functions must be compatible with the controllers' abilities and limitations.

Ease of reversion to lower levels of automation. After operating with a highly automated system for some time, returning to lower levels of automation (e.g., as in the case of system degradation or failure) may be problematic for several reasons. First, controllers may experience a loss of proficiency. If, for example, an automatic conflict-resolution function is used so extensively that controllers rarely have to resolve potential conflicts between aircraft, their problem-solving skills and strategies are not likely to be as finelytuned as when they had to prevent and solve the problems on their own. Complex skills suffer from disuse. Provisions for maintaining these skills (such as recurrent training) must be considered in these circumstances. Second, a controller's ability to respond may be affected by a loss of situational awareness. Controllers who have been left "out of the loop" will lose their situational awareness and possibly become bored and inattentive. This will increase the time required for the controller to take over from the automation and will be detrimental to the quality of the controller's response. Automated systems have to include an active, involved role for the controller, so that the controller always has enough situational awareness to handle whatever problems arise. Finally, the controller will not be able to recover from an automation failure if the number and complexity of operations being managed by the automation exceeds the controller's capabilities.

Ease of handling abnormal situations and emergencies. Controllers will always need to have the information and the means needed to intervene in emergencies or abnormal situations. The controller should never be denied access to controls or critical information that may be needed to respond in these situations. For example, even a system that assumed that the automation would handle all communications with the aircraft would be unacceptable if the controller were not left with the means to communicate with an aircraft when necessary and also had access to the critical flight information of all the aircraft in the sector.

Ease of use and learning. Automated functions should be easy to learn and easy to use. The effect of the use of these functions on training requirements should be examined. Complex systems may require extensive training on the various operational modes and limitations. Recurrent training may be advisable to deal with problems that are possible, but expected to occur rarely.

6.3.4 What is the basic set of ideas behind user-centered automation in ATC?

Before evaluating automated systems, it is useful to think in terms of a user-centered philosophy of ATC automation. Such a concept views the automation as assisting, not replacing, the controller. It recognizes that there is more of an overlap than ever in the capabilities of the human and the computer and does not divide tasks between controller and computer according to a simplistic approach, which depicts the computer as "better-than-the-human" at some tasks and the human as "better-than-the-computer" at other tasks. This philosophy advocates a task-allocation strategy that benefits the controller, rather than leaving to the controller only those tasks (or bits of tasks) that the designer finds difficult to automate.

The philosophy of user-centered ATC automation emphasizes the need for allocating functions in such a way that the controller's situation awareness is continually maintained and updated, and that the controller's expertise and creativity are exercised regularly. It suggests that dynamic or adaptive function allocation be considered as a technical solution. Under adaptive function allocation, the controller can prepare for periods of heavy traffic by allocating normally manual functions to the computer; and, s/he can prepare for lighter traffic periods by taking on some functions (or portions of functions) normally accomplished by the computer. (See Rieger and Greenspan, 1983, for a detailed discussion of adaptive function allocation. Also, see the *National Plan*, Project 2.1.B in the ATC Domain, for a description of research needed on alternative approaches to ATC automation (FAA, 1990).) Under a static allocation policy, flexibility of assigning functions is not possible.

6.4 EVALUATION ISSUES

The introduction of any new technology requires systematic testing and evaluation of its usability, operational suitability, and acceptability to the workforce. Usability testing investigates the perceptual and physical human factors of the user interface, such as display formatting, use of graphics, and design of the user-computer dialogue as well as the ergonomics of the workstation and the facility.

Operational-suitability testing evaluates the appropriateness of the design for use specifically in ATC applications.

6.4.1 What are the design-evaluation implications of increased automation?

"The burdens associated with managing automation can sometimes outweigh the potential benefits of automation to improved system performance."

Kirlik, 1993, p. 221

Both usability testing and operational-suitability testing assess the user-centeredness of design products. A clearly-identified link needs to be established between the controllers' information requirements and the design of information displays. A prototype display design must satisfy all documented information requirements. Automated human-factors engineering tools are available to assist in mapping the documented requirements to the actual design. Similarly, the design's ability to support operational tasks needs to be verified.

To some extent, user acceptance will depend on usability and operational suitability. The new technology will be judged for its reliability, the extent to which it works consistently "as advertised" in operational settings, and the extent to which it actually saves time and effort. User acceptability of the new automation, however, may also be influenced by the "incidental consequences" of information technology and automation, for example, effects on self-esteem, job satisfaction, and professional standing among colleagues (Hopkin, 1980; 1991 and 1992).

Investigations of such incidental effects focus on answering two questions (Harwood, 1994):

- Does the new automation disrupt aspects of the task that are sources of job satisfaction in the current system?

- Is there potential for new situations to emerge that will make the job less satisfying and preclude opportunities for individual achievement?

As a basis for answering these questions, it is necessary to identify sources of job satisfaction in the current system, such as problem solving, interaction with colleagues, and so forth. In keeping with the FAA policy of system enhancement with minimal disruption to operational personnel, the new technology should support or enhance aspects of the job that are satisfying and provide opportunity for individual achievement. If it does not do so, the design will be more technology-centered than it is user-centered.

Specific guidance on design evaluation is provided in other chapters of this handbook, for example, in sections on evaluation of visual and auditory displays [Chapter 7, Computer-Human Interface (CHI) Considerations], evaluation of decision-support tools (Chapter 5, Human Information Processing), and evaluation of workstation design (Chapter 9, Workstation and Facility Design and Evaluation).

6.5 REFERENCES AND SUGGESTIONS FOR FURTHER READING

American National Standards Institute. (1988). *American national standard for human factors engineering of visual display terminal workstations* (ANSI/HFS Standard No. 100-1988). Santa Monica, CA: Human Factors Society.

Ammerman, H. L., and Jones, G. W. (1988). *ISSS impact on ATC procedures and training* (CDRL C108). Colorado Springs, CO: Computer Technology Associates.

Bainbridge, L. (1987). Ironies of automation. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.), *New technology and human error* (pp. 271-283). New York: John Wiley.

Bikson, T. K. (1987). Cognitive press in computer-mediated work. In G. Salvendy, S. L. Sauter, and J. J. Hurrell, Jr. (Eds.), *Social, ergonomic and stress aspects of work with computers* (pp. 353-364). Amsterdam: Elsevier.

Billings, C. E. (1990). A concept of human-centered automation. *Abstracts of AIAA/NASA/FAA/HFS Symposium, Challenges in Aviation Human Factors: The National Plan* (pp. 3-5). Washington, DC: American Institute of Aeronautics and Astronautics.

Billings, C. E. (1991). *Human-centered aircraft automation: A Concept and guidelines* (NASA TM 103885). Moffett Field, CA: NASA-Ames Research Center.

Brown, M. L. (in press). *Risk and risk assessment: Software applications in safety-critical systems.*

Brown, S. J. (1987). The Challenges of future ATC automation and airspace design. In R. S. Jensen (Ed.), *Proceedings of the Fourth International Symposium on Aviation Psychology* (pp. 186-189). Columbus, OH: Ohio State University.

CTA INCORPORATED. (1990). *Initial Sector Suite System impact on air traffic control procedures* (Final, Report No. T104). Colorado Springs, CO: Author.

Carlson, L. S., and Rhodes, L. R. (1990). *Changes in controllers' work activities in AERA 2 and implications for controller training* (MITRE TR No. MP-89W00021). McLean, VA: MITRE.

Della Rocco, P., Manning, C. A., and Wing, H. (1991). Selection of air traffic controllers for automated systems: Applications from today's research. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and system issues in air traffic control* (NATO Advanced Science Institute, Series F, Vol. 73, pp. 429-452). New York: Springer-Verlag.

Department of Defense. (1989). *Military standard: Human engineering design criteria for military systems, equipment and facilities* (MIL-STD 1472D). Washington, DC: Author.

Ephrath, A. R., and Young, L. R. (1981). Monitoring vs. man-in-the-loop detection of aircraft control failures. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures*. New York: Plenum.

Federal Aviation Administration (FAA). (1990). *The National plan for aviation human factors*. Washington, DC: Author.

Fisk, A. D., and Schneider, W. (1981). Controlled and automatic processing during tasks requiring sustained attention. *Human Factors, 23*, 737-750.

Gabriel, R. F. (1993). Cockpit automation. In K. M. Cardosi and M. S. Huntley (Eds.), *Human factors for flightdeck certification personnel* (DOT/FAA/RD-93/5). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Gould, J. D. (1988). How to design usable systems. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 757-789). New York: Elsevier/North Holland.

Grandjean, E. (1988). *Fitting the task to the man: An ergonomic approach* (3rd ed.). London: Taylor and Francis.

Hancock, P. A. (1991). The Aims of human factors and their application to issues in automation and air traffic control. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and system issues in air traffic control* (NATO Advanced Science Institutes Series F, Vol. 73, pp. 187-200). New York: Springer-Verlag.

Harwood, K. (1994). Defining human-centered issues for verifying and validating air traffic control systems. In J. A. Wise, V. D. Hopkin, and P. Stager (Eds.), *Verification and validation of complex systems: Human factors issues* (NATO Advanced Science Institute, Series F: Vol. 110, pp. 115-129). New York: Springer-Verlag.

Harwood, K., and Sanford, B.,(1994). Evaluation in context: ATC automation in the field. In J. A. Wise, V. D. Hopkin and J. D. Garland (Eds.), *Human factors certification of advanced aviation technologies* (Aviation Human Factors Series, pp. 245-262). New York: Springer-Verlag.

Hollnagel, E. (1993). *The Phenotype of erroneous actions. International Journal of Man-Machine Studies, 39*, 1-32.

Hollnagel, E., and Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies, 18*, 583-600.

Hopkin, V. D. (1988). *Human factors aspects of the AERA 2 program*. Farnborough, U. K.: Royal Air Force Institute of Aviation Medicine.

Hopkin, V. D. (1987). Human factors implications of progressive air traffic control automation. In R. S. Jensen (Ed.), *Proceedings of the Fourth International Symposium on Aviation Psychology* (pp. 179-185). Columbus, OH: Ohio State University.

Hopkin, V. D. (1982). *Human factors in air traffic control* (AGARD-AG-275). Neuilly Sur Seine, France: NATO, Advisory Group for Aerospace Research and Development.

Hopkin, V. D. (1992). Human factors issues in air traffic control. *Human Factors Society Bulletin, 33(6)*, (pp. 1-4).

Hopkin, V. D. (1991). The Impact of automation on air traffic control systems. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and systems issues in air traffic control* (NATO Advanced Science Institutes Series F, Vol. 73, pp. 3-19). New York: Springer-Verlag.

Hopkin, V. D. (1989). Man-machine interface problems in designing air traffic control systems. *Proceedings of the IEEE, 77*, 1634-1642.

Hopkin, V. D. (1980). The Measurement of the air traffic controller. *Human Factors, 22*, 547-560.

Huey, B. M., and Wickens, C. D. (Eds.), (1993). *Workload transition: Implications for individual and team performance*. Washington, DC: National Academy Press.

Hunt, V. R., and Zellweger, A. (1987). The FAA's Advanced Automation System: Strategies for future air traffic control systems. *Computer, 20*, 19-32.

International Civil Aviation Organization (ICAO) (1993). *Human factors digest No. 8: Human factors in air traffic control* (Circular 241-AN/145). Montreal: Author.

Johannsen, G., Rijnsdorp, J. E., and Sage, A. P. (1983). Human-system interface concerns in support system design. *Automatica, 19*, 595-603.

Kantowitz, B. H., and Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships*. New York: John Wiley.

Kirlik, A. (1993). Modeling strategic behavior in human-automation interaction: Why an "aid" can (and should) go unused. *Human Factors, 35*, 221-242.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in man-machine systems research* (Vol. 5, pp. 47-92). Greenwich, CT: JAI Press.

Klein, G. A., Calderwood, R., and Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 576-580). Santa Monica, CA: Human Factors Society.

Klein, G. A., Calderwood, R., and MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man and Cybernetics, 19(3)*, 462-472.

Klein, G. A., Orasanu, J., Calderwood, R., and Zsambok, C. E. (Eds.), (1993). *Decision making in action: Models and methods.* Norwood, NJ: Ablex.

Kleinmuntz, D. N., and Schkade, D. A. (1989). *The cognitive implications of information displays in computer-supported decision making* (Working paper No. 2010-88). Cambridge, MA: Massachusetts Institute of Technology, Sloan School of Management.

Lehner, P. E., and Kralj, M. M. (1988). Cognitive impacts of the user interface. In J. A. Hendler (Ed.), *Expert systems: The user interface* (pp. 307-318). Norwood, NJ: Ablex.

Mackie, R. R., and Wylie, C. D. (1988). Factors influencing acceptance of computer-based innovations. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 1081-1106). New York: Elsevier/North Holland.

Mackworth, N. H. (1950). Researches on the measurement of human performance. Reprinted in H. W. Sinaiki (Ed.), (1961), *Selected papers on human factors in the design and use of control systems* (pp. 174-331). New York: Dover. (Reprinted from Medical Research Council Special Report Series No. 268. London: HM Stationery Office.)

Meister, D. (1987). Systems design, development, and testing. In G. Salvendy (Ed.), *Handbook of human factors*. New York: John Wiley.

Mitchell, C. M. (1983). The Human as supervisor in automated systems. In C. M. Mitchell, P. M. Van Balen, and K. L. Moe (Eds.), *Human factors considerations in system design* (NASA Conference Publication 2246, pp. 257-287). Greenbelt, MD: NASA Goddard Space Flight Center.

Mitchell, C. M., and Miller, R. A. (1983). Design strategies for computer-based information displays in real-time control systems. *Human Factors, 25*, 353-369.

Mitchell, C. M., and Saisi, D. L. (1987). Use of model-based qualitative icons and adaptive windows in workstations for supervisory control tasks. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-17*, 573-593.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. II: Cognitive processes and performance, pp. 40-1 - 40-51). New York: John Wiley.

Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Systems, 27*, 527-539.

Narborough-Hall, C. S. (1987). Automation implications for knowledge retention as a function of operator control responsibility. In D. Diaper and R. Winder (Eds.), *People and computers II* (pp. 269-282). New York: Cambridge University Press.

Nazareth, D. L. (1989). Issues in the verification of knowledge in rule-based systems. *International Journal of Man-Machine Studies, 30*, 255-271.

Norman, D.A. (1988). The Design of everyday things. New York: Doubleday/Currency.

Norman, D. A., and Draper, S. W. (Eds.) (1986). *User-centered system design*. Hillsdale, NJ: Erlbaum.

Office of the NASA Aviation Safety Reporting System (ASRS). (June 1992). Gidgets, gadgets, and gizmos...Issues in automation. *Callback, 157*, Moffett Field, CA: ASRS.

Parasuraman, R., Molloy, R., and Singh, I. L. (1993). Performance consequences of automation induced "complacency." *International Journal of Aviation Psychology, 3*, 1-23.

Pasmooij, C. K., Opmeer, C. H. J. M., and Hyndman, B. W. (1976). Workload in air traffic control. In T. B. Sheridan and G. Johannsen (Eds.), *Monitoring behavior and supervisory control* (pp. 107-118). New York: Plenum.

Perrow, C. (1984). *Normal accidents*. New York: Basic Books.

Phillips, M. D. (1987). The quantification of operational suitability. *IEEE Computer, 20*, 63-71.

Pozesky, M. T. (1990). Air traffic control automation challenges. *Abstracts of AIAA/NASA/FAA/HFS Symposium, Challenges in Aviation Human Factors: The National Plan* (pp. 8-11). Washington, DC: American Institute of Aeronautics and Astronautics.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York: North Holland.

Rasmussen, J., Duncan, K., and Leplat, J. (Eds.) (1987). *New technology and human error*. New York: John Wiley.

Rasmussen, J., and Goodstein, L. P. (1988). Information technology and work. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 175-201). New York: North Holland.

Rieger, C. A., and Greenspan, J. S. (1983). The effects of dialogue-based task allocation on system performance in a computer-aided air traffic control task. *Behavior Research Methods and Instrumentation, 15*, 208-212.

Rouse, W. B., and Boff, K. R. (Eds.). (1987). *System design: Behavioral perspectives on designers, tools, and organizations*. New York: North Holland.

Rouse, W. B., and Morris, N. (1985). *Proceedings of the Second IFAC/IFIP/IFORS/IEA Conference on Analysis, Design, and Evaluation of Man-Machine Systems* (pp. 19-24). Varese, Italy: International Federation of Automatic Control.

Shaiken, H. (1986). The human impact of automation. *IEEE Control Systems Magazine, 6 (Dec.)*, 3-6.

Sheridan, T. B. (1988a). Human and computer roles in supervisory control and telerobotics. In L. P. Goodstein, H. B. Anderson, and S. E. Olsen (Eds.), *Tasks, errors, and mental models* (pp. 149-160). New York: Taylor & Francis.

Sheridan, T. B. (1988b). Task allocation and supervisory control. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 159-173). New York: Elsevier/North Holland.

Sheridan, T. B. (1990). Theoretical approaches to human centered automation. *Abstracts of AIAA/NASA/FAA/HFS Symposium, Challenges in Aviation Human Factors: The National Plan* (pp. 15-17). Washington, DC: American Institute of Aeronautics and Astronautics.

Shneiderman, B. (1992). *Designing the user interface* (2nd ed.). New York: Addison-Wesley.

Simolunas, A. A., and Bashinski, H. S. (1991). Computerization and automation: Upgrading the American air traffic control system. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and system issues in air traffic control* (NATO ASI Series F, Vol. 73, pp. 31-38). New York: Springer-Verlag.

Sorkin, R. D., and Robinson, D. E. (1984). *Alerted monitors: Human operators aided by automated detectors* (DOT/OST/P-34/85-021). Washington, DC: Department of Transportation/Office of the Secretary, University Research Program.

Stein, E. S. (1987). Where will our air traffic controllers be in the Year 2001? In R. S. Jensen (Ed.), *Proceedings of the Fourth International Symposium on Aviation Psychology* (pp. 195-201). Columbus, OH: Ohio State University.

Swain, A. L. (1987). Four human factors problems for system operators as dependence on automation increases. In G. Salvendy (Ed.), *Cognitive engineering in the design of human-computer interaction and expert systems* (Advances in Human Factors/ Ergonomics, Vol. 10B, pp. 105-112). Amsterdam: Elsevier.

Taylor, L. T. (1991). Future air traffic control (ATC) systems require a training paradigm shift. *36th Annual Air Traffic Control Association Conference Proceedings* (pp. 407-413). Arlington, VA: Air Traffic Control Association.

Thackray, R. I. (1980). *Boredom and monotony as a consequence of automation: A consideration of the evidence relating boredom and monotony to stress* (FAA-AM-80-1). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

Thackray, R. I., and Touchstone, R. M. (1989). *A comparison of detection efficiency on an air traffic control monitoring task with and without computer aiding* (DOT/FAA/AM-89/1). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

Thackray, R. I., and Touchstone, R. M. (1982). *Performance of Air Traffic Control Specialists (ATCS's) on a laboratory radar monitoring task: An exploratory study of complacency and a comparison of ATCS and non-ATCS performance* (FAA-AM-82-1). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

Thorning, A. G. (1989). *Research in automation for air traffic control.* (SAE Technical Paper 881470). Warrendale, PA: Society of Automotive Engineers.

Weiner, E. L. (1987). Fallible humans and vulnerable systems: Lessons learned from aviation. In J. A. Wise and A. Debons (Eds.), *Information systems: Failure analysis* (pp. 163-181). Berlin: Springer-Verlag.

Weitzman, D. O. (1986a). *Definition of productivity.* (MITRE Memorandum W42-M0447). McLean, VA: MITRE.

Weitzman, D. O. (1987). *Human problem solving in complex risk technologies: Cognitive activities in air traffic control operations* (Draft). Working paper prepared for the 1988 AAAS Annual Meeting, Boston, MA.

Weitzman, D. O. (1986b). *The human element in AERA 3: The uncertainties* (MITRE Memorandum No. W42-M0602). McLean, VA: MITRE Metrek Division.

Weston, R. C. W. (1983). Human factors in air traffic control. *International Journal of Aviation Safety, 1*, 94-104.

Westrum, R. (1991). Automation, information, and consciousness in air traffic control. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and system issues in air traffic control* (NATO Advanced Science Institutes Series F, Vol. 73, pp. 367-380). New York: Springer-Verlag.

Whitfield, D., Ball, R. G., and Ord, G. (1980). Some human factors aspects of computer-aiding concepts for air traffic controllers. *Human Factors, 22*, 569-580.

Whiteside, J., Bennett, J., and Holtzblatt, K. (1988). Usability engineering: Our experience and evolution. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 791-817). New York: Elsevier/North Holland.

Wickens, C. D. (1984). *Engineering psychology and human performance.* Columbus, OH: Charles E. Merrill.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: Harper-Collins.

Wickens, C. D., and Kessel, C. (1981). Failure detection in dynamic systems. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp. 155-168). New York: Plenum.

Wiener, E. L. (1987). Fallible humans and vulnerable systems: Lessons learned from aviation. In J. A. Wise and A. Debons (Eds.), *Information systems: Failure analysis* (pp. 163-181). Berlin: Springer-Verlag.

Wise, J. A., Hopkin, V. D., and Smith, M. L. (Eds.). (1990). *Automation and system issues in air traffic control* (NATO Advanced Science Institutes Series F, Vol. 73). New York: Springer-Verlag.

Woods, D. D. (1985). Cognitive technologies: The Design of joint human-machine cognitive systems. *The AI Magazine, Fall*, 86-92.

Woods, D. D., and Hollnagel, E. (1987). Mapping cognitive demands in complex problem-solving worlds. *International Journal of Man-Machine Studies, 26*, 257-275.

APPENDIX 6A. CHECKLIST ITEMS

1. To acquire needed information, the controller only needs to look at a single, localized display i.e., switching back and forth between two or more displays is not necessary to perform an individual task. (6.1.2).

2. Increased ATC automation results in better integration of data from multiple sources (6.1.2).

3. Separately developed subsystems are effectively integrated into the operational environment (6.1.2).

4. With this design, the controller can quickly find the necessary information (6.1.2).

5. The limitations of the computer's information and advice are clear to the controller (6.2.1).

6. This design provides an active, involved role for the controller (5.3.2, 5.6.2, 6.2.1, 6.3.3).

7. The controller does not have to solely monitor automated functions for more than 20 to 30 minutes (6.2.1, 5.3.2).

8. This design will not induce complacency (6.2.1).

9. With this design, reversion to manual control will be easy; that is, the controller will have no problem stepping in when the computer cannot deal with a complex situation or when the automation fails (6.2.1, 6.3.3).

10. After system recovery from degradation or failure, a smooth return to automated operations will be possible (6.2.1).

11. With this design, the controller will be able to build and maintain sufficient situational awareness (6.2.1, 6.3.3).

12. Provisions have been made to help controllers maintain operational skills and efficiency (6.2.1, 6.3.3).

13. This design provides flexibility for the controller to offload tasks to the computer during busy periods and to take over some tasks from the computer during slow periods (6.2.2).

14. With this design, the amount and complexity of data entry is about the same as was required in the previous system (6.2.2).

15. Automated features provide explanation of their intentions, recommendations, and actions in ways that are readily understood by controllers (6.3.2).

16. Automated features behave in ways that are consistent with controller expectations (6.3.3).

17. With this design, data-entry errors can be caught and corrected before they propagate through the system (6.3.3).

18. This design makes it easy to recover from data-entry errors (6.3.3).

19. This user interface design queries the controller at critical choice points, e.g., "Are you sure you want to delete this flight plan?" (6.3.3).

CHAPTER 7.
COMPUTER-HUMAN INTERFACE (CHI) CONSIDERATIONS

Donna L. Anastasi, Brian P. Hill, Elizabeth D. Murphy, Kim M. Cardosi, Jeremy A. Guttman, and Paola Amaldi

"Users prefer to feel a sense of mastery and control over any tool at their disposal, and the computer is no exception...Users will soon gain a sense of mastery if the interface is simple, predictable, and consistent."

Mayhew, 1992

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 7. COMPUTER-HUMAN INTERFACE (CHI) CONSIDERATIONS

Designers of complex, automated system face two basic technical problems when they prepare to design the user interface:

- How to present information **to** the user.

- How to acquire information **from** the user.

Solutions to these problems can make the user's job relatively more easy or more difficult.

Critical human factors objectives for CHI design are **usability, operational suitability**, and **workforce acceptance** (Harwood, 1994). Usability is essentially the extent to which system hardware and software are compatible with human capabilities and limitations. Hardware and software may be usable but not operationally suitable, that is, not appropriate for the user's context-specific goals and tasks. Workforce acceptance requires both usability and suitability as well as a sense that job satisfaction will be enhanced. Acceptance also depends heavily on the reliability of the computer system.

The challenge to both designers and evaluators is to ensure that these critical objectives are met. Material presented in this chapter provides guidance and recommendations in many areas of CHI design. It is geared toward CHI design and evaluation in the ATC context.

7.1 OVERVIEW

Definitions of the computer-human interface range from highly specific to highly abstract. In concrete and specific terms, the computer-human interface is made up of those physical parts of the system, i.e., the computer screens and input devices, that support the transfer of information between the computer and the human. The information that is transferred provides an interface to the computer's functions (i.e., the tools and capabilities available to the user). More abstractly, we may speak of the "cognitive interface," where

Usability Guidelines:

• Use a simple and natural dialogue.

• Provide an intuitive visual layout.

• Speak the user's language.

• Minimize the user's memory load.

• Be consistent.

• Provide feedback.

• Provide clearly marked exits.

• Provide shortcuts.

• Provide good help.

• Allow user customization.

• Minimize the use of effects of modes.

• Support input device continuity.

Karat, Campbell, and Fiegel, 1992, p. 399

information presented to the human is interpreted within the context of the user's tasks and goals (Chignell, Hancock, and Loewenthal, 1989). In most current systems, the cognitive interface is in the mind of the user.

This chapter focuses primarily on the design and evaluation of the physical and functional computer-human interfaces in ATC applications. **The objective of design and evaluation activities is to make the physical and functional interfaces as compatible as possible with the cognitive interface.** The cognitive interface must be understood **before** the physical or functional interface is designed.

In this chapter, we identify the human factors issues associated with the design and evaluation of ATC displays and controls. We assume that the appropriate analyses have been performed to identify the controller's information requirements, which are tied to context-specific objectives.

CHI design and evaluation in the ATC context must address three basic problems of information transfer (Chignell, Hancock, and Loewenthal, 1989):

- Communication of the controller's intent to the computer system and communication of the system's responses and their effects to the controller.

- Control of the computer system, focusing on the questions of what the controller does, what the computer system does, and when each acts.

- Accessibility and usefulness of the full range of system tools and capabilities.

This chapter considers not only the outward appearance of displays and controls (i.e., the "look" of the user interface), but also their dynamic interaction with the controller (i.e., the "feel" of the user interface). Three key areas of controller-computer interaction are discussed: data entry, command execution, and user guidance. Recommendations are provided for consideration by operations specialists (e.g., controllers) and system designers. Finally, future CHI technologies and issues are also discussed, since they are likely to be encountered in future NAS design efforts.

7.2 VISUAL DISPLAYS

"Display design is as concerned with task design as it is with presentation symbology and display devices."

Fadden, 1993, p. 243

Which visual display design techniques produce the most useful screen displays? It depends on the task, the user, and the environment. In ATC environments, the best display is the one that most consistently and accurately enables the controller to achieve the performance goals associated with the task being performed. This is why it is necessary to first understand task objectives and related information requirements before designing the display format or symbology.

Before a display is designed, a task analysis should be conducted to answer the following questions (Fadden, 1993):

- What is the objective of the task that must be performed with this display? What other tasks will be performed concurrently?
- What information is needed by the controller at each stage of performing the task?
- What errors is the controller likely to make in conducting this task? What are the consequences of these errors?
- What other displays will the controller be working with? The symbology, layout, controls, and other characteristics of the new display must be compatible with those of other displays in use.
- How critical is it for this task to be completed immediately?

Usability testing should be conducted periodically to identify CHI problems early in design and development. A brief overview of usability testing is provided in Section 7.2.5, with detailed guidance available in many sources (e.g., Carter, 1991; Dumas, 1988; Hix and Hartson, 1993;

Mogford, 1990). Just as software must be tested and de-bugged, the user interface must be tested for usability, operational suitability, and workforce acceptance (Harwood, in press). On-going field evaluations of changes to display systems can be highly useful in identifying deficiencies and redefining information requirements (Harwood and Sanford, in press).

> "Without adequate consideration of the user interface...the final design will be driven by software or hardware requirements with little consideration for the requirements of the user."
> Gilmore, Gertman, and Blackman 1989, p. 5

Visual displays are among the controller's key sources of information. To be effective, the design of visual displays for ATC must take into account what we know in several areas (ICAO, 1993):

- Human capabilities and limitations in visual perception, information processing, and comprehension (See Chapter 3, Visual Perception, and Chapter 5, Human Information and Processing.)

- The controller's functions and tasks (See, for example, Ammerman, Claussen, Inman, and Tobey, 1987).

- Environmental constraints (See Chapter 9, Workstation and Facility Design and Evaluation.)

- Hardware/software capabilities and limitations (See, for example, Foley, Van dam, Feiner, and Hughes, 1990).

The selection and presentation of information for displays must be managed so that the displays provide only the information that is needed (and only the information that is needed), when it is needed, and in a form that is useful and operationally suitable. Ergonomic standards and guidelines are available to aid designers in meeting this objective, but these issues must be resolved within the context of the controllers' individual and team tasks. Prototype visual displays must be evaluated for usability, operational suitability, and acceptability within the ATC context.

Visual displays may be technically usable but not operationally suitable or acceptable (Harwood, 1994). That is, their physical and perceptual aspects, such as formatting and use of color graphics, may be comply with guidelines, but their content and their sequential presentation of information to the controller may be deficient (Harwood and Sanford, 1994). If visual displays do not meet criteria for operational suitability, it is unlikely that they will be acceptable to the workforce. Field evaluations, developmental testing, OT&E, and site testing are designed to identify suitability and acceptability issues, as well as to assess usability.

TERMS, TECHNIQUES, AND DESIGN STRATEGIES

In this section, we introduce terminology that is used throughout the discussion of visual displays. We also stress the importance of learning from the workforce's experience with visual displays and building on the positive aspects of that experience. Additionally, we address several critical issues faced by designers and evaluators of visual displays: selection of a visual display technology, selection of visual presentation techniques, and choice of graphic versus textual displays.

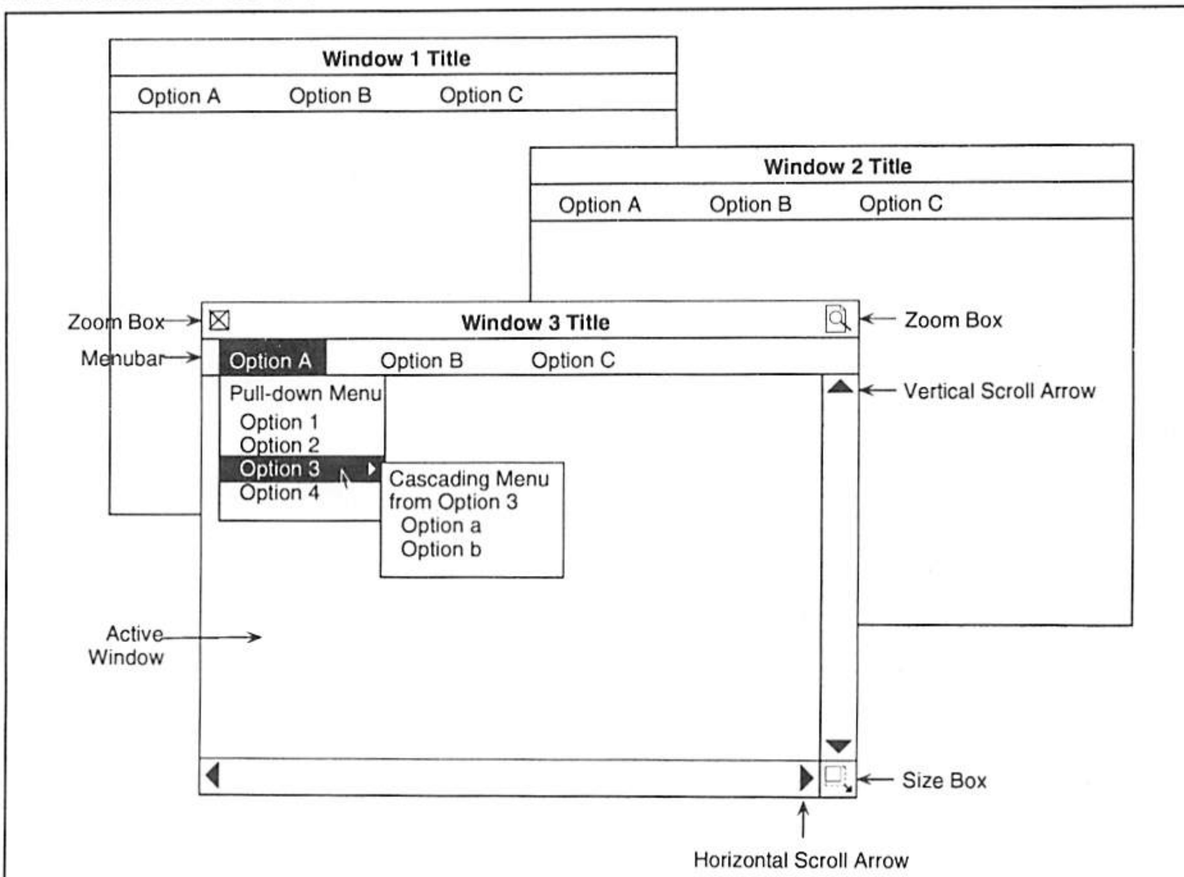
7.2.1 What elements are commonly found on visual displays?

The following elements are among those commonly found on visual displays:

- *Full page display*. If information is presented on full page displays, the user can access only one display at a time, and this display covers the entire screen area. Many of today's computer systems rely on **windows** instead of full page displays.

- *Window*. Typically, a window is smaller than the full screen size, and multiple windows can be displayed simultaneously. A window is so called because it provides a view (or a window) into a larger data set than fits within the window frame. The user can "move" the window to view different parts of the data set. Metaphorically speaking, windows are also like papers on a desk: they can sometimes be overlapped so that different parts of windows are visible. Like papers on a desk, any window can be brought to the forefront, as the active window. Figure 7-1 shows the screen format for a generic application, in which users are presented alternative choices in a window format. (For a full discussion of windows, see Billingsley, 1988).

Figure 7-1. Examples of Overlapping Windows, Pull-Down Menu, Cascading Menu, and Embedded Control Devices



- *Window controls.* Located around the window frame, window controls typically allow the user to open and close windows, to re-size them, to scroll through their contents, and to zoom in or out. The user activates these controls by using a cursor positioning device, such as a trackball or mouse.

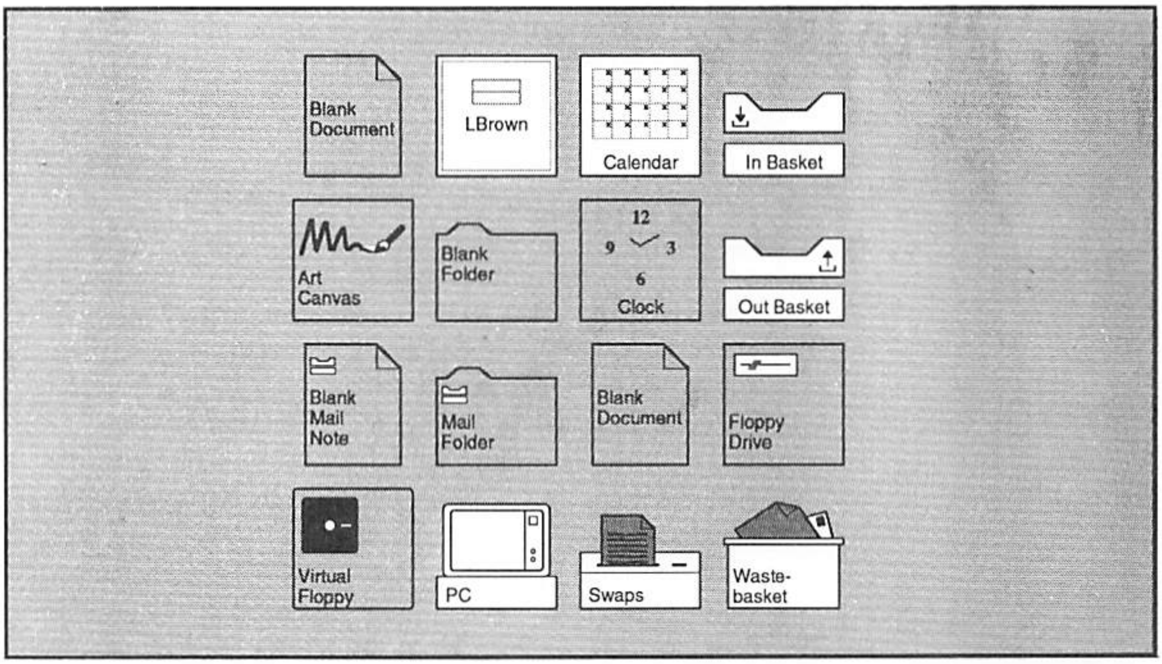
- *Symbol.* A symbol is a graphical object that may or may not allow user manipulation (e.g., move, copy, or delete); the user may be able to access detailed information on symbols. ATC displays use many symbols, such as the aircraft target symbol and the up/down arrows in the data block.

- *Menu.* Like a menu in a restaurant, a screen menu provides the user with a series of choices or options. Typically, user selection of a menu option executes a command, accesses data, or starts up an application (also called a "tool"), such as a word processing program. To preserve screen space and reduce visual clutter, many of today's computer systems use menus that are accessed on demand, rather than continually displayed. The user accesses a "pull down" window by selecting a menu title, and accesses a "pop up" menu by selecting a "hot spot" on the display (such as a graphical object that has associated with it numerous information windows which are listed in the pop up menu). Basic menu design is discussed in section 7.2.18. For thorough discussions of menu design, see K. L. Norman (1991b), and Paap and Roske-Hofstrand (1988).

- *Pushbutton.* A visually-displayed, software-generated pushbutton is designed to look like a physical pushbutton. It allows the user to execute commands associated with a display page or window.

- *Icon.* An icon is a graphical object that can be acted on, represent a command, or represent an application. As illustrated in Figure 7-2, icons are usually labeled because the meaning of an icon may not be self-evident. **If icons are used on ATC displays, the meaning of each icon should be simple, straightforward, and immediately apparent to the controller.** For example, the tower team has adopted a "rabbit" icon (Figure 7-3) to label the control for landing sequence lights. Ensuring that icons are clear requires review by representative controllers, at a minimum. Performance testing is recommended.

Figure 7-2. Some Standard Icons for a Word Processing Application
From *Human-Computer Interface Design Guidelines* by C. M. Brown, 1988, p. 11. Copyright 1988 by Ablex. Adapted by permission.



User interfaces that include windows, icons, menus, and pointing devices (e.g., a mouse or trackball) are sometimes called WIMPS (e.g., Chignell and Waterworth, 1991). No reflection is intended on the user. Rest assured, that when they are properly designed, real controllers can use WIMPS.

7.2.2 How can new displays build on the workforce's experience with the current system?

When a new system or system upgrade is being designed, it is worth reviewing how information has been represented in the past for displays supporting similar tasks. The idea is to maximize **positive transfer** of knowledge and skills from the old to the new system. For example, it would be beneficial to retain known acronyms and symbols, rather than inventing new ones unnecessarily. **Designers should not try to attach new meanings to familiar acronyms, terms, or symbols.** Building on controllers' experience can benefit the workforce through simplified and/or reduced training, particularly if familiar terminology and symbology are retained. Another advantage to this approach is that human performance expectations can be based to some extent on performance for the current display, thereby reducing the risks associated with introducing new displays.

One such risk is known as **negative transfer.** Negative transfer occurs when an individual's learned response to an event is wrongly applied to a similar event in a different context. For example, one display may indicate handoff status by highlighting the Aircraft Identification (ACID) in reverse video; a new display may use reverse video to indicate "Lifeguard" priority (i.e., top priority for clearance given to medivac aircraft and other life-and-death missions). Attempting to accept a "handoff" in the new display would be an incorrect response and an example of negative transfer. Similarly, errors would be expected if new meanings were assigned to familiar terms and acronyms. **New displays should minimize negative transfer from existing displays to new displays.**

The benefits associated with designing new displays to capitalize on controllers' experience with existing displays often make it less likely that alternate display symbology and formats will be considered in the design process. It is, therefore, important that the designer closely examine underlying tasks, because subtle differences for proposed tasks may have implications for displayed information content or formatting. In general, when considering alternative displays, designers should try to capitalize on positive transfer wherever possible through an awareness of current practice. At the same time, designers should not propagate poor human factors design or ignore important new task requirements.

7.2.3 What are the general characteristics of usable visual symbols and formats?

An ATC visual display exists to communicate information that will enable a controller to make decisions and/or take action. For example, a controller's radar scope communicates information (e.g., altitude, airspeed, aircraft type) that the controller needs to control air traffic. The designer tasked with displaying information must make decisions regarding **symbology selection** and **display formatting.** For example, how should climbing/descending aircraft be represented in the data block? How should fields in the data block be arranged?

Symbology selection determines how specific information elements will be represented within the display. **Symbols should be intuitive so that they are easily recognized and understood.**

Symbols should also be used consistently across a display set (i.e., all displays that will be used by a single controller). Otherwise, errors could be induced when the same symbol has one meaning on one display and a different meaning on another display.

Since it may be possible to represent the same element using alphanumerics or graphics, the designer must decide on the most effective symbol set. Although graphic displays are effective in conveying complex information quickly, they also use up much more space and may contribute to display clutter. **Controllers need to advise system designers on what various symbols already mean to controllers and whether other information is worth displaying graphically.** Often, a combination of alphanumerics and graphics gives the best result in terms of user performance.

Display format determines the framework within which the information will be presented (e.g., the organization and location of data). Format is influenced by the top-level tasks the display supports, while symbology selection is often guided by specific requirements of the detailed tasks. Once designers determine information requirements and general display organization, they can refer to hundreds of well-established, very specific human factors guidelines in developing the initial visual displays. These guidelines can be translated into an ATC system **style guide**, a set of specific design rules for developers to follow. CHI design guidelines are presented throughout this chapter.

Symbology and format should be consistent with the concept chosen to unify the display elements. This concept should be explicitly represented in the **conceptual model**, which conveys a simplified, coherent sense of the computer system's functionality to the user (Norman, 1986). Without such an understanding of what the computer does and how it works, users will form unfounded expectations of the computer's capabilities and inaccurate explanations for system behavior.

7.2.4 What evaluation aids can be used to assess symbology, format, and other CHI issues?

In the near future, user interface development environments will include tools for checking candidate display designs for consistency with design specifications and style guides. (See, for example, Fisher, Nakakoji, Ostwald, Stahl, and Sumner, 1993; Jiang, Murphy, Bailin, and Truszkowski, 1993.) Such automated consistency checks can go a long way toward catching and correcting potential usability problems.

Rapid prototyping and usability testing can assist designers in evaluating symbology and format to organize display elements into a unified whole. In addition, frequent consultation with controllers helps ensure the effectiveness of the design in communicating displayed information. Rapid prototyping tools enable designers to quickly transfer ideas to the computer. These software tools help designers generate screen panels that look and act like ATC displays, but information and commands entered are not really processed. Cooperative and on-going refinement of symbology and formatting through rapid prototyping, usability testing, and active controller participation becomes especially important as display complexity increases.

During usability tests, controllers try out user interface features; performance measures and subjective feedback are collected. Performance measures may include error rate and speed of response. Subjective measures may include workload ratings and questionnaires. By assessing the limitations and enhancements of alternative implementation techniques, the designer can objectively evaluate the task performance issues. Of course, relative cost and other tradeoffs will need to be considered. (See Chapter 10, Human Factors Testing and Evaluation, for a detailed discussion of test methods and performance measures.)

7.2.5 When are graphic displays preferable to textual displays?

Graphics are used to represent relationships in physical space or relationships over time (Smith and Mosier, 1986). **Graphic displays are invaluable when the information presented is pictorial in nature, such as the ATC control environment containing aircraft traffic over a geographic sector.** Situation displays are important ATC graphic displays that combine geographic/map data with overlying event data (e.g., aircraft location). Related sets of data can be compared quickly using graphics (e.g., two data sets plotted on the same axis), such as runway use by month for this year versus last year.

Monitoring changing data is also easier using graphic displays. Graphics are commonly used today to create what is called a *direct manipulation* visual display (e.g., Shneiderman, 1983b). As an example of possible direct-manipulation applications in ATC, some communications between the pilot and the controller using Mode-S will be displayed on the controller's visual displays at predetermined or controller-manipulated positions. After receipt of a message or instruction, the controller may want to acknowledge or delete the information by using the mouse cursor to click on the message or an icon for the message and drag it to the "waste basket" icon. Using this method, displays are not crowded or cluttered, and the information is deleted at the controller's option, not at the system's option.

In another instance, the controller may want to see a graphic depiction of the performance characteristics of an aircraft climb rate. The controller could display this data for information and then delete it from the display area by clicking on the data chart itself. Similarly, a supervisor may wish to see a graphical display of the current runway acceptance rate and an anticipated acceptance rate for construction on a taxiway as well as other related information. The information could be displayed on a supervisor monitor graphically, textually, or using both; it could be deleted with the mouse button; or it could be manipulated. For example, controllers could play out "what if" questions as planning aids. (What if we closed taxiway

Mike instead of Alpha? What is the delay and runway acceptance rate?)

Although users generally prefer graphics, performance is not always better with graphics. The effects of graphics on performance appears to be task-dependent. That is, some, but not all tasks, lend themselves to a textual rather than graphic presentation and vice versa. Performance testing is recommended to provide a basis for choosing text or graphics.

SELECTED HARDWARE ISSUES AND RECOMMENDATIONS

In this section, we discuss several physical characteristics of visual displays and physical features of the surrounding environment that can have negative effects on human performance. We address factors that influence the use of visual displays under low-lighting conditions. We also provide recommendations for reducing the negative effects of such factors as reflections and glare.

7.2.6 Why might one observer see a display flicker while it looks steady to another observer? Why might a display appear to flicker in peripheral vision but appear steady when viewed in one's direct line of sight?

Visual displays are actually flickering from bright to dim all the time at a constant rate. What causes the *perception* of display flicker is one's *threshold* to flicker as compared to the display's flicker rate. Under most conditions, the human threshold for perceiving flicker is under 65 cycles per second. This means that, under these conditions, **if the refresh rate of a display is at least 65 cycles per second (or Hz), the display will always appear to be steady.** (A few sources (e.g., NASA STD-3000/Vol.III) recommend a refresh rate of at least 100 Hz. for dark characters on a light background.) If the refresh rate of the display is below the observer's threshold for flicker, then display flicker will be noticeable.

Several factors can change an individual's flicker threshold. Two considerations are (1) whether the controller is using the display in a brightly lit or dimly lit room; and (2) from what angle the controller is viewing the display. In general, as the overall illumination level increases, display flicker becomes more noticeable. Further, a display viewed out of the corner of your eye may appear to flicker even though, when viewed head-on, it appears stable. This is caused, in part, by where the image falls on the retina. As noted in Chapter 3, Visual Perception, the receptors that are most prominent in the periphery of the retina are also more sensitive than are the centrally-located receptors. As a result, images viewed out of the corner of your eye may appear to flicker, even though images viewed head-on may appear stable.

Age-related factors are also associated with one's threshold to display flicker. Since the amount of light transmitted through the lens of the eye decreases with age, the apparent brightness of the image as well as the overall illumination level will be lower for older individuals. Younger individuals will, therefore, be more sensitive to flicker. The differences can be pronounced enough to cause a display to appear to flicker for one observer while an older observer would see a steady image.

7.2.7 Why might there be an apparent change in the position of a displayed object when it is viewed from different angles?

If there are two objects directly behind each other, viewing them from an angle will distort the true geometrical relationship between them. This is known as **parallax.** This can be illustrated by automobile gauges, which are arranged to accommodate the viewing angles of the driver. For example, the fuel gauge commonly uses a needle with a scale behind it. A gauge which reads 1/4 tank remaining to the driver may appear to read empty to the passenger sitting to the right of the driver, due to parallax. The distance between the needle and the scale of the gauge can influence the amount of parallax induced. A greater distance between the needle and the scale will result in more parallax and, therefore, more distortion when the gauge is viewed off-angle.

Parallax may also be noticeable on large screen CRT displays due to the wide viewing angles. In addition, the curvature of the display screen will induce even more parallax at the edges of the display. For example, on a radar display format, tracks that appear at the edges or very bottom of the screen may have alphanumerics that appear slightly fuzzy. This is due in part to parallax. When these tracks reach the center of the screen, the text information in the data blocks should appear much more readable. To prevent distortion due to parallax, critical information should be displayed in the center of the screen, not along the edges. The position and form of displayed objects should appear the same to the controller seated directly in front of the display and to the team members viewing the display from other reasonable angles.

7.2.8 What are the special considerations for visual displays that will be used in dimly lit or sunlit work environments?

There are several considerations concerning the use of displays in dimly lit rooms. First, it is important to note the difference between the **luminance** of the sources of light and the apparent brightness of the light sources. Luminance refers strictly to the physical properties of light, which include the wavelength and the amount of energy transmitted. **Brightness** refers to the appearance of a display or indicator, which can be influenced by factors other than luminance, such as the surrounding illumination in the room. To preserve dark adaptation (so that the user's eyes remain accustomed to, and can see well in, the dark), **displays should be designed to be used at the dimmest setting possible while still maintaining good image quality.**

Other considerations for using displays in dim rooms are the source of lumination and the range on the display brightness control. For dimly lit rooms (such as TRACONs), displays cannot depend on ambient light (i.e., they must emit light). For brightness control of a display used in a dim room, the range of brightness values needed is much narrower than it is for normally lit rooms. (Lighting is discussed further, as a feature of the ATC facility environment, in Chapter 9, Workstation and Facility Design and Evaluation.)

There are three main considerations for displays to be used in sunlit areas (i.e., ATC towers): glare, contrast, and use of color.

Glare can be reduced by putting a filter on the CRT screen and/or a shield (or hood) around the display. Such shields may, however, interfere with viewing the display from all necessary angles (e.g., sitting and standing). Before a decision is made to use a filter or hood, the costs and benefits should be evaluated. Measures taken to reduce one kind of glare may increase glare from other sources. There is no simple answer to eliminating glare. (Glare and how to reduce it are also discussed in Sections 9.6.1 and 9.6.2 of Chapter 9, Workstation and Facility Design and Evaluation.)

The contrast of displays to be used in direct sunlight will have to be higher than those used in more dimly-lit environments. Similarly, any colors used in a display need to be carefully selected, since sunlight washes out colors. This means that colors that would normally appear quite vivid appear very light (i.e., desaturated) when viewed under direct sunlight. (See Chapter 3, Visual Perception, Section 2.2.3, for a fuller discussion of the effect of direct sunlight on color displays.)

SELECTED SOFTWARE ISSUES AND RECOMMENDATIONS

In this section, we consider key issues in the design and evaluation of visual displays. These include selecting an effective symbol set, sizing text and symbols, and establishing consistency in the use of typographic features. A major display issue is the use or overuse of color on ATC displays. If windowing capabilities are provided, the use of windows and window management techniques are additional central issues whose resolution can positively or negatively impact controller performance. In this section, we also provide recommendations on locating visual alerts and assuring the controller's access to required information.

7.2.9 What should symbols and text look like?

Effective information transfer is critical to the success of any new ATC system. An effective symbol set is one that results in accurate and timely transfer of information from the display to the air traffic controller. An effective graphical symbol set has the following characteristics (Shneiderman, 1992; Smith and Mosier, 1986):

- Symbols are pictorial.
- Symbols and alphanumerics are legible.
- Symbols can be easily discriminated from each other.
- Symbols are simple and intuitive.
- Symbols are displayed in a logical and organized format.

Pictorial displays. Pictorial displays and display formats are already in use on ATC displays. The use of lines and segments to denote sector boundaries and airways provides a concise and easily interpretable representation of the airspace. Up and down arrows are used effectively to indicate whether aircraft are climbing or descending. Many other pictorial symbols have specific meanings for controllers in the current system.

Legible symbols and alphanumerics. Symbols and characters must be easily readable. Symbols should be simple and intuitive. Similarly, alphanumerics should be in a simple (sans serif) font. Other factors that affect readability (size, resolution and contrast) will be discussed in detail in later sections. (Also see Appendix A, Human Factors Considerations for Color Displays for ATC.) Symbols or alphanumerics that are not readable can lead to visual fatigue, controller errors, and delays in the transfer of information. One solution is to make symbology size and contrast adjustable by the controller.

Discrimination between symbols. An ATC display format can become very cluttered with graphical and alphanumeric symbology. Therefore, it is critical that display clutter be minimized and that symbols be easily discriminated from one another.

Simple and intuitive symbology. There are no hard guidelines for the development of intuitive symbology. One recommendation is to exaggerate the most salient, defining, or unique features of a symbol (Avery and Bowser, 1992). An example of intuitive symbology is the use of an aircraft symbol to represent a radar track. Complicated and very detailed symbols on a display may be hard to interpret in a timely manner, however, and may detract from information transfer. Task analysis is used to determine candidate symbol sets, which must be tested with a representative sample of controllers. Easily understandable symbology can lead to reduced training time and effective information transfer.

Logical and organized display format. The arrangement of symbols and alphanumerics plays a critical role in how well information is transferred from the display to the controller. On a radar screen, the placement of aircraft tracks indicates their location in the airspace. In addition, the data block is linked by a leader line to the aircraft track that the information pertains to. Data blocks can be rotated and adjusted within 360 degrees. They can be positioned close to or at a distance from the target, based on the controller's workload ("scope clutter") and personal preferences for data position. This format is much more interpretable than a tabular format listing aircraft positions, types, speeds, and so forth.

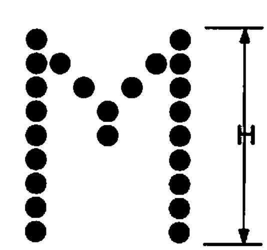
Additionally, some controllers use the placement of the data block to organize their traffic. For example, one controller observed in the field likes to organize his display so that all aircraft with the data block to the left of the target are flying west, while all aircraft with the data block on the right are flying east. Designers need to be aware of user categories like this that have developed over time, but that goes beyond the original design intent. Controllers can be very helpful in identifying display features that they use for other purposes. When there are alternative choices for the design of a symbol or a set of symbols, an evaluation of the symbols effectiveness on controller performance can help provide information in making a choice between symbols.

7.2.10 What size should text and symbols be?

Normally, we measure the size of objects in terms of yards, feet, or inches. However, the measurement of displayed information not only has to take into account the physical size of a character, but also the distance the character or symbol is from the eye. Viewing distances of displayed information can vary considerably. A television show can be comfortably viewed from as far away as 10 to 15 feet, but imagine trying to comfortably view a computer screen from that distance! To account for both the size and distance of symbols and characters, the angle that the symbol or character forms with the eye is used as the unit of measurement. This is referred to as the **visual angle.** (Visual angle and how to measure it are described in Chapter 3, Visual Perception.) Visual angles are specified in terms of minutes of arc or degrees (1 degree = 60 minutes of arc).

Character size is measured by the height of the character in terms of its visual angle. Figure 7-3 illustrates character height as the full vertical distance between the top and bottom picture elements (pixels) of a plain capital letter. Legible characters should have character heights of at least 16 minutes of arc with a preferred height of 20 to 22 minutes of arc (ANSI, 1988). At a viewing distance of 18 inches, this translates into values of approximately .08, .11, and, .12 inches, respectively. An easy formula is that the minimum height of the characters should be 1/200th of the viewing distance (Avery and Bowser, 1992). This translates to a visual angle of 17 minutes of arc (.286 degrees). It is important to note, however, that these guidelines apply only to monochromatic displays. The minimum size for characters presented in color will depend upon the colors of the text and background, and other colors used in the display. The exact size will depend upon the viewing distance and the actual task at hand.

Figure 7-3. Character Height
From *American National Standard for Human Factors Engineering of Visual Display Terminal Workstations (ANSI/HFS 100-1988)*, p. 27. Copyright 1988 by the Human Factors Society. Reprinted by permission.



If a text message needs to attract attention efficiently, then the visual angle subtended by the characters will need to be greater than would normally be required. This means that either the characters will have to be larger, or the display will have to be closer to the viewer. Other considerations are the computer monitor resolution, foreground to background contrast, and character font. Commercially successful products having high resolution, black on white characters, and a sans serif (i.e., nonornamental) font use character heights of .12, .09, .06, and .04 inches, depending on the task. **In ATC environments, symbol size should be adjustable by the controller.**

Some tasks may call for the discrimination between different colored text or between different colored symbols. **When color discrimination is important, symbols and characters need to be bigger than when they are displayed monochromatically.** This is because receptors vary with respect to their ability to detect color. For this reason, **larger than minimum character heights are recommended. Also, small blue symbols or characters are to be avoided when the task requires color discrimination.** (Color vision is discussed in Chapter 3, Visual Perception.)

7.2.11 What are the key issues in visual alerting and coding?

There are many factors that affect the effectiveness of visual alerts and coding techniques. Some of these factors are related to the physical parameters (e.g., brightness, location) of an alert and others are related to its functional parameters (i.e., does it perform as it was intended to perform). The goal of visual alerts and coding techniques are to effectively direct the controller's attention to important information that might otherwise be missed or critical situations that might otherwise take longer to recognize. Both the physical and functional parameters of alerts should be evaluated in order to ensure that the alert is effective and useful.

It is important that the alert be useful to controllers, and not be a nag or annoyance. It must provide the controller with useful information in a timely manner. If the controllers consider an alert to be annoying, then the alert needs improvement. The problems with the alert may be related to its physical parameters (e.g., the location, blink rate, or contrast of the alert) or to its functional parameters (e.g., it doesn't activate when it should or goes off when it shouldn't). In either case, the problem will have implications for human performance. For example, any alert that generates a high percentage of false alarms will be considered annoying and eventually, controllers may begin to ignore it. Similarly, an alert that continues to be activated after it is no longer needed is unnecessarily distracting.

False alarms are a potential problem with any alerting system. A high false alarm rate can lead to lack of confidence in the system and a slower response time to the alerts. Imagine a hotel fire alarm that is activated three times in fifteen minutes. After each alarm is an announcement that the alarm is false. An hour later the alarm is activated again, only this time there is no announcement. Would the guests of the hotel leave the building as quickly as they would if the false alarms had not occurred earlier?
Similarly, if a conflict alert function is repeatedly activated in situations in which the aircraft were obviously separated, controllers would begin to loose confidence in the system, and possibly not respond as quickly as if they believed that every alert was valid.

The following coding and emphasis techniques can be used to attract attention to specific information or to categorize information on a display:

- Blinking or flashing
- Reverse video
- Use of different sizes (of text or symbols)
- Color
- Location

All coding techniques have advantages and disadvantages. For example, flash coding is good for attracting attention, but flashing text is difficult to read. The critical importance of specific information can be indicated by a separate blinking symbol placed near the information that must be read (Gilmore, Gertman, and Blackman, 1989). The controller should always be able to turn off blinking or flashing. **Blink rate should be between 0.1 to 5 Hertz (Hz), with 2 to 3 Hz preferred. No more than two levels of blinking should be used because more will be difficult for controllers to discriminate. Flash coding or blinking should be used sparingly and never for text or numbers that are critical or that the controller needs to read quickly.**

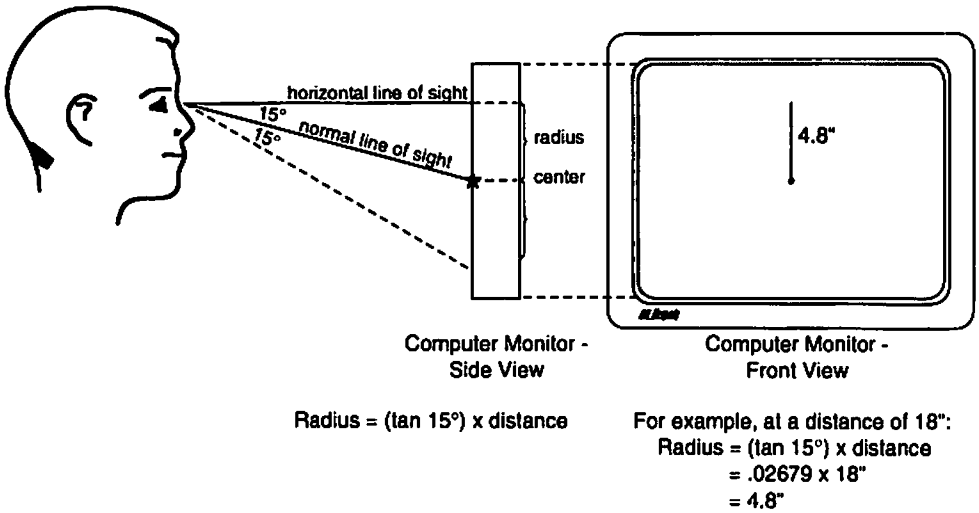
Critical, abnormal, or updated data should be emphasized or highlighted using such techniques as reverse video, shadows, brightness, or color coding (Marcus, 1992; Smith and Mosier, 1986). In each case, the levels of shadows or brightness used (as well as different colors) need to be distinctive and easily identified in order to be effective highlighting tools. Specific guidance on the use of emphasis and highlighting techniques can be found in many sources on user-interface design (e.g., Avery and Bowser, 1992; Brown, 1988; Brown and Cunningham, 1989; Carlow, 1992; Galitz, 1985; Gilmore, Gertman, and Blackman, 1989; Hix and Hartson, 1993; Horton, 1990, 1991; Mayhew, 1992; Shneiderman, 1992; Smith and Mosier, 1986).

Overuse of a coding technique or the combined use of too many techniques can result in unacceptably "busy" visual designs. **If size coding is used, ICAO (1993) recommends using only two widely different sizes.** Emphasis should be placed on simplicity and clarity.

Color coding is especially useful for highlighting or categorizing information on a display. **Highlighting must be used sparingly so that emphasized items will actually "pop out" from the others.** If too many items are emphasized, none will pop out. Use of color is discussed in detail in the next three sections of this chapter.

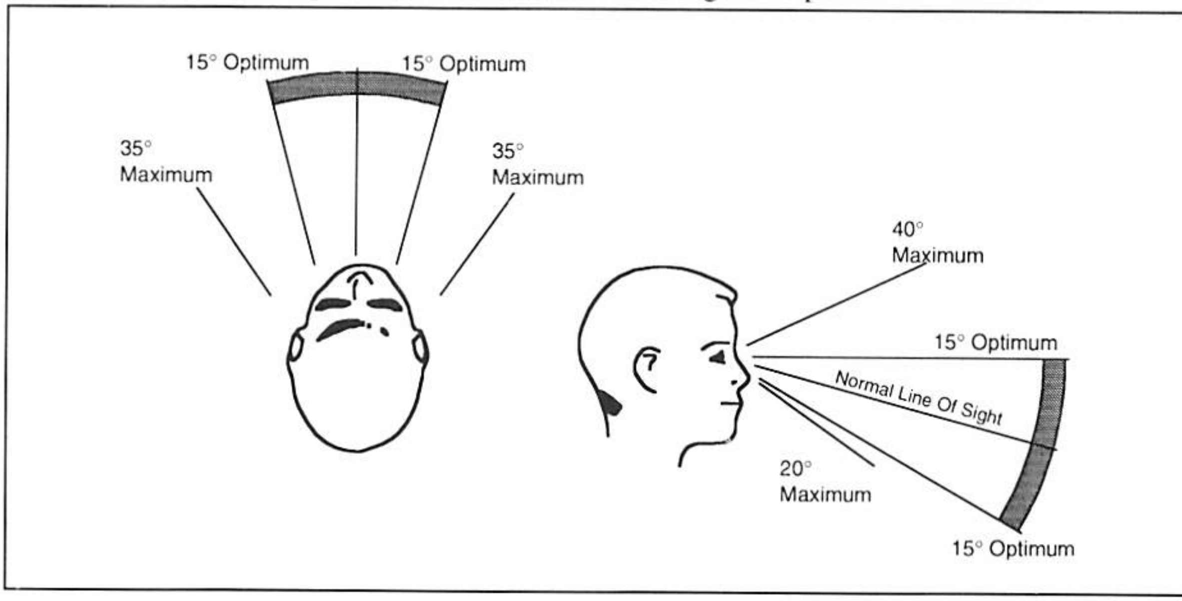
The location of information on a display is a function of both the physical area of the display surface and the position of the eyes relative to the display screen. Visual alerts will be recognized more quickly when placed in the area of the visual field that has the best visual acuity than when placed in the periphery. This optimal area can be thought of as a cone extending from the **normal line of sight** with a radius of 15 degrees. (The normal line of sight is the line 15 degrees below the line extending horizontally from the center of the pupil.) This "cone" corresponds to the portion of the eyes with the highest density of receptor cells that provide the best visual acuity and detailed vision. The density of the cells decreases towards the periphery of the eye and, as a result, visual acuity decreases as well. Signals located outside this 15 degree cone will not be detected as quickly. Figure 7-4 illustrates the determination of the circular area on the display for high priority signals.

Figure 7-4. Recommended Placement of Visual Alert Signals
Adapted from *Aircraft Alerting Systems Standardization Study (DOT/FAA/RD-81/38)* by B. L. Berson, D. A. Po-Chedley, G. P. Boucek, D. C. Hanson, M. F. Leffler, and R. L. Wasson. 1981, Volume II, Aircraft Alerting System Design Guidelines, p. 40.



Factors such as whether the controller is sitting or standing will influence the visual comfort zone. In addition, the height of the display support surface can change the eye reference point. Figure 7-5 shows the visual field with eye movements. With head movements as well, the visual field is greatly increased (to about ± 90 degrees in the horizontal and visual plane); however, visual displays should be designed to minimize the need for head movement.

Figure 7-5. Vertical and Horizontal Visual Field with Eye Rotation
From *Human Engineering Design Criteria for Military Systems, Equipment, and Facilities* (MIL-STD-1472D), Department of Defense, 1989, Figure 2, p. 31.



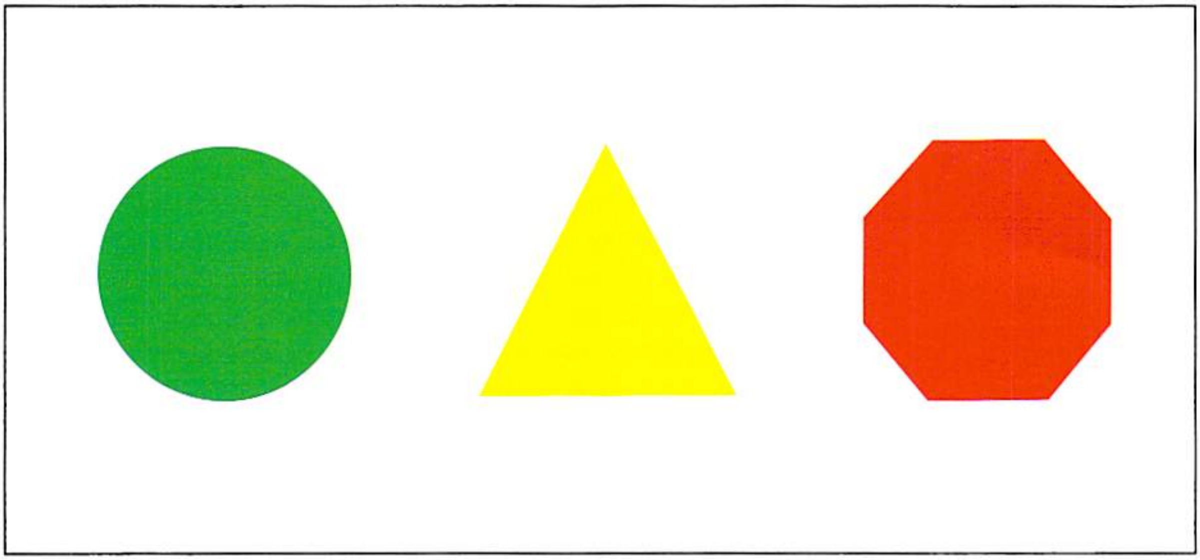
7.2.12 What are appropriate uses for color in ATC displays?

Color was introduced to CRTs in an effort to declutter displays; it was also presumed to simplify information coding problems and limitations associated with monochromatic displays. A color-coded display allows for quicker searches of information than are possible when the display is monochromatic. In fact, color codes are more effective for search tasks than are non-color codes (e.g., shapes, letters, digits). Color also "catches" attention and can reduce error, if used appropriately (Narborough-Hall, 1985). It is important that any color coding used on displays is applied consistently throughout the display set. That is, **the same color coding strategy needs to be applied to every display used by the same controller**. For example, if a flashing red data block or red button has a certain meaning on one display, it should carry the same meaning on other displays used by the same controller. Such consistent coding will help to minimize negative transfer errors.

Another advantage to using color derives from the stereotypical meanings associated with various colors by different cultures. For example, in Western cultures **red is widely associated with warnings/danger, while amber typically indicates caution, and green indicates normal status**. Other colors have more diverse cultural meanings and are more suited to general grouping rather than detailed operational information coding.

Within limits, the use of color permits higher information density without the usual decrements associated with clutter. Use of **redundant coding** (e.g., shape, position, color) can improve controller confidence in the display and help maintain performance under marginal operating conditions. In addition, redundant coding is essential to support the performance of color-deficient users. Figure 7-6 illustrates that concept of redundant coding. **It is strongly recommended that visual displays first be designed to meet human performance criteria under monochrome conditions, with color being added "only if it will help the user in performing a task"** (Gilmore, Gertman, and Blackman, 1989, p. 108). The display should be usable in monochrome because users differ in their perception of color.

Figure 7-6. Examples of Redundant Coding.
(Original figure)



Color should be used only for functional, not decorative, purposes on ATC displays. Although people generally like color, if it is used inappropriately on a display, it can make the task more difficult, instead of easier. Any proposed use of color must be thoroughly tested for its suitability to the specific task and the operational environment. For example, one important consideration for the use of color displays in the tower environment is the effect of direct sunlight on the display. This has the effect of lightening, or "washing out," the colors and makes it impossible to differentiate between some colors. (The effect of direct sunlight on a display is discussed in Section 3.2.3 of Chapter 3, Visual Perception.) Designers and evaluators need to carefully consider all of the ramifications of using color on ATC displays (see Narbrough-Hall, 1985).

For additional information on the use of color on ATC displays, see Appendix A, *Human Factors Considerations for Color Displays for ATC.*

7.2.13 How many colors should be used on a display?

Significant individual differences exist between people's ability to recognize color. In tests conducted by Boeing, pilots were unable to discriminate more than six colors on a display viewed under anticipated viewing conditions (Fadden, 1993). Carter and Cahill (1979) and other researchers found that advantages to color-coding decreased as more than five to six colors were employed. This finding is consistent with research on human short-term memory characteristics, which indicates that memory-based codes limited to five to six dimensions present the least difficulty to humans (Silverstein, 1987; Teichner, 1979).

Recommendations about the optimum number of colors that can be easily and reliably identified on a display vary from three to ten colors (Murch and Huber, 1982; Teichner, 1979; Ericsson and Faivre, 1988). Use of more than six or seven colors, however, can lead to errors in color discrimination and task performance. If the purpose of color is only to reduce clutter (as on a map display) and the task does not require absolute identification, the number of colors can increase to more than seven.

7.2.14 Which colors should be used or not used on ATC visual displays?

The selection and use of color should be firmly grounded in what we know about color vision. The designer needs to be aware of color effects, such as simultaneous and successive contrast, which can effect the controller's perception of color.

When colors are used, their meanings should be consistent with common associations, such as green for *good* or *go*. The following recommendations are made by ICAO (1993) on color usage for ATC displays:

- In general, use pastel and desaturated colors. (An obvious exception to this is in tower displays viewed in direct sunlight, since sunlight washes out colors and desaturated colors would appear white.)

- Use saturated colors only for critical and temporary information because they can be visually disruptive.

- Do not use saturated colors for small visual objects or areas.

- Do not use saturated colors, especially blue, that can induce problems such as false impressions of depth (see section 3.1.8 on chromostereopsis). Saturated red and blue should not be presented next to each other for this reason.

- Ensure that all colors meet the brightness contrast requirements (8:1).

- Use colors that are clearly different from each other to avoid confusion.

- Use colors that have obvious names familiar to all controllers.

- Use colors that allow for permissible deficiencies in color vision.

Since peripheral vision is very poor at discriminating colors, the use of color should be reserved for portions of visual displays that will normally be in the controller's

direct line of sight. Color coding is not recommended for peripheral displays.

The number of colors that can be reliably discriminated also depends on the amount of illumination in the room. As sunlight or bright light is added to the display, the number of colors that can be discriminated is sharply decreased. This would appear to the controller as "washing out" of the colors on the display. For "absolute identification," the controller should view and identify one color at a time. This is especially important in brightly lit environments, such as the tower. For more detailed information on color selection and use, see Appendix A, Human Factors Considerations for Color Displays for ATC.

7.2.15 What human factors principles apply to the typographical aspects of visual displays?

The wording of displayed text, such as labels, menu options, commands, and error/feedback messages, should be either from common English usage or ATC terminology. Wording should mimic the previous ATC system to promote positive transfer, except where such wording would propagate poor human factors design or is no longer applicable. Labels should describe the associated data content, not what to do with the data (e.g., "Metering Advisory List" rather than "Select one from the list"). Wording should be used consistently: the same term should have the same meaning within and across displays, and a single concept should always be described by the same term (Smith and Mosier, 1986).

Abbreviations should be used sparingly. Abbreviations should be reserved for terms having an abbreviation that is more common than the full term or where there is insufficient room to display the entire term. It is permissible to use the full term whenever possible and the abbreviated term otherwise. Only one abbreviation should be used for each word. Abbreviations that are familiar to the controller should be used; all other abbreviations should follow one simple abbreviation rule, such as the first three to five letters of a word. An abbreviation and acronym listing should be available to the controller, preferably within an on-line computer display page or window (Smith and Mosier, 1986).

"READING IN ALL CAPITAL LETTERS CAN TAKE LONGER BECAUSE WORDS LOSE THEIR CHARACTERISTIC SHAPES. ALL WORDS BECOME RECTANGULAR."

Brown, 1998, p. 26

Punctuation should be used conservatively. Punctuation should be used only within complete sentences, when needed for clarity, or to partition long data items (Smith and Mosier, 1986), e.g., for time "12:15," phone number "(603) 555-1212," or Social Security Number "017-60-8553." Abbreviations and acronyms should not include periods.

Standard English format and punctuation should be used for any text in sentence form (Smith and Mosier, 1986) (e.g., error/feedback messages, text messages from other controllers or pilots, lengthy text information/instruction from the computer). **Upper case should be reserved for the first letter in a sentence or a typically capitalized word, and acronyms should be upper case.** Text presented in all upper case is more difficult to read as compared to mixed case.

Upper case can be used for short items to draw the controller's attention to important text, such as for field labels or a window title. All upper-case text is sometimes used to compensate for small or illegible text: this is probably not the best solution if other options are available (e.g., increase the character size, replace the monitor with a higher resolution monitor, or increase the background-to-character contrast).

When controllers must read a lengthy amount of text, they should be given printouts. Reading from the computer screen can be 20 to 30 percent slower than reading from a hard copy. Of course, this benefit should not be nullified by the time required to print and/or to access the printout. Printers should have the capability to print in upper and lower case. Because reading print in all capital letters is more difficult and time consuming, rejecting a printer that will print only in upper case is entirely justifiable on human factors grounds.

Commands should be written in the active voice (rather than passive voice). The active voice identifies the person or entity that performs the action. Passive voice conceals that identify. For example, the following sentence, in the active voice, identifies the controller as the performer of the action:

The controller entered the aircraft's identification (ID) code. In contrast, the following sentence, in the passive voice, does not identify the performer of the action: The aircraft's ID was entered. Notice that passive voice leave the performer's identity open to question. The aircraft ID may have been entered by the controller, the supervisor, or the computer! A direct message such as, "Enter active runways," addressed to the user, is preferable to a passive message, such as "Active runways should be entered."

Commands should be written in the affirmative (i.e., tell the user what to do rather than what not to do). Affirmative commands are easier and faster to understand than those using negatives. For example, "Clear the screen by pressing <Delete>." is preferable to "The screen will not be cleared until <Delete> is pressed." Negatives should be avoided to minimize the possibility of errors and to reduce the time required to understand them. Double negatives are especially difficult to grasp quickly.

Sentence phrases should follow the order in which events will take place. For example, "Select an aircraft before accessing aircraft-specific data" is better than "Before accessing aircraft-specific data, select an aircraft." **Error messages should be direct and precise, without being cryptic, cute, or insulting** (Shneiderman 1992). **Error messages should give the user a clear instruction in what to do next to recover from the error.**

An effective method for evaluating wording, grammar, punctuation, letter case, sentence structure, and other typographical concerns is to compare the above human factors guidance to hard copies of all text displays, the command set, and an error/feedback message listing. Automated evaluation tools include spell checkers and grammar checkers. Emerging technologies will permit automated evaluation of typography and other CHI elements.

7.2.16 How should items be listed on a display?

When data display items are in a list, the list should follow an immediately apparent logical order (Smith and Mosier, 1986). For example, a menu set and options within menus should be ordered according to the most typically used sequence. Task analysis is used to determine the sequence of ATC tasks. The sequencing of menu options by task will assist controllers in learning the system and maintaining orientation within the ATC process. Where there is no logical order for a menu or data set (e.g., a list of sector names or aircraft types), items should be ordered alphabetically.

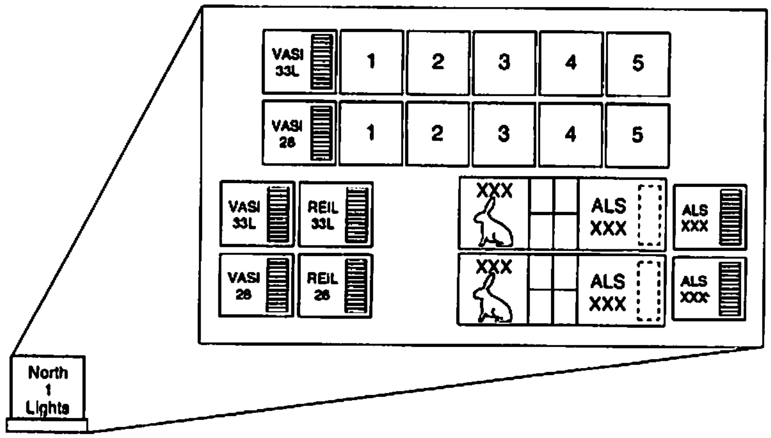
7.2.17 How should windows be implemented?

Windows allow the user to look into a portion of a data set that is too large to fit within the window frame. Typically, windows are rectangular. A window can cover part or all of the displayed area. Windows may or may not overlap each other. When windows are not permitted to overlap, they are referred to as **tiled** windows.

The user should be able to scroll the data set behind the window to make other portions of it visible within the window. A user should be able to perform the following **window management** tasks: moving windows, resizing windows, shrinking display pages into icons, and opening and closing windows (Billingsly, 1988; OSF, 1991). Figure 7-7 illustrates a proposed TCCC airport lighting display concept being reduced to an icon ("iconified").

The user should be able to enter data or execute commands within the **active window**, which is usually highlighted in some way, for example, by shading or color coding the frame (i.e., border) of the window. In most software applications, the user indicates the active window by placing the cursor in it and executing some control action, such as pressing a mouse button. When windows are allowed to overlap, the active window typically comes to the immediate foreground and is not overlapped by other windows. It appears to be in **front** of all other windows.

Figure 7-7. A Proposed TCCC Airport Lighting Display Concept Being Reduced to an Icon ("Iconified")
From TCCC Status Briefing, 1993. CTA INCORPORATED.



Designers must assume that controllers will not be familiar with window usage. The number of windows that need to be displayed on the screen simultaneously should be kept to a minimum. The user manual, training programs, and the user-system interface should all teach and promote the windowing metaphors (i.e., windows are like physical windows; data within windows are like a large paper scroll; and the window set is like papers on a desk top). **If windows are to be placed on a background that the controller is using for operational purposes (e.g., a situation display), they cannot block critical information, and their size must be kept to a minimum.**

Windows should not be used simply to increase the display's information content. Instead, the goal in using windows should be to increase the controller's **visual scope** (Norman, Weldon, and Shneiderman, 1986). Visual scope is the user's ability to integrate information contained in multiple windows into a concept or a whole. Several types of window layouts can increase visual scope, as follows:

- *Information integration layout.* Each window contains a different data set needed (in combination) to make a decision.

- *Levels of processing layout.* Each window displays information at a higher level of analysis.

- *Zoom-in/Zoom-out layout.* One window is a blow-up or gives the details of a portion of another window, which is especially important when presenting geographic data.

- *Perspective layout.* Each window shows a different angle, different attributes, a historical or future view, or some other aspect of an object or event.

- *Selective attention layout.* The controller has a primary window (or window set) but occasionally monitors a secondary window for information updates or system problems. When possible, the computer should take over monitoring the secondary window and inform the controller when s/he needs to divert attention to it.

- *Trigger change layout.* Windows are linked in such a way that if a change is made in one window, that change and calculations based on it are activated in the rest of the window set.

A key recommendation for any type of windowing layout is to clearly indicate the relationship between windows, thereby enabling the controller to maintain orientation. Orientation within visual displays (also called **visual momentum**) is the controller's ability to see relationships between windows and integrate the information contained within them (Woods, 1984). Animation is one method suggested for increasing visual momentum (Norman, Weldon, and Shneiderman, 1986). For example, with a perspective layout window, the rotation of an object can be simulated. If a window is "shrunk" into an icon, animation can show the window shrinking from a window into an icon. The few seconds a controller spends viewing window animation can contribute greatly to maintaining orientation, seeing relationships between windows, and understanding the windowing metaphors. When a person loses orientation, s/he may experience a sense of "getting lost" in the system and be unable to tie the active window to any others.

Two potentially serious problems are associated with managing overlapping windows. The first is that window management (e.g., accessing, sizing, moving, overlaying, and closing windows) takes time and effort. This can increase the controller's workload and distract him/her from the primary task of ATC. Some systems force the user to spend too much time managing windows, in some cases, even more time than is spent accomplishing the task at hand.

The second problem is a special case of window management that occurs when the number of windows in use simultaneously becomes too large. The result is a phenomenon called **thrashing** (Card, Pavel, and Farrell, 1984). So many windows are needed to perform the task at hand that nearly all of the user's time is spent finding windows and bringing them to the top of the stack. Thrashing promotes mental overload because the user must keep track of window locations and pertinent data items within covered windows.

These two problems suggest the following window design goals if windowing is to be used successfully in ATC applications: **All information that a controller needs to accomplish the current task should be located in a single window or a small set of windows.** Researchers also suggest providing the task-related commands (e.g., using push buttons within the window) so the user can recognize commands and not be required to recall them (Chignell and Waterworth, 1991). **The system designer should work with controllers to determine the most usable size and placement of windows so that the necessity for window management will be minimal.** Results of the task analysis and rapid prototyping can contribute to the process of determining window data and command content as well as window placement and size.

It is important to recognize that users seek out relationships between windows and sometimes infer relationships that do not exist. Windows that are unrelated, yet share similar features, will be grouped by the user. Some effective design methods for promoting window grouping are to display related windows in close proximity and/or have them share the same size, shape, color, or vertical or horizontal placement.

Windows that are independent of each other should have contrasting features. Users infer ordering of windows primarily by location (the windows that are above or to the left of the others are believed to come first or to be higher in the hierarchy or levels of windows). Another method to promote intended window ordering is to have windows in a window set appear sequentially over time. When windows are not dated, it is especially important that data should automatically fill sequentially in accordance with the proper ordering of windows.

7.2.18 What are the major design issues associated with menus?

Menu design looks easier than it is. Issues facing the designer include:

- the number options to offer
- how to phrase and format the options
- how to organize options within the menu
- the number of levels to use
- how to help the user maintain a sense of orientation within the menu structure.

Research on menu design has shown that design solutions to these issues are important for human performance. Both direct and indirect costs can be reduced, and benefits can be increased through menu design that complements human capabilities and protects against errors.

Number of Options. This issue has received a great deal of research attention. There is, however, no single answer to the question of how many options to present because the answer varies from one context to another (See Galitz, 1993, for a review.) Where possible, four to six options is considered optimal because it limits the demand on working memory. Having too many options in a menu increases the time it takes to find any one option. Having fewer than four probably means that levels of the underlying menu structure could be combined so that more options could be offered in one place. With few options in a given menu, the user may be required to select options from several lower-level menus in order to reach a specific location in the menu structure.

In contrast to the general recommendations just given, research on so-called "vast" menus suggests that a menu can contain up to 200 options if they are organized into logical groups (Norman, 1991). With large numbers of options, however, problems arise in developing a menu structure and graphic layout that reflect the underlying structure of the domain information. For example, if the underlying structure is a hierarchy, the menu structure should be hierarchical; but large hierarchies are difficult to traverse efficiently. Before vast menus are considered for use in ATC applications, satisfactory methods will need to be in place to resolve the problems of structure and layout.

Phrasing Menu Options. Research has shown that vague, ambiguous wording of menu options confuses the user and detracts from performance. The purpose of careful phrasing is clarity, so that users understand the meanings of the options in terms of how the system will respond. For options that are to be executed immediately upon selection, a general recommendation is to use wording that reflects the actions to be executed. Notice that the emphasis is on action. Actions are denoted by verbs or verb phrases, not by nouns or noun phrases. Common actions that might be offered in an ATC handoff menu include Accept, Deny, Initiate, and Retract.

Two other types of options are 1) routings that display a window or a cascading submenu; and 2) settings that are used to define parameters or to specify an application state. Routing options, such as Search, Customize, Sort, or Save As, should be followed by three dots (an ellipsis) to indicate that another small window, called a dialog box, will open when this option is selected. The purpose of the dialog box is to present further choices that are related to the higher-level action, for user selection before the system can proceed. For example, before carrying out a command to sort metering advisories, the system may need to know the criteria on which the sort is to be based. The alternatives should be presented in a dialog box.

Instead of imposing new terminology or using ATC terms to mean something different from what controllers understand them to mean, the wording used in menus should be consistent with ATC vocabulary as it is used in the ATC environment. This means that designers and developers should survey ATC facilities and the FAA's operations concepts to identify terms that are commonly used for system actions.

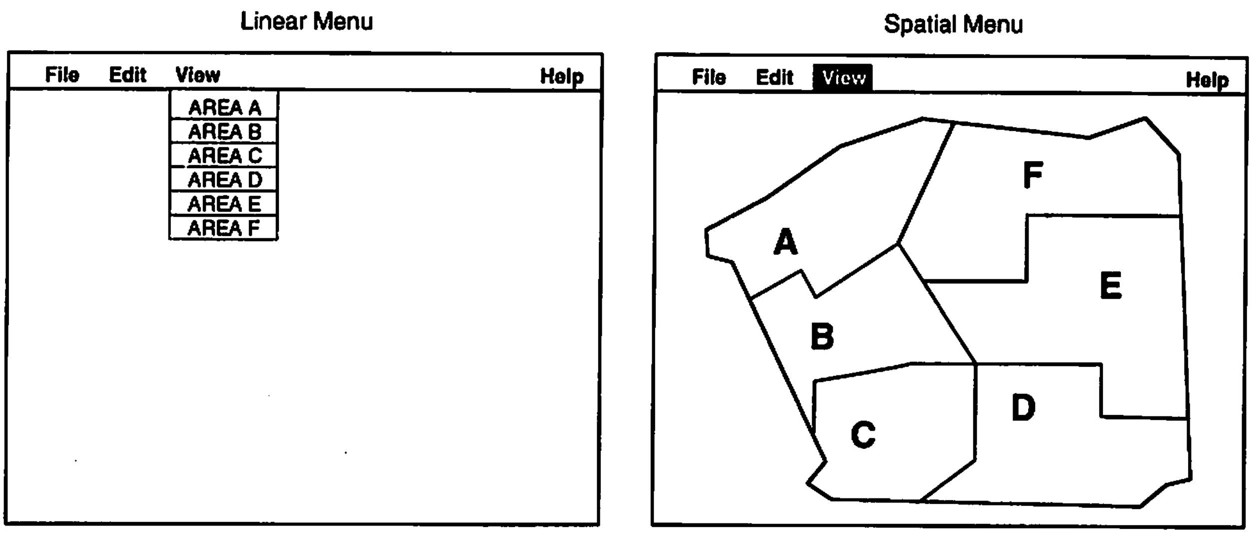
Option labels should be concise and used consistently from one menu to another. They should be distinctive and mutually exclusive (non-overlapping). Controllers can provide expert assistance in the development and review of menu phrasing.

Formatting Menu Options. Menu options can be formatted in a linear (vertical or horizontal) list or presented in a spatial format. Although most menus tend to be presented as lists, some sets of menu options might best be represented in a rectangular or circular format. Spatial menus for ATC might take advantage of existing spatial formats, such as the geometry of sectors within an area.

Whether a menu is linear or spatial, each word in a menu should be presented in upper and lower case with the first letter capitalized. When a vertical menu is displayed, the location cursor should be on the first available option closest to the top of the list. When a horizontal menu is displayed, the location cursor should be on the first available option to the left.

Menu options presented in a vertical list should be left-justified, never centered in relation to each other. Left justification allows the user's eye to follow a direct line, rather than a jagged line, in searching for a particular menu option. Other forms of justification increase search time and contribute to visual fatigue.

Figure 7-8. Examples of Linear and Spatial Menus



A window that contains a hidden pop-up menu or menus should indicate the availability of these menus to the user. One way to do this is to highlight the portion of the display that can be selected to access the hidden menu. Another way is to provide a textual message indicating that a hidden menu is available. A third approach is to change the shape of the cursor when it is located in a "clickable" or selectable area. When a pop-up menu is selected, it should be displayed in context, that is, close to associated information. For example, if there is a pop-up menu for each radar target, the menu should be displayed near the associated target, not halfway across the screen.

An option or set of options that is never available to the user should not appear in the menu. If an option is temporarily unavailable, it should be displayed in its normal place in the menu but subdued in intensity (dimmed down or grayed out).

Organizing Menu Options. There are several acceptable ways to organize menu options:

- in logical or functional groupings with clear titles
- by frequency of usage, with the most frequently used options at the top or beginning of the list
- in alphabetical, numerical order, or chronological order

If it is possible, logical or functional grouping of menu options is recommended because it is more meaningful than the other methods. (A meaningful order is easier to work with than an arbitrary order.) If frequency of usage is the basis for organizing menu options, the less frequently used options and destructive commands (such as Delete or Exit) should be placed at the bottom of the menu.

To protect against accidental activation, menu options that initiate opposing actions (such as Save and Delete) should not be placed next to each other.

For the sake of positive transfer, similar options that appear in different menus should be ordered consistently from one menu to the next. Any cascading menus should appear to the right of the parent menu, or below the parent menu if there is insufficient space to the right.

The process of organizing menu options should always occur in the context of the tasks to be performed by the controller. A menu organization that supports one set of tasks may not be as good for a different set of tasks.

Number of Levels. Research has shown that when more than three to four levels are included in a menu structure, users tend to feel lost or disoriented (Shneiderman, 1992). To reduce the likelihood that this will happen, menu users should not be required to move more than three to four levels down into a menu structure to locate a desired option.

Sometimes it may be necessary to make a tradeoff between the recommendations on number of options (menu breadth) and number of levels (menu depth). In such a case, evidence indicates that it is better to increase the number of options per level instead of adding to the number of levels (Shneiderman, 1992).

If the menu structure includes more than a few levels, designers should consider providing a way for the experienced user to by-pass intermediate levels. In satellite control centers, for example, operators can construct their own "fast-action" menus of frequently used menu options from the entire menu structure. Using fast-action keys, operators can then go directly to specific locations in the overall structure without having to step through intermediate menus.

Supporting User Navigation. When working within a large, multi-level menu structure, users often lose their bearings, forgetting where they are and which levels they have traversed. To some extent, this disorientation relates to the limits of working memory because it is difficult to keep in mind one's position four or more levels down in a menu structure while also concentrating on the task at hand. To assist users in navigating easily within a menu structure, graphical or textual aids should be provided.

A graphical aid may take the form of small schematic of the menu structure with the path already taken highlighted. This miniature menu map may be displayed continuously in a corner of the screen or made available from a pull-down menu. An alternative to the menu map is a sequential list of the options that have been selected. If they are not displayed continuously, menu-navigation aids should be available in the help function.

Resolving Detailed Design Issues. This discussion has surveyed major issues in menu design, but many other detailed issues will need to be resolved. Comprehensive treatments in the literature can provide the necessary guidance (e.g., Avery and Bowser, 1992; Fernandes, 1993; Galitz, 1993; Mayhew, 1992; Norman, 1991; Paap and Roske-Hofstrand, 1988; Shneiderman, 1992). Tentative menu designs should be subjected to systematic testing with representative groups of controllers. The key objectives of menu design, testing, and re-design are to minimize search-and-selection time and to minimize navigational errors.

7.2.19 How can visual display designers achieve consistent display formats?

Consistency allows the user to build up visual and cognitive momentum in support of smooth, efficient operational performance. **A key visual display goal is to achieve consistent display format in terms of both form and layout of data.** For example, placement of standard data fields should be consistent from one display to another, especially in a related set of displays. This allows the user to build skills based upon expectations for item placement. The format in which the information is presented within the data fields should also be consistent from one display to another (used by the same controller). One way to meet this goal is to develop and enforce design rules based on standard guidelines. Visual displays for a complex system, such as ATC, are produced by the combined efforts of many software engineers. Unless a detailed style guide is developed and followed, there will be numerous glaring and subtle inconsistencies, especially between displays having different developers.

A style guide contains design rules tailored for the specific application. Although the sample design rules that follow are not absolute for ATC, they point to factors that must be used consistently. **If a design rule is followed for one display, then it should be followed for other displays used by the same controller. The location of standard data fields and**

the format within data fields should be consistent from one display to another.

Design rules should be defined for such factors as label placement, use of punctuation, and use of separators. Typical design rules are illustrated by the following partial style guide, which was developed for a military air, mission-planning application (Lockheed Sanders, in preparation):

- Label placement is to the left of the associated field.
- A colon is placed immediately after all field labels.
- Data entry fields are surrounded by a rectangular box indicating maximum field length.
- Wherever a unit of measure is associated with a field, the unit of measure is displayed one space to the right of the field.
- Time and date are presented as MM/DD/YY HH:MM:SS.
- Helvetica font is used for all text.
- Assuming a viewing distance of 18 inches, .11 inch character height is used for the main menu title bar and main menu options; and .11 inch character height is used for window and display titles.
- .09 inch character height is used for all other text.
- Command push buttons and command icons are placed at the bottom of a display or window and left justified.

Specific format style should be determined during the system design phase so that the style guide items (i.e., design rules) are available to software engineers before software coding begins. An effective way to evaluate a system for consistent display format is to create a color hardcopy (i.e., print out) of all menus, windows, display pages, and similar display elements, and compare each element to a detailed format style guide or checklist having the same level of detail as the one presented above (CTA, 1993). Automated design evaluation tools will also be able to make such comparisons and identify inconsistencies.

Readers interested in a further discussion of CHI consistency and inconsistency are referred to papers that support differing viewpoints (Grudin, 1989; Kellogg, 1987, 1989). In the ATC context, a strong case can be made for consistency, but this does not mean blind consistency irrespective of task requirements.

7.2.20 How should visual displays balance computer versus user control of data content?

Displays should be task oriented, that is, each display page, window, or window set should support a high-level ATC function or lower-level ATC task. Data that are related to the same controller task should be displayed on the same page or window. This minimizes the number of simultaneously-displayed windows, switching between display pages/ windows, and presentation of data irrelevant to the task at hand. **It is a key design goal to ensure that the computer presents all data required to accomplish individual or team tasks without any extraneous data cluttering the display. At the same time, controllers should be able to modify the amount and detail of task-related data being presented** (Smith and Mosier, 1986). At no time, however, should the computer be solely in charge of how much information is presented on the display. **Critical information should never disappear from the screen without being deleted or suppressed by the controller.**

User control of display information can be achieved through various methods, for example, by zooming in on a sector to see more graphical detail or by using a combined graphics and window-based system. Using a combination of graphics and windows, a controller could access high-level information using graphical objects (e.g., a display showing air traffic), and the controller could request a more detailed window (or any in a series of windows having increasing detail) on each object (e.g., an aircraft). Another control feature would allow the user to name a window display setup, including the type, sizing, and organization of windows, for later access. A windowing environment provides the user a great deal of flexibility on the current data display content.

The controller may want to view different parts of a data set simultaneously (e.g., flight data relevant to the next 10 minutes and flight data relevant to a high traffic situation expected to occur in 45 minutes). The controller should be able to access a window multiple times and, by resizing the window and scrolling the data, view the two nonadjacent sections of data simultaneously.

7.2.21 How can visual displays aid the controller's planning and decision-making process?

To support planning and decision making, controllers in the future ATC system will have access to predictive displays, for example, displays that provide a picture of the predicted air traffic environment at a user-specified time in the future. A predictive display will allow the controller to anticipate an emergency or high traffic condition and give the controller more planning time.

Predictive information, which shows what will happen if the current action is maintained or changed in specific ways, can be used with monitoring, decision-making, or control tasks. Predictive information, which is generally used in conjunction with situation information reduces processing workload by integrating data (Fadden, 1993). Predictions are based on the existing or proposed control strategy.

An example of a predictive display is the "Future Situation Display," which will allow the en route controller to "try out" the effects of possible control actions as well as the effects of letting a situation run its course. Today, controllers often predict possible outcomes by thinking through complex situations and can miss unwanted consequences. In general, people are not particularly good at predicting the consequences of multiple, dynamic variables. Even chess experts cannot anticipate more than three or four moves in advance.

Predictive information should neither tell the controller directly when to act nor demand a particular task performance strategy (Fadden, 1993). The controller must have an understanding of the current situation in order to make these decisions. Predictive displays are most useful for "tasks where both deviations from some plan or standard and some form of rate information are involved" (Fadden, 1993, p. 267). Like all features of the CHI design, they must be evaluated for usability, operational suitability, and workforce acceptance.

SELECTED EVALUATION ISSUES

In this section, we discuss methods of testing and evaluating visual display designs. We recommend minimum human factors requirements for display hardware and for visual displays. We consider the issue of ATC display standardization and recommend standardization across facilities of the same type.

7.2.22 What methods are recommended for evaluating the physical characteristics of visual displays?

The following characteristics can be measured in a laboratory or bench test environment:

- Display luminance and pixel spot size
- Display contrast
- Brightness uniformity
- Color contrast
- Symbol height
- Symbol size uniformity
- Geometric stability (Jitter)
- Flicker
- Registration
- Reflections and glare.

The exact methods for performing these tests are found in the ANSI standard for *Human Factors Engineering of Visual Display Terminal Workstations* (ANSI, 1988). These methods and recommended parameters values are summarized as follows:

Display Luminance and Pixel Spot Size: Photometric measurements made on the display in both dark and illuminated test conditions provide measures of character luminance, background luminance and display pixel spot size. Luminance profiles are measured across the width of a one-pixel-wide line on the CRT. Taking these measurements requires a photometric measuring device that is calibrated against the standard for traceable luminance (ANSI, 1988).

Display Contrast: Since the background of a display is never completely black, dim characters can wash out when the display is dimmed. The contrast of foreground text and symbols against background luminance can be measured by dividing the luminance of the foreground by the luminance of the background. This number is referred to as a **contrast ratio** or **luminance ratio.** Measurements of symbol luminance, background luminance and pixel dimensions are made at the center of the display and at each of the four corners of the active area of the display screen using a photometer. Measurements are made in a dark test condition and an illuminated test condition. In office environments, the luminance of text and symbols should be at least three times that of the background for good legibility. The International Civil Aviation Organization (ICAO, 1993) recommends that **the luminance of dynamic text and symbols be eight times that of the static background symbology on ATC displays.**

Brightness Uniformity: Non-uniformity of brightness occurs when there are noticeable dark and light patches on the display screen that are not due to deliberate display formatting. Brightness non-uniformities are more noticeable at low display luminance levels than at high luminance levels. These non-uniformities often result from manufacturing processes. To adjust for non-uniform brightness, the controller has to turn up the display brightness to see the dimmest portion of the display. **The brightness uniformity of the display should not vary more than 50 percent when comparing the center of the display to the edge of the display.**

Color Contrast: The legibility of alphanumeric characters and symbols is affected by the colors and luminances of the characters and the color and luminances of the backgrounds they are presented on. For adequate legibility, one recommendation is that colored symbols should differ from their colored backgrounds by a minimum of 100 color distance units as measured by an international lighting and color standard, the Commission Internationale d'Eclairage (CIE) method. This recommendation is based on the CIE chromaticity diagram, which is highly complex and usable only be experts on color vision. For additional information on color contrast, see Appendix A, Human Factors Considerations for Color Displays for ATC.

Symbol Height: Several techniques can be used to measure symbol size. The ANSI standard recommends the use of a microphotometer for the most precise measurement. Symbol height is measured from the bottom edge to the top edge of a non-accented uppercase letter. The edge of the letter is measured by the 50 percent luminance point.

Symbol Size Uniformity: One method to test symbol size uniformity is to fill the display with the letter "M" or the letter "H." The symbol luminance and background luminance of the characters appearing at the center and four corners of the display are then measured. The heights and widths of those characters should not vary by more than ten percent.

Jitter: If a display does not meet jitter specifications, it could be difficult to read and possibly unusable. Jitter is caused by excessive variations in the picture element when the display is refreshed. Jitter refers to variations in the geometric location of a picture element on a display screen. Jitter is measured by determining the maximum movement of a picture element (pixel) over a one second period for both horizontal and vertical directions. This determination should be made for characters appearing at the center of the screen and each of the four corners of the display. Unless its position is changing dynamically, a display object should move no more than .0002 X the viewing distance (in inches) in one second (ANSI, 1988).

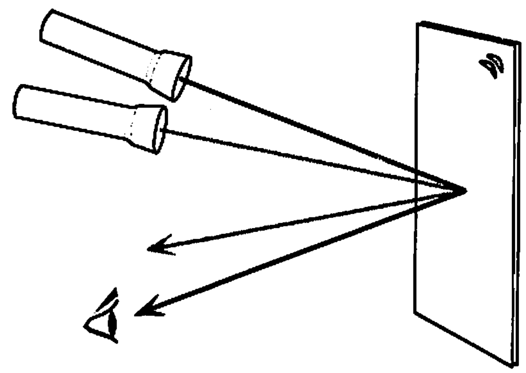
Flicker: If the display does not meet flicker specifications, the younger population of controllers could find the display unusable. Flicker can be avoided by ensuring that the refresh rate is at least 65 cycles per second (Hz). The refresh rate is the "reciprocal of the time required to produce a full screen image" (Shneiderman, 1992, p. 263). The refresh rate should be documented in the design specification. (See Section 7.2.6 and 3.1.6 of this handbook for further discussion of flicker.)

Registration: Registration refers to the accurate placement and alignment of symbols and text on the field of the display screen. If the symbology does not have proper registration, the overlap of symbols and text could make symbology impossible to read. For display of radar data, registration of radar tracks is critical as the symbols display the location of the aircraft in space. Any errors in registration will result in inaccurate placement of radar tracks. Exact measurement of horizontal and vertical placement "may be made perpendicular to a plane that is tangent to the center of the display surface" (ANSI, 1988, p. 69). there should be no more than a 5 percent horizontal or vertical variation between the location of one symbol position and another (ANSI, 1988).

Reflections and Glare: Glare and reflections contribute to the non-uniform brightness of the display since the luminance of the reflection essentially sums with the luminance of the displayed image in terms of the amount of light striking the eye. Reflections are often caused by sources of light, such as illuminated indicators reflecting off a CRT screen. Secondary reflections can also be caused by light reflecting off shiny work surfaces and onto the display screen. Glare is caused by diffuse light sources, such as bright overhead lighting or sunlight. Reflections and glare can be controlled in various ways, as follows:

- *Shifting the geometry of the light sources and the reflection surface.* Reflections from point sources of light will reflect in a mirror image as shown in Figure 7-9. The phrase "point sources of light" refers to light emanating from a single source, such as the filament of a light bulb. Shifting the geometry of the light source and reflection surface can prevent reflected light from striking the eyes. The geometry between light sources and reflective surfaces can be evaluated during the design phase, and potential problems can be eliminated. Some indicators and CRTs can be mounted on a swivel base, allowing the controller to shift geometry.

Figure 7-9. Result of Shifting the Geometry of Light Sources and Reflective Surfaces



- *Using shields or hoods to contain the light emissions from light sources.* Shields or hoods can be used when it is not practical to shift the geometry of the light source and reflective surface (such as in the cab of a control tower). However, when the display needs to be viewed from multiple angles such as from both a seated and a standing position, hoods may block a controller's view.

- *Applying anti-reflective coatings or filters to the display surfaces.* Although coatings and filters can improve display contrast, they are recommended only if other solutions are impractical. Coatings can degrade over time, (particularly with improper cleaning) and both filters and coatings can reduce image quality.

Evaluating and adjusting the physical characteristics of visual displays are fairly straightforward activities, yet they are critical to the usability and suitability of these displays in ATC facilities. *For a detailed discussion of coating and filters see section 5.3 of Appendix A, Human Factors Considerations for Color Displays for ATC.*

7.2.23 What kind of testing should be conducted to evaluate display symbology and display design in general?

A visual display exists to communicate information. The successful visual display is the one that promotes comprehension and consistent task performance. Therefore, any evaluation should first determine "whether the primary task performance defined during the early requirements phase has been achieved" (Fadden, 1993, p. 250). This is where having clearly defined and measurable performance criteria early in the development cycle benefits the designer and helps to ensure operational suitability. Testing visual displays as early as possible at field sites is highly recommended (Harwood and Sanford, 1994).

After testing has determined that the display supports the expected performance on the intended task, it is necessary to assess whether the performance of other tasks has been affected. Typical contributors to performance degradation for visual displays include the following (Fadden, 1993):

- Perceived symbol movement caused by actual movement (i.e., jitter) of nearby symbols.
- Display clutter.
- Poor legibility, discriminability of symbols.
- Excessive dominance of an unrelated nearby symbol, due to relative size, color, brightness, or shape differences.
- Use of symbology or display formatting practices that are inconsistent with controller expectations.
- Use of similar symbols to support different tasks.
- Misuse of color and color combinations.

If an already complex display is altered to support performance of a new task, it is important to confirm that the required level of performance for previous tasks can still be achieved (Fadden, 1993). After task performance has been assessed for every task associated with the integrated display, the evaluation should be expanded to include all applicable task-display combinations in the ATC facility (Fadden, 1993). One objective is to identify unintended visual effects that arise because of interactions between display elements (Tufte, 1990).

Over time, some tasks for which the display was designed may change due to changes in the ATC operating environment or skill/knowledge base of the controllers. It is important that controller comments generated through operational use of the display be solicited and evaluated against the original design intent. Controllers at some facilities may have found uses for a display that were not originally intended but that improve their efficiency and effectiveness. These operational strategies should be documented and distributed to other facilities. Future decisions about use of the display, associated training, or operational enhancements should be based on an accurate understanding of the controllers' tasks and how well they are supported by displayed information (Fadden, 1993). (Chapter 10, Human Factors Testing and Evaluation, provides guidance on planning and conducting human factors tests and evaluations.)

7.2.24 When trade-offs need to be made, what human factors standards should be considered minimum requirements?

During the design of any system, trade-offs must be evaluated and made. Cost, maintainability, and reliability can all trade-off against the incorporation of new display technology and CHI improvements. Human factors requirements, which could impact controller performance if traded off, include the following key recommendations from this chapter:

Minimal Requirements for Display Hardware Issues.

- *Jitter.* Variations in the location of the picture element should be less than 0.0002 inch horizontally or vertically per inch of viewing distance over a period of one second (ANSI, 1988).

- *Flicker.* The display refresh rate should be at least 65 cycles per second. (See section 7.2.6)

- *Registration.* Symbols and text must be accurately located and aligned, with no overlap. Horizontal and vertical displacement of a symbol position relative to adjacent positions should not vary by "more than 5 percent of the symbol box height" (ANSI, 1988, p. 23).

- *Brightness control.* Perception of brightness is influenced by several factors including age, ambient illumination, and the brightness of objects near the display surface. Controllers must be provided with a capability to smoothly adjust display brightness.

- *Ambient lighting.* Ambient lighting should not include any colored lights because they will affect the perception of colors in the display. If the facility illumination is too bright, it will degrade displayed information, but, if the room is too dim, the controller will have trouble reading important information from other sources, such as a paper printout. Under various expected ambient lighting conditions, computer displays, illuminated keyboards, illuminated indicators, other mechanical controls, and other visual displays should be clearly visible and appear evenly lit.

Minimum Requirements for Visual Display Design.

- *Minimum symbol size.* Since optimal size is task-dependent, the height of text and symbols should be adjustable by the controller. Ergonomic recommendations for minimum size and spacing should not be compromised (ICAO, 1993).

- *Redundancy of coding.* It is important to have more than one type of coding for symbols so that color-deficient users will be able to acquire necessary information from the visual display. Typically, both shape and color coding are used together to meet this requirement. Controllers must always be able to discriminate between shapes and to specify shapes verbally using obvious shape names (ICAO, 1993). If color is not available on a display or display format, this should not hinder the controller from accurately identifying the information from the display symbology. Although the need for redundant coding may differ from task to task (Narbrough-Hall, 1985), the literature generally favors redundant coding over the use of color by itself. In any case, overuse of color coding must be avoided.

- *Adequate Labeling:* A controller should be able to interpret text items and other alphanumeric information easily. Abbreviations should be used consistently on different display formats and displays. Icons should generally have text labels below them.

The primary objective of any ATC display is to communicate information to the controller in a format that promotes comprehension and consistent task performance. These objectives can best be achieved through an understanding of the ATC tasks that need to be performed and of the information required for satisfactory task performance. ATC displays must meet criteria for usability, operational suitability, and workforce acceptance. An effective display is one that supports consistent accomplishment of assigned task while providing opportunities to enhance job satisfaction.

7.2.25 What are the pros and cons of display standardization across ATC facilities?

The advantages of standardized displays include cost savings and reduced training time. Of course, these advantages are only realized if standardized displays support underlying tasks. For tasks that differ even slightly, standardized displays may result in degraded performance and require additional compensatory training. A careful task analysis is necessary to determine the appropriateness of standard displays for a particular application and to determine when displays must be tailored to one position or task.

Controllers working in ATCT, TRACON, and ARTCC facilities use different information elements in the performance of their daily tasks. Information that a tower controller needs to know, such as ceiling, will be of limited value to an en route controller, unless the ARTCC is providing Approach Control Services. Similarly, a TRACON controller's information requirements may be different from those of a tower or en route controller. However, there may be instances when the same information element is required but used in a slightly different manner at all three types of facilities. In such cases, it would be unwise to use a standard display format at each of the facilities.

ATC functions that are common across ATC facility types promote display standardization, particularly for functions that are well known and stable. In addition, NAS components to be fielded in en route facilities, TRACONs, and towers, are based on a "national" concept of ATC and will not work as intended unless standard displays and procedures are instituted nationwide. To the degree that detailed tasks can be standardized across ATC facilities of the same type, a high level of display standardization should be within reach of display designers. In addition to reducing the cost of ATC display development, the positive transfer will lead to reduced training time and error rates for controllers who transfer between facilities.

7.3 AUDITORY DISPLAYS

This section addresses the human factors issues associated with auditory displays. The term "auditory display" may sound like a contradiction because, typically, we think of a display as being viewed, not heard. **Auditory display** is the phrase commonly used in the field of human factors, however, to denote the use of sound to transmit information to a human receiver. Just as light energy is displayed or presented to the eye, sound waves are displayed or presented to the ear. Under some conditions, auditory displays are preferable to visual displays. Use of both auditory and visual displays is sometimes recommended to enhance human performance.

7.3.1 What are the basic types of auditory displays?

Auditory displays can be divided into two main types: signal (also called tonal or sound) displays, and speech (or voice) displays. Signal displays may take the form of tones, beeps, buzzing sounds, ringing bells, and so forth. Speech displays are verbal messages delivered by "live" speakers, recordings, or voice synthesizers. Typically, auditory displays are used either to present low-level information in the form of alerts, alarms, and warnings (using signal displays), or they are used to transmit speech communications (Kantowitz and Sorkin, 1983; Stokes, Wickens, and Kite, 1990).

Table 7-1 summarizes different kinds of audio displays that are appropriate for various tasks and environments. Table 7-1 is expanded upon in the following sections into a detailed discussion of the appropriateness and advantages of auditory displays; problems associated with auditory displays; and principles, standards, and guidelines applied to auditory displays.

Table 7-1. Audio Display Summary
From *Human Engineering Design Criteria for Military Systems, Equipment, and Facilities* (MIL-STD-1472D), Department of Defense, 1989, Table 3, p. 56.

Function	Tones (Periodic)	Complex Sounds (Non-periodic)	Speech
Quantitative Indication	POOR Maximum of 5 to 6 tones absolutely recognizable.	POOR Interpolation between signals inaccurate.	GOOD Minimum time and error in obtaining exact value in terms compatible with response.
Qualitative Indication	POOR-TO-FAIR Difficult to judge approximate value and direction of deviation from null setting unless presented in close temporal sequence.	POOR Difficult to judge approximate deviation from desired value.	GOOD Information concerning displacement, direction, and rate presented in form compatible with required response.
Status Indication	GOOD Start and stop timing. Continuous information where rate of change of input is low.	GOOD Especially suitable for irregularly occurring signals (e.g. alarm signals).	POOR Inefficient; more easily masked; problem of repeatability.
Tracking	FAIR Null position easily monitored; problem of signal-response compatibility.	POOR Required qualitative indications difficult to provide.	GOOD Meaning intrinsic in signal.
General	Good for automatic communication of limited information. Meaning must be learned. Easily generated.	Some sounds available with common meaning (e.g. fire bell). Easily generated.	Most effective for rapid (but not automatic) communication of complex, multidimensional information. Meaning intrinsic in signal and context when standardized. Minimum of new learning required.

Audio displays are advantageous for numerous ATC situations (DoD, 1989b; Sanders and McCormick, 1993). The most appropriate type of auditory display (speech or signal) depends on the specific tasks and information. Appropriate testing must be carried out with representative controllers to determine whether or not a particular kind of auditory display will benefit ATC task performance.

APPROPRIATENESS AND ADVANTAGES OF AUDITORY DISPLAYS

In this section, we discuss the issue of when to use auditory displays in the context of ATC tasks and conditions. We also consider the unique advantages of signal displays and speech displays. Principles and guidelines on the use of auditory displays are drawn from the human factors literature.

7.3.2 What are ATC tasks and conditions for which auditory displays are preferable to visual displays?

For many controller tasks and ATC conditions, audio presentation of information is preferable to visual presentation (Streeter, 1988). Audio displays are independent of user position and eye fixation. In the tower, for example, controllers may need to focus vision on the physical environment. As the ATC system becomes increasingly visually demanding, use of signal and speech displays provides one potential way of shifting from an over-utilized visual mode to the less-loaded auditory mode, thereby communicating critical information without increasing overall workload.

Auditory displays include beeps, buzzes, tones and speech messages. They are harder to overlook or ignore than visual displays as they automatically attract the listener's attention. **Because auditory displays can be distracting, and because so much of the controllers task already depends on auditory information, auditory displays should only be used when absolutely necessary.**

Auditory displays are particularly effective for the following tasks and conditions:

- For warnings, which should usually include an alerting signal and a signal that indicates what to do (DoD, 1989b).

- For tasks that controllers have been accustomed to performing using auditory displays (DoD, 1989b).

- When the visual display mode is heavily used in presenting information to the controller (DoD, 1989b; Hix and Hartson, 1993).

- To elicit controller response to a critical transmission (sent by pilot or computer). It is desirable, for example, to present supplementary/redundant data, such as an auditory warning presented with a textual data link message or pictorial illustration (DoD, 1989b).

- For presenting continually changing information (DoD, 1989b).

- For verbal (as opposed to spatial) tasks, since these are best accomplished by auditory input and speech output (Wickens, Vidulich, and Sandry-Garza, 1984).

- When bright illumination and/or glare make viewing visual displays difficult, and when the user must focus vision away from displays or must move away from displays as in the tower (DoD, 1989b).

- For potential controller inattention caused by the **vigilance decrement**. An experimental study had operators engage in a vigilance task, a realistic prolonged sonar monitoring situation (Coloquhoun, 1975). The detection rate for monitoring the auditory display was more than twice as quick as for the visual display. An even higher detection rate, but also a higher false alarm rate, was found when auditory and visual displays were monitored concurrently.

  In ATC, at certain times and in certain locations, the target signals (representing aircraft) may be infrequent.

The occurrence of two target signals (or aircraft) in too-close proximity is another type of rare event. These situations make the controller susceptible to experiencing a vigilance decrement. **Auditory signals in conjunction with a visual display could support early detection of potential separation problems.**

7.3.3 What are the operational costs and benefits of signal displays?

Signal displays are used for various purposes, for example, to draw attention to new information, to indicate input error, and to warn of an emergency condition. Audio signals may take the form of beeps, buzzers, or sirens, or may be everyday sounds, such the sound of static that alerts a controller to an incoming communication. Depending on the type of signal presented, the processing of signal displays may impose several kinds of task demands on controllers (Sanders and McCormick, 1993):

- Detection: determining whether or not a given signal is present.

- Relative discrimination: differentiating between two or more signals presented in the same time frame.

- Absolute identification: identifying a particular signal from a signal set, when only one signal at a time is presented.

- Localization: identifying the directional source of the signal. Detection depends largely on signal intensity (volume) but can also be affected by frequency (Pollack, 1952) and duration (Munson, 1947; Stevens and Davis, 1948).

People can make relative discriminations and absolute identifications based on several signal characteristics, such as frequency, intensity, and duration. The primary indication of a sound's origin are differences in volume (i.e., intensity) and time of detection by one ear as compared to the other (Sanders and McCormick, 1993).

Auditory signal displays typically serve to warn, alert, or cue the controller to unexpected system/pilot responses. The use of a signal display also is appropriate when the information to be presented is a sound or has a sound associated with it.

There has been much discussion in the human factors literature on the topic of an alerting tone. Such a tone can be used, for example, to draw attention to new or updated information. Several sources recommend the use of a tone to precede a voice warning (e.g., Kantowitz and Sorkin, 1983). There is evidence that a detection tone preceding a synthesized voice message is unnecessary, however, and actually increases the time it takes to respond to the message (Simpson and Williams, 1980). The characteristics of the synthesized voice may be distinctive enough to act as a detection signal. If synthesized speech serves multiple purposes, for example, is used to transmit routine and critical messages, a preceding tone can be used in the latter case to indicate message urgency (Smith and Mosier, 1986). Alternately, two different synthesized voices, such as male and female, could be implemented: one for routine and one for critical messages (Smith and Mosier, 1986).

Auditory signals are intuitive and effective when they portray the sounds of an actual event (Kantowitz and Sorkin, 1983). In many cases, such sounds are intentional design features. In other instances, unintentionally implemented sounds provide controllers with valuable information cues. For example, in the en route center, static on the line before receiving a transmission provides controllers with a "raw" indicator of an upcoming message. When designers proposed a new ATC system that eliminated this natural detection display, controllers responded negatively.

Controllers can provide valuable information to designers about their informal use of audible cues. Such usage is not likely to be documented in formal reports, but it should be identified in a task analysis of the current system. Observation, videotaping with playback, and informal debriefings with the observed controllers are useful in teasing out this kind of information, which controllers themselves may not think to mention in more formal settings.

7.3.4 What audio presentation advantages are unique to speech displays?

Like signal displays, speech displays take many forms, including live voice, prerecorded messages (with or without interchangeable phrases, also called "capture" phrases), and synthesized speech.

Speech displays provide controllers with precise information on the nature of an emergency, condition, or event. There are four major advantages to speech communication (Streeter, 1988). First is its universality. From early in life, nearly everyone masters a spoken language, and human beings are effective and efficient in processing speech. The second advantage is that receipt of a spoken message is "omnidirectional," that is, independent of the visual focus of the listener. Third, the receiver can simultaneously process input using other senses, usually, vision or touch. And fourth, spoken messages are highly likely to be attended to. Speech is much more difficult to ignore or overlook than is the written medium.

Speech displays are advantageous for the following tasks and situations (DoD, 1989b; McCormick and Sanders, 1982):

- When voice communication is a quick and effective means of transmitting complex information.
- When messages are short, since people have difficulty remembering long messages, even for a brief period.
- When messages are transitory, that is, controllers will not need to refer to the message later.
- When meaning is intrinsic in the standardized signal (e.g., words, phrases) and in the standardized context.
- When messages conveying quantitative and qualitative information require a verbal response.

- When information is concerned with time-dependent events.

- When little new learning is required.

Other sources recommend the use of speech displays for the following purposes:

- For prompts, and rare or complex warnings (Smith and Mosier, 1986).

- When there is the need for a fast information transfer rate. With speech, rates of 150 to 200 words per minute are possible. Other auditory codes, such as Morse code, have a much lower transmission rate (Kantowitz and Sorkin, 1983).

- When a wide variety of warnings must be transmitted (Cooper, 1977). When there is the need to indicate the specific nature of an event or condition, alarm and warning signals can be explained using brief speech messages.

- As a redundant warning mechanism used for important conditions in conjunction with the visual display (Simpson and Williams, 1980).

- When pilot messages are non-urgent and can be addressed at a future time. Recorded or digitized human speech technologies are methods to reduce controller workload. These speech-store-and-forward techniques enable the controller to spread out the workload by listening and responding to messages when there is time available (Shneiderman, 1992).

- In order for speech displays to be effective, speech warning messages must be distinct from other voices in the control room; synthetic speech can create voice warnings that are highly distinctive (Kantowitz and Sorkin, 1983).

- Slightly increasing the length of a speech message can occasionally reduce the overall time to complete a controller-pilot communication. Although an extra word adds about 300 milliseconds, having the added word in a very short message can reduce the time it takes to react to the message (Simpson and Williams, 1980).

- Data link communication when combined with the use of prerecorded messages (rather than visual text) produces messages that can be quickly sent and received and that do not overload visual capabilities (Kerns, 1991).

The name of the controller's position and a short message can be used to divert attention from current auditory involvement to a critical message (e.g., "Newton High, handoff"), in much the same way that flashing symbols grab visual attention. ATC procedures include phraseology for critical messages (FAA/ATO, 1987, FAAH 7110.65).

LIMITATIONS OF AUDITORY DISPLAYS

In this section, we review the limitations of auditory displays and recommend that they not be used to represent spatial information. We also discuss problems in implementing auditory displays.

7.3.5 What are the limitations of auditory displays, and when is their use inappropriate?

There are several limitations and drawbacks to the use of auditory displays:

- Comprehending spoken language is slower than reading. For spoken words, people prefer about 160 words per minutes (wpm) (Simpson and Navarro, 1984) versus up to 300 wpm for reading text.

- It is not possible to scan auditory displays to pick out critical data (Smith and Mosier, 1986).

- Spoken language is transitory. The listener cannot review or preview, but options to rewind and fast forward prerecorded message can be provided.

- A poor spoken interface is more annoying (since it is more difficult to ignore) than a poor textual interface.

- A controller (or pilot) can hold in working memory only a limited number of sequentially presented speech messages and might remember only one or two (Smith and Mosier, 1986).

- There are limited standards for auditory display design, and these are not consistently followed (Kantowitz and Sorkin, 1983).

- Auditory signals can distract other controllers and can be overheard by adjacent co-workers (Smith and Mosier, 1986).
- Signal displays offer limited categories of warning and cautionary codes (Smith and Mosier, 1986)

- Auditory displays are inappropriate as a means to represent spatial information, but can be used to supplement pictorial displays (Smith and Mosier, 1986).

These limitations suggest that designers should use auditory displays conservatively and that such displays should be verified for usability, operational suitability, and workforce acceptance.

7.3.6 What are some design problems that have been encountered in implementing auditory displays?

Research has shown that certain aircraft cockpit auditory displays present uncomfortably loud, adverse auditory signals (Cooper, 1977; Patterson and Milroy, 1979). In both cockpits and air traffic control rooms, speech warnings can sometimes interfere with communications between controllers and pilots (Shneiderman, 1992). In nuclear power plant control rooms, auditory displays have failed to convey information about the cause of the alarm. There has been inadequate implementation of volume; some signals have been masked by noise, and other signals have been so loud that they have been unpleasant and disruptive.

Another problem in the nuclear power control room environment has been the occurrence of "nuisance" or false alarms. The result of numerous false alarms is a "cry-wolf" syndrome, which leads to a loss of confidence in the computer and a casual attitude towards the ringing of alarms (Kantowitz, 1977; Seminara, Gonzalez, and Parsons, 1977).

False alarms are a potential problem with any alerting system. A high false alarm rate can contribute not only to lack of confidence in the system but also to a slower response time to the alerts. Imagine a hotel fire alarm that is activated three times in fifteen minutes. After each alarm is an announcement that the alarm is false. An hour later, the alarm is activated again, only this time there is no announcement. Would the guests of the hotel leave the building as quickly as they would if the false alarms had not occurred earlier? Similarly, if a conflict alert function is repeatedly activated in situations in which the aircraft were obviously separated, controllers would begin to loose confidence in the alert, and possibly not respond as quickly as if they believed that every alert was valid.

It is important that controllers perceive the alert to be reliable, useful, and not a nag or annoyance. An alert that gives the controller useful information in a timely manner would never be considered annoying. If the controllers consider an alert to be annoying, then the alert needs improvement. The problems with the alert may be related to its physical parameters (e.g., the pitch is too high, the alarm is too loud, not loud enough, or too long) or to its functional parameters (e.g., it doesn't activate when it should or goes off when it shouldn't). In either case, the problem will have implications for human performance. For example, any alert that generates a high percentage of false alarms will be considered annoying and eventually, controllers may begin to ignore it. Similarly, an alert that continues to be activated after it is no longer needed, or provides more information than is needed, is unnecessarily distracting. Auditory alerts should be cancelable by the controller.

PRINCIPLES AND GUIDELINES

In this section, we present design principles and guidelines that can help to overcome the limitations of auditory displays. The objective is for auditory displays to be distinctive, concise, clear, audible, and invariant. Design based on these principles and guidelines can help to ensure usability, but any proposed auditory displays for ATC must be evaluated for usability, operational suitability, and acceptance to the workforce.

7.3.7 What are some commonly applied principles for the implementation of auditory displays?

The following principles are derived from human factors research (e.g., Mudd, 1961; Licklider, 1961).

- *Don't overload the auditory channel*: For any situation, it should be impossible for more than a few auditory displays to be presented simultaneously. If there is one underlying problem, rather than activating multiple alarms, the problem should produce one high-level alarm message indicating the root cause. As an example of what not to do, the Three Mile Island nuclear crisis activated over 60 auditory warning displays, adding to the confusion (Sanders and McCormick, 1993).

- *Avoid conflict with previously used signals*: Where practical, newly installed signals should agree in meaning with previous ATC system signals.

- *Promote intuitiveness*: Auditory displays should make use of natural or learned relationships familiar to controllers, e.g., wailing signals mean emergency.

- *Promote discernability*: Auditory displays should be easily recognized among the other signals and noise. Use of synthetic speech or tones of varying frequencies will usually create a highly distinct warning.

- *Don't overload short-term, working memory*: The auditory display should not provide more information than is necessary.

- *Promote invariance*: The same signal should always be used to indicate the same information and should not be used for any secondary purpose. The meanings of specific auditory displays should be documented in a design standard or style guide to ensure consistency in their implementation.

- *Avoid extremes of auditory dimensions*: For example, a high intensity alarm could startle the controller and disrupt his or her performance. All alarms should be of moderate intensity (or volume). Alarm signals should be spaced relatively far apart and kept within a moderate range.

- *Establish intensity relative to ambient noise level*: Signal volume should be loud enough to be heard over background noise. Depending on the frequencies present in the signal and noise, warning signals need to be 5 to 15 dB more intense than the ambient noise. Messages presented in natural speech should be at least 6 dB more intense than the surrounding noise. Messages produced by synthetic speech should be at least 8 dB above ambient noise.

- *Use interrupted or variable signals*: To minimize perceptual adaptation, use interrupted or variable signals rather than steady signals.

- *Test the signals to be used*: Use a representative controller sample to make sure these users can detect the signals in their working environment. All signals should be tested in all realistic ATC work conditions to ensure discriminability. They should also be evaluated for suitability and acceptability.

- *Facilitate changeover from previous display*: When auditory displays are replacing visual displays, display both simultaneously during the transition period. In some cases, it is preferable to maintain both permanently, for example, a visual display to illustrate spatial relations and a voice message to provide a verbal reference.

- *The controller should be able to cancel auditory signals*: A capability to cancel signals will help to minimize annoyance caused by auditory alerts and prevent possible disruption to the controller's processing of other information.

In summary, auditory signals need to be distinctive, concise, clear, audible, and invariant. Testing is necessary to ensure that they meet these usability criteria.

7.3.8 What are some specific auditory display standards and guidelines?

"Ensure that [auditory] alerts and other signals related to the interface are presented in another modality, such as tactile or visual."

Avery and Bowser, 1993, p. 3-27

Guidelines and standards on user-computer interface design provide specific direction for implementation of auditory displays (e.g., Avery and Bowser, 1992; DoD, 1989b; Gilmore, Gertman, and Blackman, 1989; Smith and Mosier, 1986). These guidelines, however, were not written specifically for air traffic controllers. Audio signal displays generally serve as warnings, cautions, or as mechanisms to draw attention to visual or speech displays. While auditory warning signals that do not require immediate attention may make sense for users of other systems, extreme care should be taken when considering their use for controllers. If not used judiciously, auditory warnings may be unnecessarily distracting and take attention away from other important information. **Since most of the information that a controller has to process comes through the ears (e.g., communications with pilots and other controllers), auditory signals that compete with this should only be used when necessary and only when immediate action is required.** They should also be used redundantly with a visual display. That is, **the information contained in the auditory warning should also be displayed visually.**

A loud signal can cause a startle reflex, enhancing performance of a simple task where there is one signal and one response to it. For example, a fire bell rings, and firemen slide down a fire pole; or a bell rings, and a boxer comes out swinging. The startle reflex hinders performance in a complex environment, such as ATC, where there are many types of signals and numerous choices to be made.
Here, signals having a relatively medium level of intensity (volume) are preferable (Deatherage, 1972).

To be effective, warnings must be loud enough to be heard and understood under the worst of circumstances, yet not be so loud that they startle or annoy the controllers. How "loud" a warning signal will sound will depend upon many factors such as the age of the controllers and their history of exposure to loud noises, the momentary level of workload, distractions, and expectations. However, these are not factors that the designers of the signal can control. Rather, they must measure the frequencies contained in the signal and in the background noise, and the absolute intensity of the signal (measured in dB). **The frequencies contained in the warning signal should be between 500 and 3000 Hz** since these are the frequencies that our ears are most sensitive to. **The frequencies contained in the alarm should also be different from other alarms and background noise so that the signal will not be masked.** Modulated signals (one to eight beeps per second) and warbling sounds (varying from one to three times per second) are effective because their distinctiveness attracts attention (Deatherage, 1972; Mudd, 1961). In terms of **signal duration**, DoD guidelines suggest a **minimum of 0.5 seconds. This signal should continue (or repeat with a maximum pause of three seconds between repetitions) until either the problem is corrected or the controller switches it off** (DoD, 1989b).

Most published standards suggest that alarms should be at least 10 dB above ambient noise (e.g., see Woodson, Tillman, and Tillman, 1992). Research conducted by Boeing for airline cockpits suggest that the intensity level of auditory signals should be maintained between 5 and 11 dB (8 ± 3, depending on the frequencies contained in the signal) over the critical bandwidth ambient noise with an automatic gain control (Berson, Po-Chedley, Boucek, Hanson, Leffler, and Wasson, 1981). **An automatic gain control adjusts the intensity of the signal to compensate for ambient noise so that it is audible in a noisy situation, but not too loud in a quiet environment.** This is a very desirable feature in any environment where the ambient noise is highly variable or unpredictable.

As mentioned in Section 7.3.3, there are four basic task demands associated with auditory signal displays: detection, relative discrimination, absolute identification, and localization.

For detectability, where the controller must determine whether or not a signal occurred, a signal should have the following physical characteristics. In a quiet environment, an intensity of 40 to 50 dB above absolute threshold is needed for detection, depending upon the signal's frequency and duration. The signal (especially when a pure tone is used) should to be presented for at least 500 ms for detection (Sanders and McCormick, 1993).

For relative discrimination and absolute identification, where there are multiple signal types, differences in physical characteristics of signals can be used to inform the controller of different conditions. Possibilities for signal variations include different levels of intensity, frequency, pitch, and duration and different types of beats or harmonics.
Various warnings can be conveyed well and appropriately by voice warning systems (Cooper, 1977). Speech output displays could use different voices types (e.g., low pitch versus high pitch) to indicate different data categories. Different pitches must be used with care, however, to avoid impacts on intelligibility.

Absolute identification is necessary when a single signal is presented and each signal represents an independent category rather than a higher or lower level on the same dimension. Each signal presented has an independent meaning. Absolute identification occurs, for example, when a high-intensity signal means system shutdown and a low-intensity signal means incoming message. Absolute identification does not apply when the signal changes continuously over some range solely to indicate gradations in a condition (e.g., the higher the intensity the greater the aircraft altitude danger level).

When absolute identification is necessary, the number of audio signals used to represent the discrete, independent categories of information should be kept to a minimum.

Researchers provide varying views on the exact number of signals that can be discriminated. Depending on the researcher and the study, the recommended number is four to five (Cooper, 1977), or five to six (DoD, 1989b; Patterson and Milroy, 1980). Because of the limits of working memory, especially under time pressure, it is probably wise to stick with three to four. If more than four signals require absolute identification, it would be prudent to have the ATC system designer demonstrate, through simulation, testing that controllers can quickly and easily identify each signal in the set in the environment in which they will be used.

Localization of sound is probably not as demanding for controllers as it is for pilots, but directional cues may be useful in helping controllers associate specific communications messages with specific aircraft. A sound-localization technology under consideration for military cockpits is known as a **head-coupled auditory display** (Furness, 1986). With this kind of display, which varies the volume and timing to each ear, the pilot hears sounds through his or her headphones as if they were coming from their actual origins in the real world. This approach improves the pilot's situation awareness and makes it possible to associate each communication source with an apparent origin (Sanders and McCormick, 1993). Application of this technology in ATC environments has yet to be fully explored.

COMBINED USE OF VISUAL AND AUDITORY DISPLAYS

In complex, automated environments, such as ATC, visual and auditory displays are typically used in combination. In this section, we discuss the benefits of using a mix of display modalities. Systematic testing and evaluation are needed to ensure that any proposed mix of visual and auditory displays truly benefits controller performance.

7.3.9 When and how should the ATC system combine the use of auditory and visual displays?

For many complex systems, performance can be improved and a person's workload decreased by using a mix of auditory and visual displays (Simpson and Williams, 1980). (As pointed out in Chapter 8, Workload and Performance Measurement in the ATC Environment, however, workload should not be decreased below a moderate level needed to maintain situation awareness and alertness.)

Multiple Resources Theory accounts for the benefits of combining visual and auditory displays by asserting that humans possess several separate pools of information processing resources (Wickens, 1984; 1992). According to this theory, resources cannot be swapped between resource pools. The implication is that when different controller tasks tap into different resource pools, tasks can be accomplished without overloading the controller's processing capabilities. Generally, tasks will interfere more and, thus, impose greater human workload if they require resources from the same resource pool. (For further discussion of multiple resources theory, see Chapter 5, Human Information and Processing.)

We can often accomplish two tasks that require us to divide our attention between eye and ear with more ease than we can perform two tasks that draw upon only auditory resources or only visual resources. This is why we can easily to listen to a news broadcast while operating a vehicle, but would find it much more challenging to read that same information from newsprint while driving. Splitting tasks between modalities offers the following advantages (Wickens, 1984; 1992):

- Timesharing is generally more efficient.
- Changes in difficulty of one task are less likely to affect the other.
- We can tap into resources to perform an auditory task that are not available for accomplishing a visual task.

In some cases, splitting a task into auditory and visual components is warranted because one channel is overloaded.

A display to the other channel can solve the problems of excessive visual scanning or extreme auditory masking. Even when problems of visual scanning or auditory masking do not exist, there is still a performance enhancement to be gained from using both visual and auditory displays. Studies support this assertion for the task of detecting simultaneous pairs of targets and for a tracking and reaction time task (Rollins and Hendricks, 1980; Treisman and Davies, 1973). Any proposed mix of visual and auditory displays should be carefully evaluated for unforeseen effects in the ATC operational context.

7.4 INPUT DEVICES

The primary human factors issue in selecting and evaluating input devices is operational suitability. This means that first and foremost the input devices selected must be appropriate for the task. Also, where multiple data-input devices are use, the devices must be compatible and carefully integrated into the workstation design. This section discusses physical dimensions and recommended functions in detail, however, it is important to keep in mind that the compatibility between tasks and devices is critical for optimum task performance.

7.4.1 What are some key issues in the selection of input devices for the operational environment?

Input devices allow the user to transfer information to the technical system that performs automatic computation and control functions. Collectively, input devices are often referred to as **data entry and control devices, controls, or interaction devices**. Information input controlled by these devices can be either continuous or discrete. *Continuous* input sets the system to a value along a continuum, for example, by adjusting the brightness of the radar scope. *Discrete* input selects values among a finite set of alternatives, for example, pushing the microphone button on and off. Selection of continuous or discrete input devices should be based on the controller's task objectives.

The design, arrangement, and functioning of the controls influence the effectiveness and safety of the system. Because of the many possible interactions between human, device, and technical system, input devices cannot be regarded simply as machine elements. They must be selected on the basis of operational requirements as well as ergonomic criteria.

Design dimensions, such as shape, size, and material, must be compatible with human anatomical and physiological characteristics. Other dimensions relate to the user's task needs, such as the degree of accuracy, force, precision, and manipulation. For example, the selection of an input device has to take into account how quickly any adjustment needs to be made, the amount of force to be applied, tactile feedback, and risks involved in accidental activation. Comparisons of controls according to these dimensions are available in the published literature (e.g., Bullinger, Kern, and Muntzinger, 1987; Greenstein and Arnaut, 1988; Sherr, 1988).

Interactions among properties and dimensions sometimes make it hard to select the best control for specific requirements. In such cases, a computer-assisted decision-making aid may be useful in the selection of controls. Such a program can be designed so that all the properties and dimensions of various controls and their associated benefits are stored in a data base. A decision-aiding program could also be designed to help the engineer identify which control device does, indeed, meet the specified characteristics.

7.4.2 What are the human factors issues in the design of input devices?

Several issues need to be considered in the design and use of input devices. Some of these issues are common across devices, while others are unique to the device. Some common considerations are discussed first, followed by considerations unique to the various devices controllers are likely to use in the future. Such devices include keyboards, touchscreens, and trackballs. Although the selection of a specific type of input device will depend upon the task at hand, in no case should an overall design of input devices require frequent switching between devices. Frequent switching between devices not only adds to controllers' workload, but also sets up a situation that can induce errors.

Clearly, the chief human factors issues are usability and operational suitability. At issue is not only whether a controller can use a device to perform a task, but also

"Standard ergonomic recommendations should be applied to the controller's workspace, including the positioning, spacing, sensitivity, feedback and visual appearance of input devices, and appropriate reach distances and forces required to operate them."
ICAO, 1993, p. 15

whether the device promotes optimal task performance. An example may serve to illustrate this point. In a tower, a controller could use the keyboard to enter or manipulate data. However, the need to maintain situation awareness outside the cab window suggests that input devices requiring less heads-down time would be more appropriate or operationally suitable.

Some human factors issues that impact the operational suitability of interaction devices may be identified during developmental testing. For example, the shaft girth of a portable interaction device may be too large for a small-handed controller. This can be detected during developmental testing. However, a portable device that meets weight requirements may still be too heavy for a controller to hold during the typical four or five hours on position. This problem may not be identified during developmental testing, but it surely would surface during extended field evaluations of the kind conducted by the Center TRACON Automation System (CTAS) project (Harwood and Sanford, 1994).

7.4.3 How should keyboards, touchscreens, and other input devices be designed in support of the controller's task performance?

Controllers frequently update information in the computer and exchange information with other controllers through the computer system. Consequently, any change in the input device needs to be carefully examined because it affects critical aspects of ATC performance. In this section, several input devices are described in terms of their general operating features. This section also summarizes selected studies that examined how controllers' performance was affected by the introduction of new input devices. Both general and device-specific evaluation guidelines are provided.

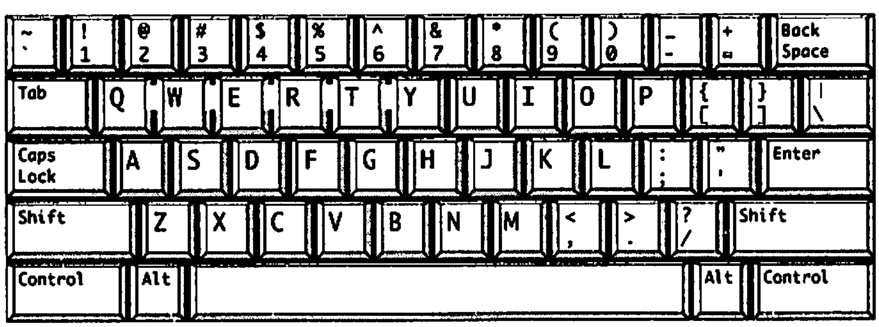
General Evaluation Guidelines for Input Devices. When assessing input devices for use in ATC, evaluators should focus on two key questions:

- Does the device meet the needs of the controller?
- Do new capabilities change the ways controllers organize their tasks?

To ensure that the performance of the device matches the needs of the controller, positioning accuracy and time lag between input and output should be taken into account when evaluating the device. Performance parameters must be assessed on tasks that closely resemble those performed by air traffic controllers, and the participants in the study should be controllers who are representative (e.g., in level of skill) of the user population.

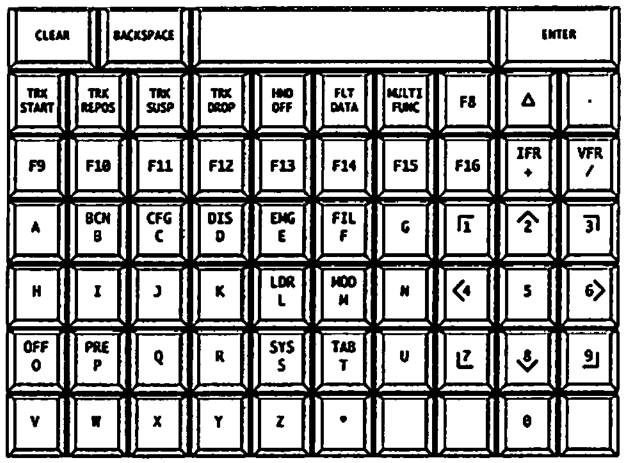
Keyboards. Standard QWERTY keyboards are still the major input devices used by en route controllers to update information on visual displays. (The term "QWERTY" comes from the arrangement of six keys to the left, above the home row [ASDF] for the left hand, as shown in Figure 7-10.) In ARTS III terminals, controllers use an alphabetically arranged keyboard as well as a QWERTY keyboard. The typical keyboard also includes arrow keys and function keys.

Figure 7-10. Typical QWERTY Keyboard Arrangement



Both QWERTY and alphabetic keyboards are sequential, with one key representing one and only one character (e.g., a number or letter), and with individual keys being activated in a sequence to produce a word or some other character string. The ARTS III alphabetic keyboard is represented in Figure 7-11.

Figure 7-11. ARTS III Alphabetic Keyboard Arrangement



Good keyboard design is essential to provide accurate and timely data entry input. The design of a keyboard can make practical differences in the speed and accuracy of data entry input. In addition, use of poorly designed keyboards can lead to muscle fatigue in the arms and possible nerve damage in the wrist and fingers.

The following factors play major roles in the design and use of standard keyboards (ANSI, 1988):

- Keyboard height (i.e., the height of the keyboard from the floor and in reference to the position of the person using the keyboard) and keyboard slope can affect comfort and performance.

- The required format of the data for keyboard entry can influence the speed and accuracy with which it is keyed into the system.

- Keyboard hardware design, which includes items such as key size, shape, spacing, and feedback, can affect error rate and typing speed. Arrangement and labeling of keys are other important issues, along with key travel (displacement), the force required to depress a key, and keyboard stability.

One way to evaluate which keyboard is best suited to carry out a set of tasks is to assess controllers' performance when using different keyboard dimensions and layouts. Such an evaluation can identify issues in the three key areas of usability, operational suitability, and user acceptance.

Evaluation Guidelines for Standard Keyboards. Because the sequential keyboard is the primary means of textual data input, its human engineering characteristics must be thoroughly evaluated. Extensive research on keyboard design has yielded numerous design guidelines (DoD, 1989a; DoD, 1989b; Gilmore, Gertman, and Blackman, 1989; Lueder, 1986; Smith and Mosier, 1986). Some of these guidelines are presented here:

- *Height, Thickness, and Slope.* Keyboard height should be adjustable by the controller within a range from 23 to 32 inches (i.e., it should be possible to raise and lower the surface that supports the keyboard). The keyboard should be less than 30 mm. thick from its base to the home row of keys, and its slope should be adjustable between 15 and 25 degrees from the horizontal (Gilmore, Gertman, and Blackman, 1989).

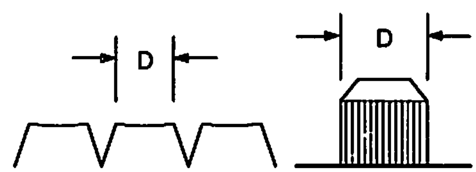
- *Arrangement, Labeling, and Feedback.* Alphanumeric keys should be logically arranged and must meet standards for dimensions, displacement, and separation, as provided in Figure 7-12. If extensive numeric data are to be entered, a separate numeric keypad should be provided, visually distanced from the main keyboard, and arranged in a 3 X 3 + 1 matrix with zero (0) centered on the lowest row. To reduce syntax error, function keys should be provided for frequently invoked commands.

Figure 7-12. Keyboard Dimensions, Resistance, and Displacement
From *Human Engineering Design Criteria for Military Systems, Equipment, and Facilities* by the Department of Defense, 1989, Table X, p. 95.

	DIMENSIONS*	RESISTANCE Numeric	RESISTANCE Alpha-numeric	RESISTANCE Dual Function
Minimum	10 mm (0.385 in.)	1 N (3.5 oz.)	250 mN (0.9 oz.)	250 mN (0.9 oz.)
Maximum	19 mm (0.75 in.)	4 N (14.0 oz.)	1.5 N (5.3 oz.)	1.5 N (5.3 oz.)
Preferred	13 mm (0.5 in.)			

	DISPLACEMENT Numeric	DISPLACEMENT Alpha-numeric	DISPLACEMENT Dual Function	SEPARATION
Minimum	0.8 mm (0.03 in.)	1.3 mm (0.05 in.)	0.8 mm (0.03 in.)	6.4 mm (0.25 in.)
Maximum	4.8 mm (0.19 in.)	6.3 mm (0.25 in.)	4.8 mm (0.19 in.)	
Preferred				6.4 mm (0.25 in.)

*Refers to dimension D shown below.



Function keys should be clearly labeled to identify their functions, and these functions should be consistent throughout the system. In addition, all keyboards to be used by the same controller should conform to one layout standard, preferably the QWERTY layout.

Requiring controllers to use keyboards with different layouts risks transfer errors of many kinds and is likely to increase the time needed for data entry. Nonactive keys should be blank; mechanical overlays should not be used to restrict access to nonactive keys. The key used to initiate a command must be clearly labeled "Enter." Keyed data should be quickly displayed or "echoed" on the screen. Tactile feedback should verify keystrokes and inform the controller when the next action may be initiated.

- *Location.* The keyboard should be located a comfortable distance from the controller and directly in front of and below the associated visual display. Forearm and wrist support should be provided to reduce discomfort, and the possibility of the controller's developing nerve damage in one or both wrists.

- *Other Considerations.* Keyboards must be readable under all operating conditions. For dimly-lit ARTCC or a darkened TRACON, appropriate backlighting should be used. Guards should be considered for any key that would present a problem if inadvertently activated.

Detailed requirements for keyboards are given in MIL-STD-280 (DoD, 1989a).

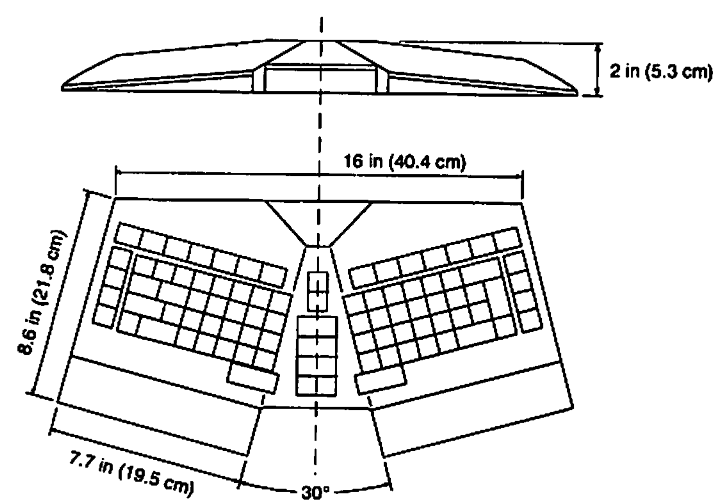
Alternative Keyboards. In addition to alphabetically-arranged keyboards, other keyboards have been developed that differ greatly from the standard QWERTY keyboard. For example, chordic keyboards contain fewer keys than standard keyboards, but two or more keys must be activated together to produce a number or letter (e.g., Gopher and Raij, 1988). One key can participate in producing several alphanumerics depending on how it is activated (e.g., up, down, side to side). With the Dvorak keyboard, named for its developer and shown in Figure 7-13, a skilled typist can enter many more frequently-used characters from the home row, as compared to the QWERTY layout (Lu, 1983).

Figure 7-13. Dvorak Keyboard Arrangement

As illustrated in Figure 7-14, split keyboards have recently been developed to reduce the strain on users' wrists. An even more recent development is the half-QWERTY keyboard, which permits one-handed typing on a keyboard half the size of the standard keyboard (Matias, MacKenzie, and Buxton, 1993). These alternative keyboards are not suitable for the current ATC environment for many reasons but may be considered in future ATC environments. Workforce acceptance is another issue to be resolved.

To avoid increasing controller workload, important objectives of CHI design should be to minimize requirements for keyed data entry and to minimize shift keying (DoD, 1989).

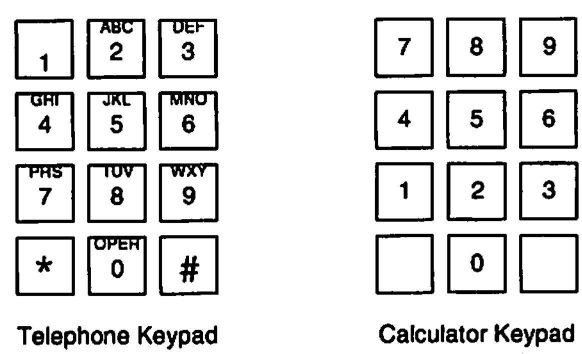
Figure 7-14. Example of a Split Keyboard
From *Human Factors in Engineering and Design* (7th ed.) by M. S. Sanders and E. J. McCormick, 1993. Copyright 1993 by McGraw-Hill. Adapted by permission.



Data-entry Keypads. Most computer keyboards incorporate a separate, numeric keypad for data entry. Telephone keypads include letters as well, usually placed on keys arranged from left to right. The arrangement of keys is the primary human factors consideration for keypads.

Determining the best layout for keypad numbers and letters has been the object of research conducted by human factors engineers. Some studies have compared performance with the telephone layout and the calculator layout, both of which are illustrated in Figure 7-15. In the telephone layout, keys are arranged in three rows of three keys with a single key placed at the bottom of the block of nine keys. In the calculator layout, the arrangement is the same but the first row has the 7, 8, and 9 digits.

Figure 7-15. Telephone Keypad and Calculator Keypad Layouts



A study tested air traffic controllers on both types of layout and found that the telephone layout was better for entries and mixed data, but performance was the same for entry of numbers only (Paul, Sarlanis, and Buckley, 1965). This implies that if keypads are used by controllers only for number entry, then both the telephone and calculator layouts are equally suitable. Any proposed keypad layout should be evaluated for usability and suitability by representative controllers.

Touchscreens. A touchscreen device generates an input signal in response to a touch or movement of the finger on the display. Typically, touchscreen devices operate by producing X and Y position data whenever a touch is detected by the system. The most common version has two conductive layers, each with an electrode grid in both the X and Y directions. When pressure is applied, the two surfaces touch, and the circuit is completed.

A newer touchscreen technology uses acoustic waves that are reflected across the active surface of the glass screen. The waves are gathered by receiving strips at the outer edges of the glass and converted into an electronic signal. When a finger, or other energy-absorbing object (such as a pencil eraser) touches the screen, part of the acoustic wave is absorbed. This results in a change in the electronic signal that is analyzed by a microprocessor and transmitted to the computer as an X and Y coordinate pair. A Z-axis may also be available that tells how hard the user is pressing the touchscreen. Thus, a light touch could be used as one type of input and a hard touch could be used as another. One caution that must be mentioned about using such a system is that while a finger and an eraser are energy absorbing objects, fingernails and the red and blue ATC pencils are not. A touch with a fingernail alone would not be recognized by the system. Thus, users will be successful with such a system only if their fingernails are short.

Touch-sensitive screens are used in data input for ATC systems, where space is at a premium and where use of a keyboard is not recommended (e.g., in a tower where head-down time must be minimized). For example, the primary VSCS user interface is a touchscreen. Using a touchscreen, controllers can be guided through complex activities. Some early problems associated with touchscreens have been alleviated (e.g., hand obscuring screen, fatigue, screen smudging), but **human-factors issues must be addressed with every specific application of touchscreen technology to ensure that it is appropriate for a particular ATC situation.**

An advantage of a touchscreen is that all valid inputs are displayed on the screen, and the relationship between the user's input and displayed output is straightforward. Although the size of controllers' fingers may tend to constrain the size of targets for land-on strategies, research has shown that touchscreen users can successfully select very small target items when using a lift-off strategy (Sears and Shneiderman, 1989).

Several strategies allow a user to select a target or menu item on a touchscreen. One study examined human performance with three such strategies: **land-on**, **first-contact**, and **lift-off** (Potter, Weldon, and Shneiderman, 1988). Using the land-on strategy, the user selects an item by touching its location on the screen or landing on the item. Using the first-contact strategy, the user drags a finger to a selectable item, which is selected on contact. The lift-off strategy allows the user to drag a cursor onto the intended target, which becomes selected when the finger is lifted off the touchscreen. This study found that the fastest strategy was first-contact, and the slowest was land-on, although the land-on strategy had the lower error rate. Thus, there was a trade-off between speed and accuracy, which is also a relevant relationship to consider when choosing among different options.

These results must be considered in the context of the test problem. In this case, participants (computer science students) had to select a target from a set of closely-spaced items. If controllers are not trying to select items that are as closely spaced, then their relative error rate is likely to be lower across the three strategies than what was found in the study. In addition, while the relative difference in error rate among the three strategies was interpreted as reflecting differences among the strategies themselves, the absolute error rate may have been inflated by the participants' lack of familiarity with the task. **The best selection strategy for controllers will depend upon the task and the conditions under which the task is to be performed. Determining the best strategy required performance testing with representative controllers under realistic conditions.**

We continue here with a discussion of additional research on touchscreens, since this technology is fairly new to many controllers. Readers interested in going directly to evaluation guidelines for touchscreens may turn to page 358.

Experimenters in an earlier study selected the target identification procedure as an example of the application of touch-sensitive screens to the ATC environment (Gaertner and Holzhausen, 1980). The touchscreen designed for this study worked in the following way. After the participant touched a displayed object (e.g., an aircraft or other item), a virtual function keyboard appeared on the lower right part of the screen, temporarily covering the information displayed there at that time. The function keyboard had a rectangular shape divided into 18 control square switches. Each function became active by simply touching its corresponding switch. The virtual keyboard could carry out computation as well. For example, one of its functions allowed calculation of the heading to be flown in order to arrive at a certain location.

When several different displayed target objects were very close together, operators marked the target area; the virtual function keyboards appeared; and the operators activated the zoom function. This allowed the aircraft targets to be clearly separated and any planned operation for any of the targets could be easily and safely performed.

The authors of this study concluded that touch input control devices will save input time compared with other control devices because these other devices require learning of control manipulation. Another advantage is that operators of touchscreens should be able to concentrate mainly on the display area without requiring their attention to be drawn to other locations. If the touchscreen is not the primary display, however, there may be some concern about head-down time devoted to the touchscreen. Performance testing is needed to explore this issue.

As part of a larger study, other researchers examined data from questionnaires on controllers' acceptance of a new system for data transfer and display (Stammers and Bird, 1980). The major aim of the new system was to provide all the relevant information on a single display unit, thus replacing the transfer of flight strips and verbal transfer of information from controller to controller. By touching areas on the display, controllers could transfer data between areas on their own screens or to other controllers' screens. Updating the display replaced the task of writing on and rearranging strips. Of course, there may be uses of strips that cannot be replaced with computer displays.

Two primary sources provided information on controller performance and assessment of the system: (1) the questionnaire, which included specific questions on displays and controls as well as general questions about the system; and (2) video recordings of controllers dealing with problems and making errors.

Study results showed that all of the users favored the touch input system. When touchscreens were rated against alternative modes of input, several advantages were found over keyboards, but most advantages were found for using a touch-input system instead of marking flight strips. Physical features of the display, however, such as screen size, distance, and angle, were not rated very highly.

Data derived from the video recordings showed the kinds of errors most commonly made in using the system. Basically, there were three types of errors: "missed touches," where the controller touched the area of the surface adjacent to the appropriate label; "error touches," when wrong labels were touched; and "illegal touches," where a touch was executed out of sequence. Controllers also reported problems with "double touches," perhaps due to parallax between the display surface and the touch-sensitive surface. The participants commented that a system providing better feedback and a faster rate of change following a human input will reduce error rate. They recommended early field evaluation and performance measurement.

These findings are not consistent with the description of advantages generally attributed to touchscreens. Minimal training, immediate feedback, limited error possibilities, and the availability of only valid options are among the advantages typically attributed to touchscreens (e.g., Boff and Lincoln, 1988). Possibly, the degree to which these characteristics are realized depends on the complexity of the task for which the touchscreen is programmed. In the study just discussed, controllers had invalid options available, and they made recommendations for improving the system feedback and for training (Stammers and Bird, 1980).

In summary, this study provided a rich body of data derived from controllers' questionnaire responses and comments (Stammers and Bird, 1980). Both quantitative and qualitative analyses generated useful information regarding the advantages and disadvantages of the touch input device. The need to include such analyses during the development of a new system design was evident. Further information on touchscreen research is available in many sources including the following: Plaisant and Sears, 1992; Plaisant and Wallace, 1990; Sears, 1990; Sears, Plaisant, and Shneiderman, 1993; Sears and Shneiderman, 1991; and Shneiderman, 1991. ATC experience with the VSCS touchscreen can provide some valuable "lessons learned."

Evaluation Guidelines for Touchscreens. The following human-factors issues should be evaluated to assess the usability and suitability of touchscreens for ATC applications:

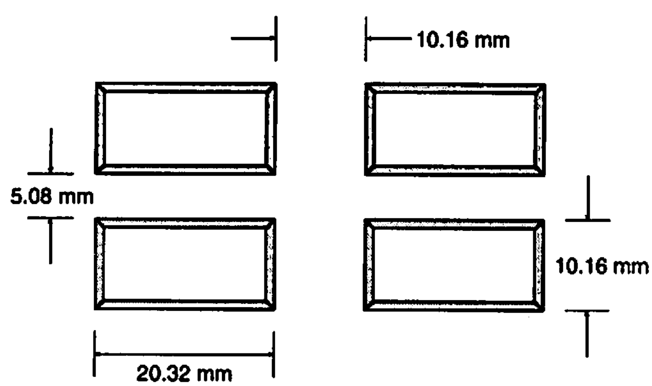
- *Device suitability.* The first issue that needs to be addressed is the suitability of touchscreen technology for the task to be performed. For a task involving considerable keyboard input, a touchscreen may be inappropriate due to the requirement to remove a hand from the keyboard home row position. Frequent switching between input devices can lead to slower task completion times and/or increased error. Another consideration in the choice of an input device is the dialogue type of the interface. Although touchscreens are suitable tools for form-filling dialogues or menu selections, they may be less suitable for other dialogue types such as command entry. Other touchscreen issues to be considered include the size of the touch area and the level of touch accuracy required in the operational setting.

- *Physical characteristics and layout.* The sensitive areas of touchscreens should be large enough to permit activation by fingers if the land-on strategy is being used. Parallax caused by the curvature of the CRT screen can be minimized by using touchscreens mounted as close as possible to the CRT surface. Touchscreen displays should have sufficient luminance to be read easily under all operating conditions. A positive indication should be provided within about 100 ms. to acknowledge activation. As with keyboards, touchscreen displays should conform to applicable minimum standards for dimensions, separation (of active areas), and resistance (DoD, 1986b). Figure 7-16 illustrates recommended dimensions for separation between touch keys. Touchscreens should not be used if the task will require controllers to keep their arms up and unsupported for lengthy periods (Gilmore, Gertman, and Blackman, 1989).

- *Input strategy.* For land-on strategies, the software accepts the touch immediately, preventing the user from verifying the correctness of the spot before activation (Shneiderman, 1992). Lift-off strategies allow the user to touch the surface, drag the cursor to the desired location, and lift off after verifying the correctness of the spot before activation. The appropriate activation strategy should be determined in the context of the controller's tasks.

Finally, as with other visual displays, the touchscreen needs to be readable under all anticipated lighting conditions; color coding must be used conservatively and redundantly with another coding technique.

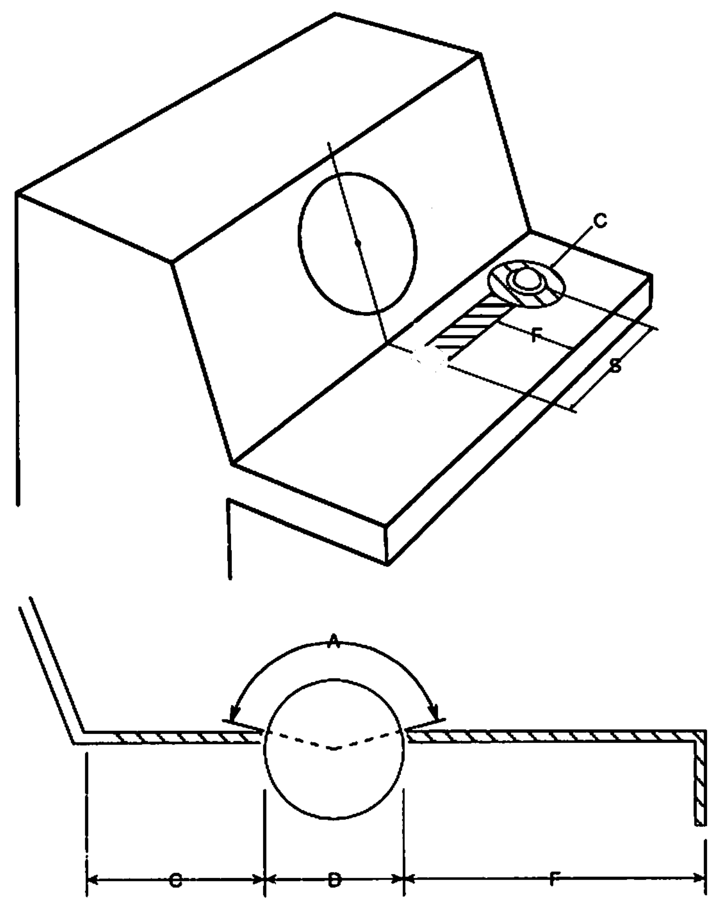
Figure 7-16. Recommended Dimensions for Size and Separation Between Touch Keys
From a study by R. Beaton and N. Weiman as cited in *Human Factors in Engineering and Design* (5th edition) by M. S. Sanders and E. J. McCormick, Figure 11-19, p. 361. Copyright by McGraw-Hill. Used by permission.



Trackballs. Ball controls called trackballs or rollballs can be used for positioning the cursor and selecting data. A typical trackball design is shown in Figure 7-17. Research conducted in the area of battlefield automated systems (Parish, Gates, Munger, Grimma, and Smith, 1982) indicated that trackballs are excellent for designating and moving symbols on a display because they are fast and accurate. Trackballs also have the advantage of familiarity. Since controllers have used them for so long, the workforce is highly practiced in their use.

Central issues with respect to trackballs are the handedness of the user (i.e., left- or right-handed) and the relationship between trackball movement and movement of the cursor on the screen. This relationship should be consistent, so that the controller will be able to predict cursor movement based on force exerted on the trackball. Precise cursor positioning depends on a predictable relationship between cursor and trackball control.

Figure 7-17. Trackball Design
From *Human Engineering Design Criteria for Military Systems, Equipment, and Facilities* (MIL-STD-1472D), Department of Defense, 1989.



	DIMENSIONS		RESISTANCE		CLEARANCE		
	D Diameter	A Surface Exposure	Precision Required	Vibration or Acceleration Conditions	S Display CL to Ball CL	C Around Ball	F Ball to Shelf Front
Minimum	50 mm (2")	1745 mrad (100°)			0	50 mm (2")	120 mm (403/4")
Maximum	150 mm (6")	2445 mrad (140°)	1.0 N (3.6 oz.)	1.7 N (6 oz.)	320 mm (12-5/8")		250 mm (9-3/4")
Preferred	100 mm (4")	2095 mrad (120°)	0.3 N (1.1 oz.)				

mm - millimeter
N - Newton
mrad - milliradian

Evaluation Guidelines for Trackballs. The following human-factors guidelines should be met by the design of any trackball for use in ATC operations:

- *Dynamic Characteristics.* The trackball should be able to move the cursor in any direction without displaying any cross-coupling (i.e., cursor movement in the opposite direction). Cursor control ratios should permit both rapid gross positioning and smooth, precise fine positioning. When it is possible to drive the cursor off the edge of the display, indicators should be provided to advise the controller about how to drive the cursor back onto the display (DoD, 1989b).

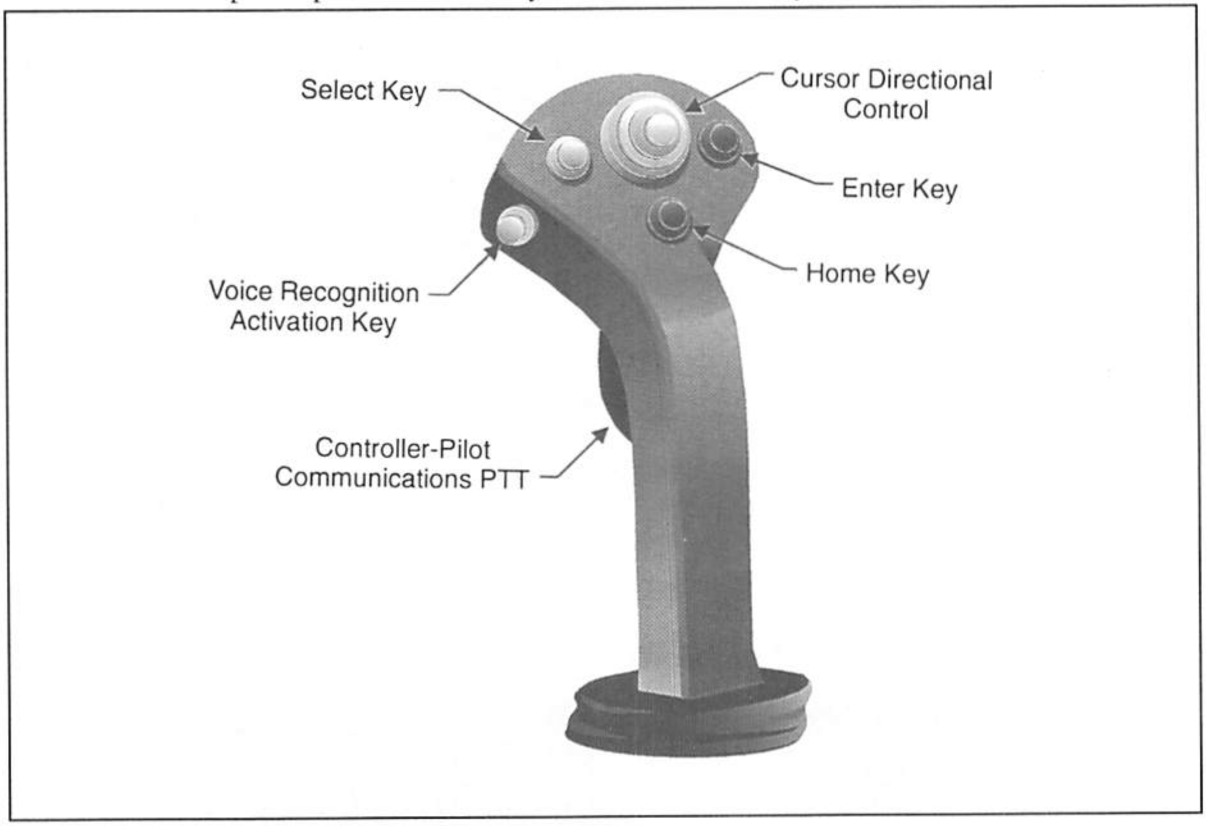
- *Limb Support.* Support should be provided for the controller's wrist or arm when the trackball is used for precise or continuous movement (DoD, 1989b).

- *Dimensions, Resistance, and Clearance.* Physical dimensions, resistance, and clearance on the work surface should meet the criteria given in Figure 7-16 from MIL-STD-1472D (DoD, 1989b).

In addition, a discrete mechanism should be provided to activate and deactivate the device. In general, smaller trackballs should only be used where space is at a premium and precise positioning is not required (DoD, 1989b).

Control Grip Device. The Control Grip Device (CGD) is a portable data input tool that may be installed in control towers when the Tower Control Computer Complex (TCCC) is fielded. A prototype CGD is shown in Figure 7-18. (The prototype CGD does not display the same button arrangement found on CGDs typically found in ATC systems.) The CGD is intended to provide the controller with the functionality of a trackball in a package that will allow freedom of movement in the tower cab. Design issues include weight, shaft width, and button/key positioning.

Figure 7-18. Prototype Control Grip Device
Modified from a photo provided courtesy of Measurement Systems, Inc.

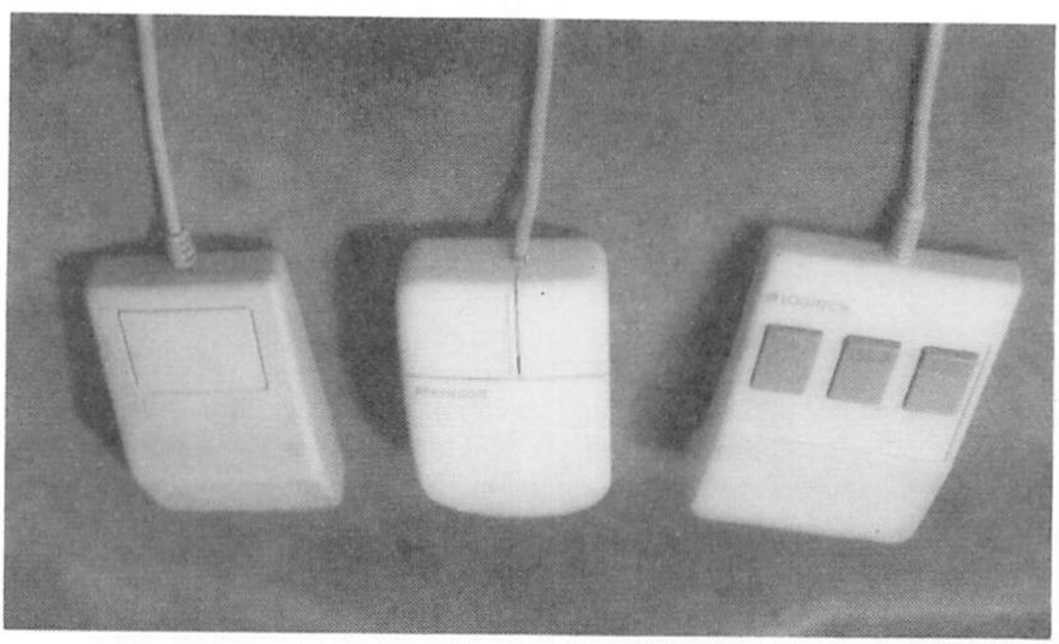


The development of the CGD is an example of task demands influencing the input device and, to a greater extent, a significant portion of the CHI design for an ATC application.

Human factors evaluations have resulted in modifications to the original CGD design. Incorporating the various cursor control mechanisms originally resulted in a device too heavy to be comfortably held by controllers while on position for three to four hours. In addition, the shaft of the earlier CGD was too wide for small-handed controllers to grasp, and its buttons were hard to reach without rotating the CGD. Controllers recommended the redesign of the CGD's size and button arrangement. They also recommended updates to the software, which eliminated the need for some functionality. This resulted in a CGD with acceptable characteristics. The re-designed CGD satisfies the tower controllers' functional needs and meets criteria for usability and workforce acceptance.

Mouse/Mice. Also known as a free-moving XY controller, a mouse is a mechanical, hand-sized, data input device, with one to three buttons on top. Figure 7-19 shows three variations on the mouse concept. A mouse usually has a wire ("tail") connecting it to a computer, although tailless mice are recommended (DoD, 1989b). Movement of the mouse on a flat surface produces cursor movement. Buttons are pressed quickly ("clicked") to perform functions such as changing menus, drawing lines, or confirming input. A mouse button may need to be pressed only once (single click) or twice (double click) for a specific purpose. In some cases, a mouse button is held down while the mouse is moved to "drag" data or graphical objects from one location to another on the screen.

Figure 7-19. Three Versions of the Mouse with One, Two, and Three Buttons



A mouse is an appropriate device for cursor positioning or for data selection. The mouse is fairly easy to learn and use. A mouse is easier to use than other input devices for some tasks such as altering display attributes (e.g. window size). Operational difficulties can arise from frequent switching between the mouse and other input devices (e.g., keyboard). For this reason, the mouse, like other input devices, should be matched to the task for which it was intended. Sources in the human factors literature disagree on whether or not the mouse is suitable for high-precision tasks. Thoroughly understanding the types of data manipulation the user will be performing will help determine properties of the mouse itself, such as the number of input buttons. Controllers who used the mouse in a field study of airspace complexity learned to use it quickly (Mogford, 1990; Murphy and Guttman, 1993). Systematic evaluation is recommended for any mouse design proposed for ATC applications.

Evaluation Guidelines for Mice. The following human-factors issues should be evaluated to assess the usability of mice for ATC applications (DoD, 1989b):

- *Physical Characteristics.* The mouse must have no sharp edges and should be essentially rectangular. Minimum dimensions for width, length, and thickness, given in Figure 7-20, should be met. The mouse should have at least two buttons so that different responses can be associated with different buttons. The mouse should be usable by left- and right-handed controllers.

Figure 7-20. Minimum Dimensions for Mouse Width, Length, and Thickness
From *Human Engineering Design Criteria for Military Systems, Equipment, and Facilities* (MIL-STD-1472D), Department of Defense, 1989, p. 114.

	Minimum	Maximum
Width	40 mm (1.6 in.)	70 mm (2.8 in.)
Length	70 mm (2.8 in.)	120 mm (4.7 in.)
Thickness	23 mm (1.0 in.)	40 mm (1.6 in.)

- *Cursor Positioning.* The controller should be able to orient the cursor to within 10 degrees of the desired orientation without visual reference to the mouse. A mouse should allow the controller to move the cursor easily in any direction without a change in hand grasp, and movement of the cursor on the screen should be smooth. Controllers should be able to maneuver the mouse with either hand. Movement of the mouse from side to side of the work surface should result in similar cursor movement from side to side of the display. If it is possible to drive the cursor off the edge of the display, indicators should be provided to assist the controller in bringing the cursor back onto the display. An appropriate mouse pad must be provided to support positioning operations.

Use of a mouse is not recommended for generation of free-hand graphics (DoD, 1989b).

Issues of operational suitability and workforce acceptance of a free-moving XY control device require early field assessment and involvement of controllers in all phases of testing.

Graphic Tablets. Graphic tablets are flat surfaces that represent an extension of the main visual display. Movement of the finger or a stylus on the tablet provides cursor location information.

A touch-sensitive tablet can sense where it is being touched. Because there is no mechanical intermediary between hand and tablet, this input device can be provided with multi-touching capabilities by placing templates over the tablet to define special regions, which can be touched in parallel (e.g., Lee, Buxton, and Smith, 1985). If used for ATC applications, templates might correspond to parameters such as altitude of aircraft, speed, and direction. The multiple touch sensitive tablet allows for simultaneous adjustment of all the parameters. With respect to a single touch sensitive tablet, the multiple-touch sensitive tablet has longer response time delay as the number of touches increases.

Graphic tablets raise several human factors issues, including the mapping of movement from the finger or stylus to the cursor; steady positioning of the cursor; and the appearance of a continuous stroke if the tablet is used for freehand drawing. Recommendations on these and other technical issues, based on research, are provided in MIL-STD-1472D (DoD, 1989b). For example, the direction of movement of the finger or stylus should produce smooth cursor movement in the same direction; the cursor should remain steady at its point of placement on the tablet until the finger or stylus moves it; and the cursor refresh rate should be high enough to ensure that the user perceives a continuous (not dashed) line in freehand drawing.

If a graphic tablet is used, the user should not be expected to switch frequently to another input device, such as a keyboard. Finding space for a graphics tablet on the worksurface may be a challenge in ATC environments. Controllers should not be required to hold graphic tablets on their laps.

Evaluation Guidelines for Graphic Tablets. If a graphic tablet is being considered for use in an ATC application, the following usability issues should be evaluated:

- *Cursor Movement.* There is a direct correspondence between movement of the finger or stylus and movement of the cursor.

- *Cursor Positioning.* The cursor is maintained in a steady position where it has been placed until further action is taken to move it.

- *Stroke Continuity.* A continuous, solid line is generated when the stylus is used in freehand drawing.

- *Use with Other Input Devices.* It is not be necessary for the controller to switch frequently between the tablet and other input devices.

- *Space Requirements.* Sufficient space is available on the workstation for the graphic tablet to be placed directly to the left or right of the keyboard. If operationally suitable, it should be possible for the keyboard to be stowed and the tablet to be placed directly in front of the controller.

Like other input devices, the graphic tablet should be usable by both left- and right-handed controllers. Any response-time delays must be acceptable to controllers. **Performance measurement and early field evaluation are highly recommended to evaluate the potential positive and negative effects of graphic tablets (and other input devices) on controller performance.** A question for investigation is whether or not a graphic tablet or touchpanel could substitute effectively for the standard keyboard in ATC facilities.

Pushbuttons. Mechanical pushbuttons are single-action controls, which operate in one direction only. Sizes vary, but those employed in the radar operators' workstation are small and finger-controlled. There are various types of push-buttons. The one commonly used by controllers is the "momentary" type, which has "push-on" and "release-off" functions, used by controllers to speak to pilots. The key human factors concerns are pushbutton size, operating force required, and the provision of feedback. Recommended values are given in Figure 7-21.

Evaluation Guidelines for Mechanical Pushbuttons. The dimensions of the fingers that will operate the push button determine its size. Dimensions should fall within the minima and maxima given in Figure 7-21. One study investigated how differing button sizes affected performance on various keyboards arrangements (Deninger, 1960). The findings showed that increasing the dimension of a button from 7.5 to 27.4 mm. produced a reduction in keying times and a reduction in error. Pushbutton dimensions should be evaluated, along with other pushbutton characteristics, by representative controllers.

Figure 7-21. Design Criteria for Mechanical Pushbuttons
From *Human Engineering Design Criteria for Military Systems, Equipment, and Facilities* (MIL-STD-1472D), Department of Defense, 1989, p. 91.

DIMENSIONS (D) / RESISTANCE

	D: Fingertip	D: Thumb or Palm	Resistance: Single Finger	Resistance: Different Fingers	Resistance: Thumb or Palm
Minimum	9.5 mm (3/8 in.)	19 mm (3/4 in.)	2.8 N (10 oz.)	1.4 N (5 oz.)	2.8 N (10 oz.)
Maximum	25 mm (1 in.)		11 N (40 oz.)	5.6 N (20 oz.)	23 N (80 oz.)

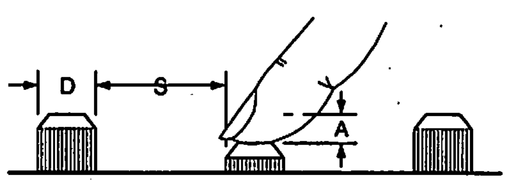
DISPLACEMENT (A)

	Fingertip	Thumb or Palm
Minimum	2 mm (5/64 in.)	3 mm (1/8 in.)
Maximum	6 mm (1/4 in.)	38 mm (1-1/2 in.)

SEPARATION (S)

	Single Finger	Single Finger Sequential	Different Fingers	Thumb or Palm
Minimum	13 mm (1/2 in.)	6 mm (1/4 in.)	6 mm (1/4 in.)	25 mm (1 in.)
Maximum	50 mm (2 in.)	13 mm (1/2 in.)	13 mm (1/2 in.)	150 mm (6 in.)

N = Newton



The second human-factors design factor is the button's required operating force or its key resistance, which serves two purposes:

- Providing the user with tactile and kinesthetic feedback about the extent to which the button has been pressed.
- Helping to guard against accidentally activating the button.

If too much or the wrong type of resistance is used, the operator could experience muscular fatigue.

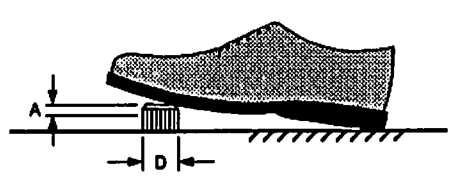
To prevent an accidental slippage on hard pushbuttons, their surfaces should be rough or concave. If the feedback is given by feel or by an audible click, then the eyes can be free for other work.

Foot Switches and Pedals. In place of a hand-operated push-to-talk button, a foot-operated switch or pedal may be used to key the microphone, so that both the controller's hands can be free for other tasks. Usability issues associated with foot switches and pedals include dimensions, resistance, displacement, positioning, and feedback. Figure 7-22 illustrates foot-switch operations and provides standards for their diameter, resistance, and displacement. Foot switches and pedals should also be evaluated for operational suitability and acceptability.

Figure 7-22. Foot Switch Operations
Adapted from *Human Engineering Design Criteria for Military Systems, Equipment, and Facilities* (MIL-STD-1472D). Department of Defense. Figure 12, p. 93.

	DIAMETER (D)	RESISTANCE		DISPLACEMENT (A)		
		Foot Will Not Rest On Control	Foot Will Rest On Control	Normal Operation	Ankle Flexion Only	Total Leg Movement
Minimum	13 mm (0.5 in.)	18 N (4 lb.)	45 N (10 lb.)	13 mm (0.5 in.)	25 mm (1 in.)	25 mm (1 in.)
Maximum		90 N (20 lb.)	90 N (20 lb.)	65 mm (2.5 in.)	65 mm (2.5 in.)	100 mm (4 in.)

mm - millimeter
N - Newton



Evaluation Guidelines for Foot Switches

In addition to meeting the requirements given in Figure 7-21, foot switches should comply with standards for positioning and feedback, as given in MIL-STD-1472D (DoD, 1989b): Foot switches should be positioned so that they are operated by the toe or ball of the foot, not by the heel. A foot switch should be clear of obstructions so that the controller can center the ball of his/her foot on the switch. To aid the controller in locating and operating the switch, a pedal may be placed over the switch button. Positive feedback should be provided to indicate activation of the foot switch. Feedback may take the form of an audible click, a change in the audio display, or another indicator that is acceptable to the workforce. Requiring the controller to operate more than one switch with the same foot is not recommended (DoD, 1989b). Because controllers are likely to resist the introduction of foot switches, early field evaluation and testing is recommended to identify issues of operational suitability and acceptability of foot switches or foot pedals.

Software-generated ("Virtual") Controls. Until fairly recently, there has been a clear distinction between displays and controls. Controllers have configured their displays, switched communication frequencies, and entered data into the computer by means of hard "knobs and dials," such as keyboards, physical switches, and trackballs. In the case of advanced graphical user interfaces (GUIs), the traditional distinction between displays and controls becomes somewhat blurred. GUIs include components that allow the user to enter and manipulate data and to configure frequencies and audio/visual displays by selecting visually displayed, software-generated controls. These controls are also called objects or **widgets.** The set of controls may include various objects or widgets, whose names and appearance will vary depending on the particular software development environment. Many commercial vendors have developed their own sets of these software-generated or "virtual" controls (so-called because they behave virtually like physical controls).

A typical control set is likely to include window controls (e.g., a close box, a size box, a zoom box, scroll bars, scroll arrows) as well as pushbuttons and other controls that allow the user to interact with the application software. Virtual controls are activated via a cursor-positioning and selection device, such as a mouse. In some commercial software development environments, for example, the user activates pushbuttons to provide information needed before a command to the system can be executed. When such information is needed, a control called a dialog box is presented, and the user indicates the options s/he wishes to activate. When a command will display a dialog box for further information, an indication to that effect should be given in the command menu.

Usability issues in the design of virtual controls include their visual appearance when active and inactive, their layout and separation, consistency of usage and layout throughout an application, and the extent to which their functions are self-explanatory.

Evaluation Guidelines for Virtual Controls.

Each commercial software environment has its own set of guidelines for the design and use of virtual or "soft" controls. In general, the visual design of these controls should convey a basic sense of what the control does when it is activated. A virtual pushbutton, for example, should be labeled with text or a pictorial image that indicates what will happen upon activation. **The inactive and active states of a visually-displayed, virtual control should be visually distinct.** For example, an active pushbutton may be filled in with a dark color to make it appear pushed in. Virtual controls that differ in function should be made visually distinct through differences in shape. Use of shadowing to give a three-dimensional appearance must be consistent throughout the application.

In the following discussion, we consider the design of several kinds of virtual pushbuttons.

Togglebuttons are a type of pushbutton that can have only one of two states, on or off. Togglebuttons are typically grouped within dialogue boxes to allow the user to set various parameters that are changed only occasionally. A further distinction is made between *radio buttons* and *check buttons* (also called check boxes). Both are togglebuttons, meaning that they can be only on or off. Togglebuttons become radio buttons if they are grouped and programmed so that only one of the buttons can be on. Radio buttons should be used when an exclusive choice must be made among several options, as in selecting a station on a car radio. Check buttons should be used, however, when more than one option can be active at a time.

Radio buttons and check buttons should be visually distinct. For example, a filled square can be used to indicate the on state of a check button, while an empty square indicates the off state. In contrast, a radio button may be a diamond or a circle, which is filled to indicate the on state and empty to indicate the off state (OSF, 1991). Moving the cursor to the button and clicking activates or de-activates the button.

A consistent set of labels should be used to indicate the choices available to the user. Groups of radio buttons and check buttons should be separated and laid out in such a way that the user is guided logically through the choices that need to be made. When default settings are used for the buttons in dialogue boxes, the default buttons should be highlighted.

A detailed comparison of commercially available virtual control sets, their appearance and operation, is provided by Marcus (1992). **Prior to adopting a particular software package, the design team should evaluate virtual controls to determine their usability, operational suitability, and acceptability for use in ATC applications.** Customized control sets may need to be developed and evaluated. Controllers with various levels of experience and from a cross-section of ATC facilities should participate in these evaluations.

7.4.4 What kind of input device should be used?

Input devices have been distinguished according to how they are operated, in a continuous or discrete fashion. For certain types of tasks, such as text editing, both continuous and discrete devices can be used. The question arises as to which kind is more appropriate for which type of function. For text selection on a CRT, continuous input devices, such as a trackball, mouse or joystick, have proven to be superior to discrete devices, such as text keys (Card, English, and Burr, 1978). The methods used by researchers provide examples of how to conduct comparative evaluations.

An early study of input devices looked at which types of devices were best suited for specific tasks (Card, et al., 1978). The task used to compare discrete versus continuous input devices was text editing, which required participants to precisely identify the target, reach for the input device, acquire the target, and select the target. Four input devices were tested: a mouse, a joystick, step keys, and text keys. (The step keys were the familiar five-key cluster, with a central HOME key surrounded by the keys to move the cursor in each of the four directions. The text keys had functions to move the cursor to one paragraph, line, word, and character.)

Participants in the experiment had never used any of the devices previously. They had to edit text by first directing the cursor onto a highlighted target word or phrase, and then pressing a button to edit the target. This task is quite similar to the editing tasks controllers would perform if they had to update the information contained in the data block associated with each aircraft or information shown on another display.

Two performance variables were measured: speed and accuracy, defined in terms of error rate. The reaction time to perform the task was measured in two parts:

- Interval A: Time needed to detect the target, reach for the device, and start moving the cursor by operating the device. (This interval is called "latency.")

- Interval B: Time needed to reach the target by actually moving the cursor by use of one of the input devices.

The time required to reach the input device and start moving the cursor after detecting the target was longer for participants who used the mouse and the text keys. In the case of the mouse, this longer time reflected the time for the hand to move from the keyboard and reach the mouse.

Study results also showed, however, that the time needed to move to the targets, was always shorter, and the error rate was lower when test participants used the continuous devices. Performance was best with the mouse, followed (in descending order from a speed and accuracy viewpoint) by the joystick, the text keys, and the step keys. The authors of the study recommend the use of the mouse because it outperformed all other devices (Card, et al., 1978).

These results imply that continuous devices (such as trackballs) are most suitable for ATC tasks involving the movement of a cursor on a screen. With a continuous device, the motion of the cursor is more related to the motor movements of the hand than it is when discrete devices are used. Of course, the specific results of this study cannot be directly applied to today's ATC environment because controllers have had extensive experience with the input devices currently in use (i.e., trackball and keyboard). Recall that participants in the 1978 study had never used any of the devices that were tested. Because of controllers' experience, switching from a trackball to a mouse, for example, would not be advisable. Still, the results of the study are valuable and should be considered in the development of future ATC systems.

A later study addressed the issues of compatibility of a pointing device with the requirements of both person and machine (Kley, 1983). This study critically reviewed conclusions of the previous work based on the fact that only a few factors had been taken into account as a basis for recommending the use of the mouse over the other input devices. **A satisfactory comparison of input devices should**

not only focus on speed and accuracy associated with each device, but on several other elements that are important from a human factors standpoint, including positioning speed, resolution capability, space requirements, fatigue, and long-term effects of repetitive movements (e.g., carpal tunnel syndrome).

A thorough evaluation of any input device should include assessment of basic human factors issues and testing of some performance parameters that are valued in ATC operational environments. Because certain human factors issues and performance parameters might not have the same relevance in different contexts, findings from the research literature should be generalized with caution. The nature of the task used, the purpose, and the situation in which controls are employed constitute some of the facts that must be taken into account when evaluating the relevance of any piece of research. Although findings in the literature can suggest what the results may be, there is no substitute for actual usability testing with controllers who are representative of the workforce.

7.4.5 What kind of testing should be conducted to evaluate the design of interaction techniques and devices?

Testing for usability, operational suitability, and acceptance should be conducted throughout the developmental cycle of an interaction technique or device. Early field evaluations are recommended during development (Harwood and Sanford, 1994). Iterative testing is appropriate and necessary to evaluate the capabilities of an interaction technique/device to assist the controller in the performance of ATC duties.

Developmental testing is conducted early in the design phase when modifications are more possible. A "Proof of Concept" test is frequently conducted before more formal testing begins. During developmental testing, techniques and devices are often evaluated against published standards. Compliance and noncompliance are established, and the potential system impacts of noncompliance are identified. When the operational impact is judged to be significant, redesign is often required.

Developmental testing can also be accomplished using early versions of the device, called prototypes. Prototypes allow the users to interact with a reasonable representation of the device in a semi-realistic setting or in the actual setting in which the device is to be used. In this sense, developmental testing can resemble operational testing. The advantage is that the design is still relatively flexible when the price of modifications is not prohibitive.

Traditionally, operational test and evaluation (OT&E) is conducted after a design has proceeded through the development cycle and is ready for fielding. Operational testing has been used to identify additional suitability and usability issues that were missed during developmental testing. Early field testing can, however, reduce the burden on OT&E. Problems identified during iterative field evaluations can be incorporated into future versions of the interaction device or technique. Using this approach, as pioneered by the CTAS program, subtle issues of usability, operational suitability, and acceptance can be identified and resolved well before OT&E (Harwood and Sanford, 1994). A discussion of methods of testing and data analysis can be found in Chapter 10, *Human Factors Testing and Evaluation.*

7.5 CONTROLLER-COMPUTER INTERACTION

Computer workstations are primary features of modern ATC facilities. From a controller-centered perspective, the computer is there to assist the controller's efficient and effective accomplishment of ATC objectives. The controller is the responsible party, and the computer provides a suite of tools. Designing these tools to meet the needs of controllers is a challenge to software designers. Another challenge resides in designing a natural, intuitive interaction (or "dialogue") between the controller and each software tool.

This section presents design principles and evaluation guidelines in three key areas of controller-computer dialogue design:

- *Interactive Control* (also referred to as "Sequence Control"). User commands/actions that initiate, interrupt, or terminate transactions (i.e., computer processing).
- *Data Entry*. User actions that input data to a computer and the computer response to these inputs.
- *User Guidance*. Computer-generated error messages, alarms, prompts, and instructional materials.

The design of these features of the system can have critical consequences for usability, operational suitability, and workforce acceptance.

INTERACTIVE CONTROL

In this section, we define interactive control and review selected approaches to managing the transfer of information between the controller and the computer. We also present basic human factors guidelines for the design of this interactive transfer. Evaluators need to pay close attention to the effects of design choices in this area because they will have critical implications for controller performance.

7.5.1 What are the characteristics of modern approaches to interactive control?

Interactive control or **sequence control** is a two-way process wherein a controller enters commands to the computer and the computer responds by initiating, interrupting, or terminating some transaction. Today's major commercial computer systems are converging on a standard graphical, window-based user system interface design including standardization in the area of interactive control. Typically, to start an **application** (i.e., a specific software program, environment, or software tool), the user selects an **icon**. An icon is a pictorial representation of a physical entity or concept that represents the application. Unless their meaning is unmistakable (as shown by testing), icons should be labeled to clarify their meaning.

Sometimes, applications are initiated by selecting **menu options** rather than icons. (Menu options are the alternative actions listed for user selection.) Once an application is accessed, a window or window set appears containing the information and commands associated with the application. An application window may cover the entire screen area or just part of the screen area.

Application commands are typically generated using application specific menu(s). In addition to activating commands, menu options may be used to access lower level windows for data entry or to specify command parameters into predefined labeled fields.

Three types of menus are **permanent, pull down**, and **pop up**. Permanent menus are always "opened," that is, the menus options are continuously displayed on the screen. Pull down and pop up menus conserve screen area by only showing the available options at the user's request. Pull down menus are activated by selecting a **menu title** that describes the menu. Several menu titles are usually available on a window's or application's **menu bar**. Pop up menus are typically accessed by pressing a special mouse button while the cursor is on a graphical symbol or **hot spot**. (Pop up menus are generally not used with a one button mouse system.)

The lower-level application windows used for data entry or to specify command parameters may contain graphical controls called command buttons to initiate or cancel data processing and command execution. Command buttons are also called **pushbuttons** because of their physical appearance. Pushbuttons may be shaded or colored to appear three dimensional. This coding enables them to show a pushed-in (i.e., activated) versus pushed-out state on the visual display, depending on whether they have been selected or de-selected via an input device.

Frequently, a menu option or pushbutton command can be executed using keyboard short cuts, such as **function keys** or key combinations (e.g., "Return" key to activate the currently highlighted pushbutton or "Alt" and "F4" to close an open window). Before the widespread use of the mouse, function keys were used extensively as the only or primary method to activate menu options.

Another method to execute commands within today's applications is by directly manipulating graphical objects on the screen. For example, to send ATC mail, the controller could place an envelope symbol into a mailbox symbol. Movement and placement of graphical objects is accomplished by using the mouse to place a cursor on the object, selecting the object by pressing a mouse button, and dragging the highlighted object by holding down the mouse button while moving the mouse. This form of interaction, known as **direct manipulation**, gives the user a powerful sense of acting on the displayed environment (Shneiderman, 1992).

More detailed information on interactive control methods and objects (also called **widgets**) is available within general and specific style guides (e.g., Avery and Bowser, 1992; Carlow, 1992; OSF, 1992 Apple Computer, 1987;).

7.5.2 What are the basic human factors guidelines for the design of commands and menus?

Numerous human factors guidelines can be applied to promote simple, intuitive, consistent, error-free command execution. A few key guidelines are presented below with more detailed and comprehensive sets available in many sources (e.g., DoD, 1989b; Gilmore, Gertman, and Blackman, 1989; Hix and Hartson, 1993; Mayhew, 1992; Shneiderman, 1992; and Smith and Mosier, 1986).

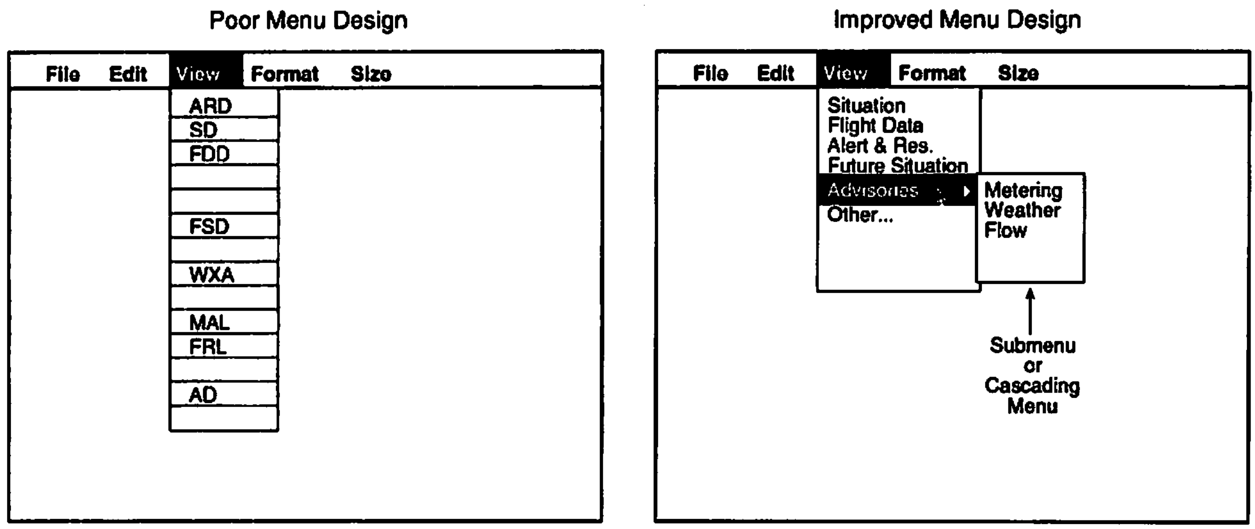
- **The computer should indicate which commands are currently available by displaying the command set relevant to the task at hand.** Often, commands that are applicable to the current task but that currently cannot be executed are displayed but are "grayed-out" (e.g., dimmed so that the menu option is shown in faint grey letters). Graying out is preferred to omitting menu options so that the normal order of the options is preserved.

- **Minimal, simple controller actions should be required to execute commands.** If many parameters must be specified, an easy mechanism should be provided for the controller to enter this information, such as a command parameter entry window with predefined, labeled fields and a list of values to select from for each.

    - **When the controller types in a command or a command parameter, the computer should treat upper- and lower-case letters as equivalent.** Otherwise, the computer sets up a condition that promotes operator error, and users will be frustrated with the computer's refusal to process what they know is a valid command.

    - **Menu option selection can be simplified by modifying the number of menus and the number of menu options and levels within each menu.** If there are very few menu options (e.g., three or fewer), the menu may be merged with another short menu that is functionally similar. If there are more than 10 menu options, the menu may need to be broken down into two menus (unless they are simplistic and related options, such as font type or colors). Hierarchical or **cascading** menus can be used to produce a simplistic, high-level menu that has a few or several sub-options for each menu options. It is difficult for a person to find options or memorize option placement when the menu levels get too deep (i.e., more than four levels). Figure 7-23 shows a hypothetical "bad" menu, and the same menu re-worked to reflect the preceding design rules. Detailed guidance on menu design is available in many sources (e.g., Avery and Bowser, 1992; Hix and Hartson, 1993; Mayhew, 1992; Norman, 1991b; Paap and Roske-Hofstrand, 1988; Shneiderman, 1992).

Figure 7-23. Poor and Improved Menu Design Examples



- **Commands should be consistent in placement, wording, and method of activation.** Use of a style guide promotes a standardized look and method for command activation, e.g., menu options and push buttons. Still, the designer must ensure that wording and usage of commands are consistent throughout. One means to achieve this is by listing each command, a one-line description of its method of activation, and the result it produces. The developer should modify any two commands that perform the same or very similar functions or any one command that performs multiple functions. All commands of a given type should have the same activation method.

- **The computer should indicate the current operational mode.** A **mode** is an internally defined state or condition of operation, such as keyboard input mode, help mode, or save mode. Commands that are accepted in one mode may not be accepted or may cause unexpected actions if used in another mode. The current mode should be clearly and continually indicated since the method and result of command execution could depend on the current operational mode.

- **Command ordering should be consistent from window to window.** When commands, such as menu options or pushbuttons, are displayed in a group, they should follow a logical order, such as typical task sequence or frequency of use. For example, one style guide arranges pushbuttons in the following order for all windows: OK, Apply, Cancel, Help (OSF, 1992). Another style guide may use a somewhat different order. The point is to use the same order consistently to help users develop a natural rhythm in performing tasks.

- **The computer should not tax the controller's memory by requiring entry of a long sequence of command parameters.** The controller should indicate the high-level command and then be provided a fill-in-the-blank window or other simple means to enter parameters. When the number of possible parameter values is limited, the controller should be able to select a value from a list.

- **Command execution should be by explicit controller action: the computer should not take control by activating a command as a by-product of some user action.** When the controller does not execute a probable command, the computer should not take control, but should prompt the controller to perform the action. This example prompts a likely action: "Save changes to flight plan before Exit?" with the options "Yes" to save, "No" to not save, and "Cancel" to not save and not exit.

- **For any type of destructive command (e.g., where data will be lost that would be difficult and time consuming to recover), the computer should explain the consequences of the command execution and have the controller confirm the action.** The confirmation action should be an additional action different from the command initiation action to reduce the possibility that a controller will inadvertently confirm.

- **The controller should receive some type of feedback so that it is clear that the computer has executed a command.** If there is a lengthy delay, a "waiting" indicator is useful, and another indication should be given when processing is complete. Allowing the controller to suspend/interrupt or cancel/undo a transaction that is in progress is recommended as a valuable capability.

Evaluators should systematically assess design choices made in implementing interactive control. These design decisions can have unforeseen critical effects on controller performance. Interactive control design should foster the controller's interaction with the ATC situation, not require excessive attention to managing the computer.

DATA-ENTRY DIALOGUE

In the area of person-to-person communication, people generally refer to *dialogue* as the interchange of information between speakers. During a simple dialogue, speakers take turns speaking and listening. In the realm of the ATC system, data entry is the controller's side of the "dialogue" with the computer. The controller communicates his or her intentions to the computer by means of the input devices provided for this purpose. In response, the computer first "echoes" the entered data by displaying a readable representation of the input and then acts upon the input.

Designing the form and format of this controller-computer dialogue is the topic of this section. We present human factors principles to guide dialogue design and evaluation. We also sample standard methods used to support effective, efficient human-computer dialogues.

7.5.3 What human factors principles apply to designing data entry dialogues?

Depending on the system, the required amount of data input to the computer could be extensive or minimal. For ATC, the controller needs to focus on monitoring situation displays, communicating with pilots, and making data entries. Therefore, the methods for data entry should be simple, quick, intuitive, and not require special skills, such as touch typing.

Before we discuss specific data entry forms and methods, let's consider general data entry principles based on DoD, 1986:

- **The controller should have to enter a data item only once.** Once an item is entered, the computer should be able to insert that value wherever else it may be needed. Repeated entries of the same information should never be required. Once the controller has entered the assigned altitude for an aircraft, for example, the system should be able to insert that value in other places, as needed.

- **The computer should perform all necessary calculations from entered data and fill in derived data fields.** Let's consider a situation in which this principle was violated. Every 20 seconds, operators of a NASA satellite control system received the total number of telemetry blocks (data being downloaded from the spacecraft). Since operators were interested in the rate of telemetry download, it was necessary for them to subtract the old from the new number of telemetry blocks being received by several pieces of satellite communications equipment for each currently transmitting satellite (Mitchell and Saisi, 1987). Controllers should never be required to transform data from one unit to another; nor should they be required to perform any mental calculations. **All data should be presented in directly usable form.**

- **The controller must receive feedback by viewing each key stroke as it is immediately echoed on the screen.** This is especially important for user populations, like controllers, who may not be touch typists. An exception to this principle is for private or secured entries, such as a log-in password.

- **Once data are entered, the controller should receive feedback on data acceptance or rejection.** The computer can provide feedback on a field by field basis or can wait until the controller submits an entire data set for processing. (In the latter case, all erroneous fields, not just the first, should be indicated.)

- **The computer should not erase an erroneous entry since the user needs to see why the entry is incorrect and may only need to fix a piece of the entry.** The computer should not interrupt an in-progress data item entry, but should wait at least until the controller leaves the current field and enters a new field. Users find it annoying to be corrected before they are given the opportunity to review and fix a field themselves.

- **Data entry should be user paced.** The computer should not impose time limits or time outs.

- **The computer should not restrict the order in which data items are entered.** The controller should be able to defer required data and at some later time have the computer prompt him/her to provide all deferred values.

- **Data processing should be initiated only after an explicit command from the controller.** The controller should be able to cancel data modifications and have fields revert to their previous values. The controller should be able to "undo" the new or modified value entered in an individual field (i.e., the current field) and have it revert to the initial value.

- **The computer should present a special cursor (usually an "I" beam) to indicate that the controller is in keyboard data entry mode; i.e., the user can type an entry into the current field using the keyboard.** Other data entry modes should each have a distinctive cursor that by shape, color, blinking, or other means can be easily located on the screen. There should be minimal use of multiple cursors available simultaneously, and there should be a simple means to switch among using cursors.

An example of multiple cursors is the data entry "I" beam for data entry in the current field of the active window, and an arrow cursor used to select another field in the window, activate a new window, select a graphics symbol, or initiate other action.

- **Initially, the keyboard data entry cursor should appear in the first data entry field or location.** The assumption should be made that data entry will begin at the beginning to relieve the user of having to position the cursor.

- **In addition to typing in an entry the controller should be able to select an entry from a menu (i.e., list of options).** To conserve screen space, the list associated with a field can pop up at the controller's request and disappear once a selection is made.

- **The data entry method should be designed to minimize errors and enable quick, simple data editing/correction.** Where possible, the system should provide default values that the controller can confirm, modify, or replace. Correction of entries should always be simple (e.g., with an "undo" function).

- **The method for data change should be the same as or consistent with the method initially used to enter data.**

- **The controller should be able to edit all or part of a data field.**

- **The computer should ignore excess spaces and decimal points following integers. The controller should not have to enter leading zeros for numeric entries.**

- **There should be limited requirements for controller-entered delimiters.** When delimiters are needed, for example to partition long entries, the computer should present the required format and punctuation, which should be consistent with common ATC practice (e.g., __/__/__ (mm/dd/yy)).

- **Every input by the controller should produce some perceptible response from the computer.** Frequent or common actions should produce fairly subtle responses and critical actions should produce more monumental responses. Absence of computer response is not an acceptable means of signaling acceptable entry.

These principles should be followed no matter what specific method is used to support the controller-computer dialogue.

GENERIC DESIGN OBJECTIVES IN DEALING WITH HUMAN ERROR

- Exclusion of Error: The design makes it impossible for the user to commit particular errors.
- Increased Difficulty of Error: The design makes it difficult, but not impossible, for the user to commit particular errors.
- Reduced Consequence of Error (Error Tolerance): The design reduces the severity of the consequences of error without affecting the probability of errors.

After Sanders and McCormick, 1993, p. 660

7.5.4 What standard methods of data entry support the human-computer dialogue?

The simplest data entry method is **free text entry** into a large field (e.g., entering a data link message to be transmitted to a pilot). Generally, little or no data entry restrictions are placed on such a field, and little data validation is performed by the computer. The user may be able to execute commands on such text (e.g., check spelling or search for a key word).

A **fill-in-the-blank form** allows the user to enter data into predefined, labeled fields. **Labels should be used consistently within and throughout all forms and should use specific, user-oriented wording** (e.g., accepted ATC terminology). Data entry items should be logically ordered and arranged according to user expectation.

When the same set of data must be entered for several entities (e.g., aircraft flight information for all current traffic), a **table** is often the best data entry form. Labels are required for both columns and rows. The first column should contain the most salient discriminating feature that identifies and distinguishes the row (e.g., flight number or aircraft identifier). There should be row and column scanning aids, (e.g., on every third row/column a blank or shaded row/column). If the user needs to run a finger down the table or hold a ruler up to it, this is indicative of poor table design. The table (and other forms) used for data entry should be compatible with, if not identical to, display data on tabular and other displays.

In general, another sign of poor design is the "sticky" note posted in or around the workstation area. If controllers need to post notes, the design has failed to capture and provide all necessary information.

Data entry may also be achieved via **direct manipulation** of graphic objects on the screen (e.g., Shneiderman, 1992). For example, to enter a revised flight route for the pilot to fly, the controller could select points on a map (e.g., by pointing with the cursor/mouse movement and selecting by pressing and releasing the mouse button) or could "bend" the pilot's currently planned flight path (e.g., by selecting a point on the route and pressing and holding the mouse button while moving the mouse).

For cluttered screens, where many selectable graphic objects are in close proximity, it helps to provide users the ability to **zoom** (or enlarge) a portion of the display for precise pointing. Usually, the selectable objects remain the same size while other information (e.g., a background map) is magnified so that selectable objects appear farther apart.

Once selected, a graphical object should be highlighted, so that the user knows the next action will be applied to it. For certain transactions, the user should be able to select a set of objects (e.g., by holding the Shift key and clicking on several objects or drawing a box around a grouping or specifying a characteristic that all objects in a group share). Once the grouping is specified, the user can initiate a transaction on all group members. For example, the controller may want to transmit the same message to several aircraft.

Because direct manipulation interfaces support a highly natural flow of communication between the user and the computer, they can be considered "intelligent" interfaces (Chignell, Hancock, and Loewenthal, 1989). We discuss intelligent interfaces in more detail under Section, 7.6 Future CHI Issues.

USER GUIDANCE

User guidance includes alarms, prompts, labels, error messages, and instructions provided to guide the user's interaction with the computer. Effective user guidance contributes to meeting several of the CHI design's general objectives (Smith and Mosier, 1986):

- Consistency of operational procedures
- Efficient use of full system capabilities
- Limited memory load on the user
- Reduced learning time
- Flexibility in supporting different users.

In this section, we consider the formulation of error messages to address the kinds of errors typically made by novice and expert users. Error messages are just one kind of guidance message (i.e., a message from the computer to the user about their interaction, not about the external situation). We provide recommendations on the wording and physical construction of guidance messages in general. We discuss the appropriate uses for guidance messages and methods for evaluating their effectiveness.

Because it is often difficult for a software developer to create messages that are clear to the user community, we recommend that a team including a software developer, human factors specialist, and controller representative collaborate on guidance message development. Candidate messages should be evaluated by a wider controller audience.

Guidance messages should provide any required context information (e.g., a history of previous entries). Both on-line and hard copy forms of a dictionary of abbreviations and legends for all symbols and codes should be provided. In addition to on-line help, the controller should have access to a complete list of error messages with detailed explanation in hard copy form (e.g., within the user manual).

7.5.5 What kinds of error messages should be provided to address both novice and expert errors?

Typically, users commit two types of errors: expert error and novice error. **Expert errors** or "slips" occur when the user knew what should have been done but inadvertently performed the wrong action. This can occur without direct awareness and is not likely to be remembered. In such a case, the user needs the computer to report *what s/he did.* **Novice errors** or "mistakes" occur when the user knew what s/he did, but thought that the action sequence was correct.[1] In such an instance, the intention is flawed and the user needs to know *what s/he should have done.* Any user is likely to commit both types of errors, although with experience the novice errors will be less prevalent. To address both types of errors, **messages should summarize the problem and propose a solution.** Expert users may choose to downsize an error message window so that only the problem statement is in view. **When an error is detected, the computer should display an error message stating both what is wrong and what can be done. When the set of acceptable responses is small, these should be provided.**

To address the needs of both the novice and the expert as well as an individual's specific needs, **the computer should support multiple levels of error messages.** There are many ways to accomplish this. One possibility is to provide a controller-defined capability to set message level as abbreviated or detailed. Another is to provide an error message window that can be sized so that only the beginning of a message—which should be the most informative part—is shown. A third is to provide a capability with each message to expand the text with explanation, examples, and perhaps illustration. Another is to tie messages closely to a help application (e.g., an on-line user manual where the controller can click on "hotspots" to jump to various topics and/or tutorials) so that the controller can quickly transfer into the appropriate area of help. For example, each message window can include a "Help" pushbutton through which the controller can access specific information associated with the current problem.

For training problems, the initial computer setup should be geared towards the developmental controller (i.e., provide extensive feedback). By the time a trainee is expert enough to be annoyed by excessive messages, s/he should know how to set preferences that tailor the displays in order to limit the number and content of computer messages.

7.5.6 How should guidance messages be worded?

Always written in the active voice, error messages should be as specific as possible while, at the same time, minimizing wordiness (Smith and Mosier, 1986). Error messages that are too generic (e.g., "Invalid Entry" or "Invalid Drive Path Specification") or too obscure (e.g., Error #E211A07) should be avoided. **Error messages should describe the problem and the recommended solution(s) in very specific terms, that is, they should tie the problem to a process, object, action, file, data entry field, or other data element.** Terms used in error messages should be task oriented and familiar to controllers. **Error messages should be unique, that is, only one possible message for each error, and should not be redundant, that is, each error generates only one message.**

Guidance messages should promote the idea of the computer as a tool, not a person, and should affirm the controller's sense of being in charge (Smith and Mosier, 1986). Wording of feedback messages should reflect fact rather than make value judgments (e.g., instead of "Flight Conditions =

Excellent," state "Visibility = 50 miles"). Prompts should indicate a ready mode (e.g., an I beam for data entry) rather than telling the controller what to do (e.g., ENTER DATA>). The computer can prompt the controller for a logical action (e.g., "Save training session 'Alpha' before Exit?"), but should never automatically perform a task or switch the controller into or out of a mode without being explicitly instructed by the controller to do so.

Since many of the controllers who will use the ATC system may be unfamiliar with computers or may be accustomed to modern day "friendly" applications, computer messages should be constructive phrases written in a positive tone. Messages should be worded in the affirmative, not the negative. They should instruct the controller on what can be done, rather than what is wrong or not possible. Any blame should be placed on the computer, not the user; e.g., a value should be described as "Unrecognized" or "Not within range" rather than "Illegal" or "Invalid."

Message wording and tone are key components in error message development (Smith and Mosier, 1986). Historically, computerized messages were written in condemning and extreme language. It was not uncommon to get a "FATAL" or even "CATASTROPHIC" error or have a "RUN ABORTED" or "TERMINATED." (Messages were often in all capitals to further emphasize the severity of the error.) User actions and entries were deemed "Bad," "Illegal," in "Error," or "Invalid." This approach can be unsettling to a person not familiar with the programming world in which operating systems and compilers, even today, commonly display harsh messages. Wording should reflect the controller's understanding of the system and avoid any codes, references to computer user manuals, or other information that may not be readily available.

7.5.7 How should guidance messages be physically constructed and presented?

In general, users prefer mixed upper and lower case. People can read mixed case messages more quickly than messages in all capitals, and mixed case makes error messages appear less harsh. Messages should preserve consistency in grammatical form, terminology, phrasing, abbreviations, punctuation, and other grammatical elements (Smith and Mosier, 1986). Visual format and placement should also be consistent. Error messages may be presented in a pop-up message window, which the controller should be able to quickly dismiss, for example, by pressing a confirmation pushbutton, or perhaps by clicking anywhere on the screen. Error messages can also be presented as a closeable message-log window, which should be displayed upon log-on unless otherwise specified by the controller, or via an audible beep.

Audible "beeps" provide one alternative to textual error messages. These are most appropriate for minor errors committed by expert users, who need to be informed of the problem's occurrence, but not given the solution. The sound of the audible tone should be subtle and pleasant. Since users often find an audible cue to be embarrassing or annoying, they should be able to turn off the tone or replace it with a visual cue, for example, very brief dimming of the screen (Smith and Mosier, 1986). If the controller repeats the error two or more times successively, a textual message should be provided.

Error messages should be presented immediately after the error's occurrence. In order not to disrupt the user's task, the computer should allow the controller to finish entering data in a field before the computer validates the entry; e.g., the user indicates completion by moving to another field (Smith and Mosier, 1986). A special key (e.g., a function key or the "Return" key) can be used to allow the controller to move quickly to the next erroneous field using the keyboard. Guidance messages should present a series of actions in the sequence that the controller would normally follow.

7.5.8 In what situations should computer messages be provided?

The following are some areas where computer messages or feedback should be provided (Smith and Mosier, 1986):

- Processing delay: acknowledge receipt of request, specify process and length of delay, indicate percent of processing complete, and provide feedback when processing is complete.

- Limits not met or exceeded: specify appropriate range.
- Unrecognized command action (e.g., mouse click) or data entry value.
- Status information, e.g., current mode.
- Prompting for logical action, e.g., save before exit.
- Confirmation of destructive entry, especially if the action is irreversible (e.g., delete).
- Cautionary messages or highlighting for questionable entries.

To maintain consistency and promote ease of error message evaluation, a single file of computer messages should be created. This allows software developers to access a common set of computer messages or to create new ones using templates provided in this file, e.g., "[task type] processing will be complete in [time value] [minutes or seconds unit of measure]."

7.5.9 What methods are recommended for evaluating computer messages?

The complete set of computer messages should be reviewed by controllers who are only minimally familiar with the new system (e.g., not those who have been involved in system development). Evaluation and redesign of messages should occur where appropriate to eliminate uncommon or unclear words, phrases, abbreviations, and acronyms. We suggest the following evaluation methods:

- Show an error message and ask a controller to describe the problem that caused it and the solution to remedy it. This will determine whether the message is intuitive with explicit remedies or needs to be changed to make it more readily understood.
- Describe a problem with several potential error messages and have the controller select the best one.

- Present a problem and an error message and have the controller note any aspect of the message that is confusing, superfluous or incomplete. Once error messages are implemented on the ATC computer, a performance data collection routine can be implemented to run in the background. This routine could record the frequency of each error message. Any message that occurs frequently in a session or across controllers points to user-interface issues that need redesign.

Guidance messages should provide any required context information, e.g., a history of previous entries. Users should have access to an on-line and hard copy dictionary of abbreviations and legends for all symbols and codes.

7.6 FUTURE CHI ISSUES

Technology is evolving at a rapid pace. The products that are being tested today had their designs frozen many years ago. The tremendous lead time can sometimes result in products that are outdated before they reach their intended environments. This means that technologies emerging today may not be seen for several years. What may seem like science fiction today may well be reality tomorrow. Some of the high technology likely to find its way into ARTCCs, TRACONs, and ATCTs includes electronic flight strips, expert systems, intelligent and adaptive user interfaces, hypermedia, multimedia, and virtual reality. These technologies are discussed below with emphasis on human factors issues.

7.6.1 How can changes in technology affect display characteristics of a new or enhanced system?

Display technology changes may permit useful changes in symbology selection even for tasks and information requirements that remain the same. Such would be the case when the change involves a new technology such as graphics capabilities for creating symbology. Advanced graphics capabilities may permit the display of data in ways that promote visualization of the information. Helping controllers visualize data may reduce the cognitive demand of interpreting alphanumeric data. For example, representing altitude data graphically may speed information processing by relieving the controller of the cognitive task of comparing the current value (220) with the previously displayed value (240) to reach a conclusion regarding status. Arrow icons (level, upward and downward pointing) can be very helpful.

7.6.2 What are the key human factors issues associated with electronic flight strips?

Paper flight strips measure approximately one inch by eight inches that includes aircraft and flight information used by controllers in ATC facilities. Flight strips can be written on, cocked, sorted, and held. They represent an efficient aid to controlling air traffic. Paper flight strips, however, require a great deal of manual handling and a great deal of real estate in an ATC facility.

On the not-too-distant horizon, pending testing outcomes, electronic flight strips, called Flight Data Entries (FDEs), will display flight strip information on a video monitor and will extend flight strip functionality to include prompting, indications of status change, and updating without human intervention. The final format of FDEs has not yet been determined. They will be electronically transferred, not passed manually, between controllers. With FDEs, it will be possible to tailor the information content, based upon stage of flight, to the needs of the controller.

Several suitability issues remain to be worked out by designers and controllers. A major concern is ensuring that subtle information cues provided by today's paper flight strips are not lost when strips are converted to an electronic medium (Hopkin, 1991). Paper flight strips and strip boards possess qualities which may not be easily replicated on a video screen. A strip board portrays relationships between aircraft as well as the pattern of traffic under the jurisdiction of a controller in an open form that can be viewed by others. As such, a flight strip board is an indicator of workload that can aid supervisors in staffing. Because markings are not erased from a flight strip, each strip represents a complete flight history on one comprehensive and observable record.

Ongoing research at the FAA's Civil Aeromedical Institute (CAMI) is investigating the cognitive functions of flight strips so that necessary functions can be identified and retained in any approach to providing timely flight data information to controllers and to the computer (Vortac, 1991; Vortac, Edwards, Jones, Manning, and Rotter, 1993; Vortac, Edwards, Fuller, and Manning, 1993).

7.6.3 How are developments in artificial intelligence likely to find their way into ATC systems?

Artificial Intelligence (AI) purports to capture the knowledge or intelligence of human experts and make this expertise available to the operator of the system. AI or "expert" systems are particularly good at performing complex data searches and comparisons and can complete such tasks far faster more accurately than humans can. Some AI systems are even capable of rudimentary learning and of altering their behavior accordingly. Well-known AI programs include artificial chess players that allow a person to play against a computer and medical diagnostics in which medical personnel can input data on a patient (e.g., medical history, symptoms, and test results) and receive information on a probable diagnosis or further tests that should be conducted.

AI systems are limited, however (see Dreyfus, 1972). Since these systems are based on rules and data, they are only as good as the rules and supporting information. Another problem for expert systems is the extent to which the subtleties and nuances of human knowledge can actually be captured and encoded. In the case of an expert system for ATC, for example, it is not currently possible to encode everything a controller might know about a dynamic ATC situation and the interactions of many changing ATC variables. A variable not known to the expert system might make a world of difference to the controller. For the foreseeable future, any ATC expert system should be considered an assistant, which offers advice within the constraints of its limited knowledge.

The most likely application of AI in the ATC domain is decision-aiding. AI programs could provide controllers and traffic management personnel with "look-ahead" predictions of traffic and allow them to see "what if..." by simulating the results of possible strategies. Another possibility, computer-aided conflict resolution, is a concept that has been discussed and prototyped for over 15 years (Ball, Lloyd, and Ord, 1976; Ball and Ord, 1983; Kulik and Welles, 1990; Cardosi, et.al., 1992). While computer-generated resolutions could be generated from logical rules alone, resolutions developed without the information and strategies used by expert controllers may not be readily accepted (and used) by controllers, even though they would effectively solve the potential conflict. The algorithms that generate the resolutions, should not only be based on separation standards, the laws of physics, and physical data (e.g., aircraft performance characteristics, altitude, and winds). They should also incorporate the formal and informal rules used by controllers (such as, avoid assigning altitudes that would be wrong for the direction of flight or other altitudes that are not routinely assigned; turn a slower aircraft behind, rather than in front of, a faster aircraft, etc.).

The many serious human factors issues associated with expert systems for ATC are identified and discussed in a highly recommended source (Hopkin, 1988). Among these issues are the need for controllers to evaluate the resolutions or other suggestions offered by an expert system, rather than simply accepting and implementing the highest-ranked suggestion; the effects of increased automated aiding on controllers' decision-making skills; and the implications for training and ATC teamwork. (See Chapter 6, Issues in ATC Automation, for a more detailed discussion of the implications of automation.)

The concept of an intelligent, adaptive interface is another AI development that might be applicable in future ATC environments. Research on intelligent interfaces uses AI techniques for the purpose of improving user interfaces to highly automated systems that are capable of solving complex problems on their own (Miller, Sullivan, and Tyler, 1991). An intelligent, adaptive interface is one that can adjust its selection and presentation of information depending on what it knows about the user. Intelligent tutoring systems, for example, are designed to give additional instruction in students' areas of weakness. (See Hancock and Chignell, 1989, for further discussion of intelligent interfaces.)

The full range of human factors issues applies to intelligent, adaptive interfaces. These issues can be summarized in terms of three problems (Card, 1984): the role problem, the automation problem, and the communication problem. All of these problems make intelligent interfaces extremely difficult to design.

The role problem has to do with the relationship between the user and the AI components of the system: Is the AI component the user's partner? If so, which partner makes the decisions? Does the AI component present itself as a tool, as an environment, or as an artificial person with a recognizable style? Designing a role for intelligent components requires attention to issues of operational suitability and workforce acceptance. Decisions in this area should not be made without seeking operational input or without evaluating alternatives.

The automation problem is essentially the function-allocation problem revisited. In its original form the question was, "Which tasks should be allocated to the user and which to the computer?" Getting beyond this form of the question, researchers are asking, "Can tasks and functions be shared in a way that does not require predefined allocation of functions? How should the user and the computer cooperate?" Cooperation requires communication.

The communication problem involves finding a framework for user-computer communication that is "open, but natural and predictable" (Card, 1984). If intelligent, adaptive interfaces are considered for ATC, these will be some of the major challenges to designers and evaluators.

7.6.4 What are some of the advanced concepts of presenting information that may have potential for application in future ATC systems?

Three advanced concepts for enhancing the user's experience of the information presented are **hypermedia, virtual reality** and **multimedia.** As compared to current approaches, all of these technologies permit wider exploration and deeper cognitive processing of available information.

Hypertext and Hypermedia. The idea behind hypertext and hypermedia is linkage of related information (e.g., Shneiderman, 1992). In hypertext systems, the user can go from one place to another in an on-line document by following a set of software links between items that may be widely separated in the actual document. For example, if this handbook were an on-line hypertext document, you (the reader/user) could select a cross-reference that refers you to another part of the document, and go there immediately.

Hypertext systems are generally structured in a hierarchical fashion so that users can find more detailed information on available topics. Selecting a highlighted item, such as *virtual reality*, takes the user to other information on that topic. In hypermedia systems, the concept of linkage is extended to include video and sound as well as text and graphics (Shneiderman, 1992).

There are nearly unlimited possibilities for hypertext and hypermedia applications in air traffic control. Imagine, for example, a capability to select an aircraft target symbol and have detailed information appear about that aircraft, beyond what is already coded in the data block. Linkages between aircraft, routes, weather, flow restrictions, procedures, and so forth could be made more readily available to the controller than they are today, easing some of the burden on long-term and short-term memory. The ATC handbook, letters of agreement, and other documents could be available on-line in hypertext formats. [ex. of aircraft maintenance manual] Designers of hypermedia systems must, however, beware of creating too many links, which can overwhelm the user (Shneiderman, 1992).

A key usability issue in hypertext and hypermedia is ease of navigation through the hyperstructure. Users can easily get disoriented and forget where they have and have not been in the structure (i.e., they can become lost in "hyperspace"). This is similar to the loss of orientation that can occur in a large, hierarchical menu system. Aids to navigation should be provided to help the user maintain a clear sense of where s/he is in the structure and where s/he has been. In addition to the navigation issue, designers and evaluators must consider issues of content and interaction (Shneiderman, 1992).

Virtual reality. Virtual technologies create artificial but highly realistic computer-generated worlds of sight, sound, and touch (e.g., Smith, 1992). They permit interaction with a 3D world which may or may not yet exist. For example, virtual technologies could be used to design and interact with a prototype ATC workstation before one is even built. Virtual reality is being explored by NASA and the U.S. Air Force for applications in space exploration and fighter cockpit design, respectively (Furness, 1986). Virtual reality can be used by pilots to "fly" their proposed route prior to takeoff. Similarly, controllers might find it useful to preview upcoming traffic dynamics by means of virtual reality.

While the pace of technology is rapid, a controller's capabilities, being human, are likely to remain relatively stable. For this reason, it is important that tomorrow's advanced technology be evaluated against today's capabilities. In the end, it will still come down to whether the controller can use the tool effectively and consistently to complete a task. Displays will still have to be evaluated for their usability, operational suitability, and acceptability. For example, if a controller envisions aircraft positions in two dimensions (latitude and longitude) or in three dimensions (i.e., altitude, direction, and airspeed), a 4D display may be inappropriate for an ATC application. (It cannot be assumed, however, that all controllers build and maintain 2D or 3D images that map directly to the external ATC situation.) Virtual reality may become a useful tool in ATC training because it will allow the trainee to experience, explore, and manipulate the virtual ATC environment in ways that strongly reinforce classroom training.

Multimedia. Multimedia software interacts with its users via two or more of the following media: text, graphics, animation, video, or sound. An application including only text and still graphics, however, is not usually considered to be multimedia. NASA Ames and the FAA are investigating multimedia interfaces for tower and cockpit applications (Blattner and Dannenberg, 1992).

Complex landing patterns such as triple parallel approaches require the controller to integrate high volumes of air traffic into an already congested space. To aid controllers in this task, acoustic displays are being proposed in which controllers hear communications from incoming traffic in positions that correspond to the actual location of the aircraft in the terminal area. Such a display may make potential collisions more obvious to the controller as air/ground communications would be heard in their true spatial locations and their routes could be tracked. Another multimedia application involves ATC alerting systems, whereby an auditory icon could be used to warn of potential incidents such as runway incursions. As with the first example, the auditory icon could be processed to convey directional information. (Auditory icons are sometimes called "earcons.")

Future CHI technologies can benefit the controller if adequate attention is paid to human factors principles and adequate testing is conducted. Compensating for poor CHI design should **never** become the responsibility of training departments or procedures developers. Controllers should not have to develop work-arounds to compensate for design flaws or oversights.

A usable, operationally suitable, and acceptable CHI design can go a long way toward mitigating controller workload and performance issues.

7.7 REFERENCES AND SUGGESTIONS FOR FURTHER READING

Adelman, L., Cohen, M. S., Bresnick, T. A., Chinnis, J. O., Jr., and Laskey, K. B. (1993). Real-time expert system interfaces, cognitive processes, and task performance: An empirical assessment. *Human Factors, 35*, 243-262.

American National Standards Institute (ANSI). (1988). *American national standard for human factors engineering of visual display terminal workstations* (ANSI/HFS 100-1988). Santa Monica, CA: Human Factors Society.

Ammerman, H. L., Claussen, C. A., Inman, E. E., and Tobey, W. K. (1987). *FAA air traffic control operations concepts* (Vol. III: ISSS en route controllers, DOT/FAA/AP-87-01). Washington, DC: Department of Transportation/Federal Aviation Administration.

Apple Computer, Inc. (1987). *Human interface guidelines: The Apple desktop interface.* New York: Addison-Wesley.

Aucella, A. F. (1989). *Design guidelines for graphical user interfaces* (Human Factors Society tutorial). Hamilton, MA: AFA Design Consultants.

Avery, L. W., and Bowser, S. E. (1992). *Department of Defense human-computer interface style guide* (Version 2.0). Fort Lewis, WA: U. S. Department of Energy, Pacific Northwest Laboratory.

Avery, L. W., Benel, D. C. R., and Perkins, J. C. (1985). *U. S. Army Test and Evaluation Command test operations procedure soldier-computer interface* (DRSTE-RP-702-100 Test Operations Procedure 1-1-059). Alexandria, VA: Essex.

Baber, C., Usher, D. M., Stammers, R. B., and Taylor, R. G. (1992). Feedback requirements for automatic speech recognition in the process control room. *International Journal of Man-Machine Studies, 37*, 703-719.

Ball, R. G., Lloyd, R. B., and Ord, G. (1976). Interactive conflict resolution in air traffic control. *Proceedings of the NATO AGARD Conference on Plans and Developments for Air Traffic Systems* (No. 188). Cambridge, MA: NATO AGARD.

Ball, R. G., and Ord, G. (1983). Interactive conflict resolution in air traffic control. In M. E. Sime and M. J. Coombs (Eds.), *Designing for human-computer interaction* (pp. 261-283). New York: Academic Press.

Beaton, R., and Weiman, N. (1984). *Effects of touch key size and separation on menu-selection accuracy* (Technical Report No. TR 500-01). Beaverton, OR: Tektronix, Human Factors Research Laboratory.

Benbasat, I., and Todd, P. (1993). An experimental investigation of interface design alternatives: icon vs. text and direct manipulation vs. menus. *International Journal of Man-Machine Studies, 38*, 369-402.

Berson, B.L., Po-Chedley, D.A., Boucek, G.P., Hanson, D.C., Leffler, M.F., and Wasson, R.L. (1981). *Aircraft alerting systems standardization study* (Vol. II - Aircraft alerting system design guidelines, DOT/FAA/RD-81/38). Washington, DC: Department of Transportation/Federal Aviation Administration.

Billingsley, P. A. (1988). Taking panes: Issues in the design of windowing systems. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 413-436). New York: Elsevier-North Holland.

Blattner, M. M., and Dannenberg, R. B. (Eds.) (1992). *Multimedia interface design.* New York: Addison-Wesley.

Bodker, S. (1991). *Through the interface: A Human activity approach to user interface design.* Hillsdale, NJ: Erlbaum.

Boff, K. R., Lincoln, J. E. (Eds.) (1988). *Engineering data compendium.* Wright-Patterson Air Force Base, OH: Harry G. Armstrong Aerospace Medical Research Laboratory.

Brown, C. M. (1988). *Human-computer interface design guidelines.* Norwood, NJ: Ablex.

Brown, J. R., and Cunningham, S. (1989). *Programming the user interface: Principles and examples.* New York: Wiley.

Bullinger, H.J., Kern, P., and Muntzinger, W. F. (1987). Design of controls. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 577-600). New York: John Wiley.

Burnett, M. S., and Barfield, W. (1991). Perspective versus plan view air traffic control displays: Survey and empirical results. *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 87-91). Santa Monica, CA: Human Factors Society.

Card, S. K. (1984). Human factors and the intelligent interface. In G. Kohl and S. J. Nassau (Eds.), *Combining human and artificial intelligence: A new frontier in human factors* (Symposium proceedings). New York: Metropolitan Chapter of the Human Factors Society.

Card, S. K., English, W. K., and Burr, B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics, 21*, 601-613.

Card, S. K., Moran, T. P, and Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Erlbaum.

Card, S.K., Pavel, M., and Farrell, J.E. (1984). Window-based computer dialogues. *Proceedings of INTERACT '84, First IFIP Conference on Human-Computer Interaction* (pp. 355-359).

Carlow International Incorporated. (1992). *Human-computer interface guidelines* (DSTL-92-007). Greenbelt, MD: NASA-Goddard Space Flight Center.

Carroll, J. M., Mack, R. L., and Kellogg, W. A. (1988). Interface metaphors and user interface design. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 67-85). New York: Elsevier-North Holland.

Carter, L. E. (1991). *Usability: Improving the human interface* (Unpublished manuscript).

Carter, R. C., and Cahill, M. C. (1979). Regression models of search time for color-coded information displays, *Human Factors, 21*, 293-302.

Chapanis, A. (1982). (Unpublished keynote speech). *Human factors in computing systems.* Gaithersburg, MD.

Chignell, M. H., Hancock, P. A., and Loewenthal, A. (1989). An introduction to intelligent interfaces. In P. A. Hancock and M. H. Chignell (Eds.), *Intelligent interfaces: Theory, research and design* (pp. 1-26). New York: Elsevier/North Holland.

Chignell, M. H., and Waterworth, J. A. (1991). WIMPs and NERDs: An extended view of the user interface. *SIGCHI Bulletin, 23(2)*, 15-21.

Christ, R. E. (1975). Review and analysis of color coding research for visual displays, *Human Factors, 17*, 542-570.

Coleman, W. D., Williges, R. C., and Wixon, D. (1985). Collecting detailed user evaluations of software interfaces. *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 240-244). Santa Monica, CA: Human Factors Society.

Colquhoun, W. P. (1975). Evaluation of auditory, visual, and dual-mode displays for prolonged sonar monitoring in repeated sessions. *Human Factors, 17*, 425-437.

Coombs, M. J. (1986). Artificial intelligence and cognitive technology: Foundations and perspectives. In E. Hollnagel, G. Mancini, and D.D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 407-420). Berlin: Springer-Verlag.

Cooper, G. E. (1977). *A Survey of the status of and philosophies relating to cockpit warning systems* (NASA-CR-152071). Moffett Field, CA: NASA Ames Research Center.

CTA. (1993). *Production qualification test procedure for the Air Force Mission Support System (AFMSS) Mission Planning Subsystem (MPS).* (Vol. II: Procedure M4314, USI/Human factors verification; Document No. ATP4080670; Draft). Bedford, MA: Author.

Dannenberg, R., and Blattner, M. (1992). Introduction: The trend toward multimedia interfaces. In M. M. Blattner and R. B. Dannenberg (Eds.), *Multimedia interface design* (pp. xvii-xxv). New York: Addison-Wesley.

Deathrage, B. H. (1972). Auditory and other sensory forms of information presentation. In H. P. Van Cott and R. G. Kinkade (Eds.), *Human engineering guide to equipment design* (pp. 123-160). Washington, DC: American Institutes for Research.

Department of Defense. (1989a). *Keyboard arrangements* (MIL-STD-1280). Washington, DC: Author.

Department of Defense. (1989b). *Human engineering design criteria for military systems, equipment, and facilities* (MIL-STD-1472D). Washington, DC: Author.

Dreyfus, H. L. (1972). *What computers can't do.* New York: Harper and Row.

Dumas, J. S. (1988). *Designing user interfaces for software.* Englewood Cliffs, NJ: Prentice-Hall.

Dzida, W. (1989). The Development of ergonomic standards. *SIGCHI Bulletin, 20*, 35-43. (Special Interest Group on Computer & Human Interaction of the Association of Computing Machinery (ACM Press)).

Eberts, R. (1987). Human computer interaction. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 249-304). New York: Elsevier/North Holland.

Ellis, S. R., McGreevy, M. W., and Hitchcock, R. J. (1987). Perspective traffic display format and airline pilot traffic avoidance. *Human Factors, 29*, 371-382.

Ericsson, K. A. and Faivre, I. A. (1988). What's exceptional about exceptional abilities? In L. K. Obler and D. Fein (Eds.), *The Exceptional brain: Neuropsychology of talent and special abilities* (pp. 436-473). Guilford Press.

Fadden, D. M. (1993). Display Design. In Cardosi, K. and Huntley, M.S. (Eds.), *Human Factors for Flight Deck Certification Personnel* (DOT/FAA/RD-93/5). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Federal Aviation Administration/Air Traffic Operations Service. (1987). *Air traffic control* (7110.65E). Washington, DC: Author.

Federal Aviation Administration. (1991). *Federal Aviation Administration operations concept* (Volume I. ATC background and analysis methodology). Washington, DC: Department of Transportation/Federal Aviation Administration.

Fernandes, K. (1993). *User interface specifications for the Joint Maritime Command Information System (JMCIS)* (Version 1.3). San Diego, CA: Naval Command, Control, and Ocean Surveillance Center; Research, Development, Test, and Evaluation Division.

Fischer, G, Nakokoji, K., Ostwald, J., Stahl, G., and Sumner, T. (1993). Embedding computer-based critics in the context of design. *INTERCHI '93 Conference Proceedings* (pp. 157-164). New York: Association for Computing Machinery (ACM).

Foley, J. D. (1983). Human-computer dialogue: Interaction tasks and techniques -- A survey and categorization. In C. M. Mitchell, P. M. Van Balen, and K. Moe (Eds.), *Human factors considerations in system design* (NASA Conference Publication 2246, pp. 89-106). Greenbelt, MD: NASA/Goddard Space Flight Center.

Foley, J. D., VanDam, A., Feiner, S. K., and Hughes, J. F. (1990). *Computer graphics: Principles and practice* (2nd ed.). New York: Addison-Wesley.

Furness, T. A. (1986). Fantastic voyage. *Popular Mechanics, December*, 63-65.

Furness, T. (1986). The Super cockpit and its human factors challenges. *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 48-52). Santa Monica, CA: Human Factors Society.

Galitz, W. O. (1985). *Handbook of screen format design* (2nd ed.). Wellesley Hills, MA: QED Information Sciences.

Galitz, W. O. (1993). *User-interface screen design.* New York: Wiley-QED.

Gaertner, K. P., and Holzhausen, K. P. (1980). Controlling air traffic with a touch sensitive screen. *Applied Ergonomics, 11*, 17-22.

Gilmore, W. E., Gertman, D. I., and Blackman, H. S. (1989). *The User-computer interface in process control: A Human factors engineering handbook.* New York: Academic Press.

Gopher, D., and Raij, D. (1988). Typing with a two hand chord keyboard–Will the QWERTY become obsolete? *IEEE Transactions on Systems, Man, and Cybernetics, 18*, 601-609.

Greenstein, J. S., and Arnaut, L. Y. (1988). Input devices. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 495-519). New York: Elsevier-North Holland.

Grudin, J. (1989). The Case against user interface consistency. *Communications of the ACM, 32*, 1164-1173.

Hancock, P. A., and Chignell, M. H. (Eds.). *Intelligent interfaces: Theory, research and design.* New York: Elsevier/North Holland.

Hargreaves, W., Rempel, D., Halpern, N., Markison, B., Kroemer, K., and Litewka, J. (1992). Toward a more humane keyboard. *Proceedings of CHI '92: Human Factors in Computing Systems* (panel abstracts, pp. 365-368). New York: Association for Computing Machinery (ACM).

Harwood, K. (1994). Defining human-centered system issues for verifying and validating air traffic control systems. In J. A. Wise, V. D. Hopkin, and P. Stager (Eds.), *Verification and validation of complex systems: Human factors issues* (NATO Advanced Science Institute, Series F: Vol. 110, pp. 115-129). New York: Springer-Verlag.

Harwood, K. and Sanford, B. (1994) Evaluation in context. In J. A. Wise, V. D. Hopkin, and D. J. Garland (Eds.), *Human factors certification of advanced aviation technologies* (Aviation Human Factors Series, pp. 245-262). New York: Springer-Verlag.

Hillman, D. (1985). Artificial intelligence. *Human Factors, 27*, 21-32.

Hix, D., and Hartson, H. R. (1993). *Developing user interfaces: Ensuring usability through product & process.* New York: John Wiley.

Hollnagel, E. (1992). The reliability of interactive systems: Simulation based assessment. In J. A. Wise, V. D. Hopkin, and P. Stager (Eds.), *Verification and validation of complex and integrated human-machine systems* (Preprints of NATO Advanced Science Institute, pp. 27-42). Daytona Beach, FL: Embry-Riddle.

Hopkin, V. D. (1991). Automated flight strip usage: Lessons from the functions of paper strips. *Challenges in aviation human factors: The national plan* (pp. 62-64). Washington, DC: American Institute of Aeronautics and Astronautics.

Hopkin, V. D. (1989). Man-machine interface problems in designing air traffic control systems. *Proceedings of the IEEE: Special Issue on Air Traffic Control, 77*(11), 1634-1642.

Hopkin, V. D. (1988). *Human factors aspects of the AERA 2 program.* Farnborough, UK: Royal Air Force Institute of Aviation Medicine.

Horton, W. K. (1990). *Designing and writing online documentation.* New York: Wiley.

Horton, W. (1991). *Illustrating computer documentation.* New York: Wiley.

Ilg, R. (1987). Ergonomic keyboard design. *Behaviour and Information Technology, 6*, 303-309.

International Civil Aviation Organization (ICAO). (1993). *Human factors digest No. 8: Human factors in air traffic control* (Circular 241-AN/145). Montreal, Quebec: Canada.

Jiang, J., Murphy, E. D., Bailin, S. C., and Truszkowski, W. F. (1993). Automating a human factors evaluation of graphical user interfaces: An Update on CHIMES. *SpaceOps '92: Proceedings of the Second International Symposium on Ground Data Systems for Space Mission Operations* (JPL Publication 93-5, pp. 685-690). Pasadena, CA: Jet Propulsion Laboratory.

Jorna, P. G. M. (1992). Operator workload as a limiting factor in complex systems. In J. A. Wise, V. D. Hopkin, and P. Stager (Eds.), *Verification and validation of complex and integrated human-machine systems* (Preprints of NATO Advanced Science Institute, pp. 97-106). Daytona Beach, FL: Embry-Riddle.

Jubak, J. (1992). *In the image of the brain: Breaking the barrier between the human mind and intelligent machines.* New York: Little, Brown.

Kantowitz, B. H. (1977) Ergonomics and the design of nuclear power plant control complexes. In T.W. Kvalseth (Ed.), *Arbeidsplass og miljobruk av ergonomiske data.* Trondheim, Norway: Tapir.

Kantowitz, B. H. and Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships.* New York: Wiley.

Karat, C.-M., Campbell, R., Fiegel, T. (1992). Comparison of empirical testing and walkthrough methods in user interface evaluation. *CHI '92 Conference Proceedings* (pp. 397-404). New York: Association for Computing Machinery (ACM).

Kellogg, W. A. (1987). Conceptual consistency in the user interface: Effects on user performance. In H. J. Bullinger and B. Shackel (Eds.), *Human-computer interaction-INTERACT '87* (pp. 389-394). New York: Elsevier/North Holland.

Kellogg, W. A. (1989). The dimensions of consistency. In J. Nielsen (Ed.), *Coordinating user interfaces for consistency* (pp. 9-20). New York: Academic Press.

Kerns, K. (1992). Data-link communications between controllers and pilots: A Review and synthesis of the simulation literature. *The International Journal of Aviation Psychology, 1,* 181-204.

Kinney, G.C., and DeCicco, M.J. (1982). *Discussion of trials of four-color plan-view displays at Washington Air Route Traffic Control Center* (MTR-82W37). McLean, VA: MITRE.

Kley, V. (1983). Pointing device communication. *Computer Graphics World, 11,* 69-72.

Kloster, G. V., and Zellweger, A. (1987). Engineering the man-machine interface for air traffic control. *IEEE Computer, 20,* 47-62.

Kulik, C., and Welles, M. (1990). *AERA 2 computer human interface specification* (MTR-89W24, Revision 1). McLean, VA: MITRE.

Lee, S. K., Buxton, W., and Smith K. C. (1985). A multi-touch three dimensional touch-sensitive tablet. *Proceedings of CHI '85: Human Factors in Computing Systems* (pp. 21-25). New York: Association for Computing Machinery (ACM).

Lewis, C., and Norman, D. A. (1986). Designing for error. In D. A. Norman and S. W. Drapter (Eds.), *User centered system design: New perspectives on human-computer interaction* (pp. 411-432). Hillsdale, NJ: Erlbaum.

Laios, L. (1978). Predictive aids for discrete decision tasks with input uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-8,* 19-28.

Licklider, J. C. R. (1961). *Audio warning signals for Air Force weapons systems* (TR 60-814). Wright Patterson Air Force Base, OH: U. S. Air Force, Wright Air Development Division.

Lockheed Human Factors Engineering. (1993). *The MIL-STD-1472D Checklist.* Sunnyvale, CA: Lockheed Missiles and Space Co. (Available from the CSERIAC Program Office, Wright-Patterson AFB, OH: 45433-6573; (513)255-4842).

Lockheed Saunders, Inc. (1994). *Air Force Mission Support System (AFMSS) A/W/E Developers' Documentation Set* (Vol. 4 USI Style Guide, ISD 150/A020-2). Merrimack, NH: Information Systems Division.

Lu, C. (1983). Will there ever be an efficient keyboard? *High Technology, Jan.*, 22-23.

Lueder, R. (1986). Work station design. In R. Lueder (Ed.), *The Ergonomics payoff: Designing the electronic office.* Toronto, Ont., Canada: Rinehart & Winston.

Marcus, A. (1992). *Graphic design for electronic documents and user interfaces.* New York: ACM Press.

Matias, E., MacKenzie, I. S., and Buxton, W. (1993). Half-QWERTY: A one-handed keyboard facilitating skill transfer from QWERTY. *Proceedings of INTERCHI '93: Conference on Human Factors in Computing Systems* (pp. 88-94). New York: Association for Computing Machinery (ACM).

Mayhew, D. J. (1992). *Principles and guidelines in software user interface design.* Englewood Cliffs, NJ: Prentice Hall.

McCormick, E. J., and Sanders, M. S. (1982). *Human factors in engineering and design* (5th ed.), New York: McGraw-Hill.

Michaelis, P. R., Miller, M. L., and Hendler, J. A. (1982). Artificial intelligence and human factors engineering: A necessary synergism in the interface of the future. In A. Badre and B. Shneiderman (Eds.), *Directions in human/computer interaction* (pp. 79-94) Norwood, NJ: Ablex.

Miller, J. R., Sullivan, J. W., and Tyler, S. W. (1991). Introduction. In J. W. Sullivan and S. W. Tyler (Eds.), *Intelligent user interfaces* (pp. 1-10). New York: Addison-Wesley.

Mishkoff, H. C. (1985). *Understanding artificial intelligence.* Dallas, TX: Texas Instruments

Mitchell, C. M., and Saisi, D. L. (1987). Use of model-based qualitative icons and adaptive windows in workstations for supervisory control tasks. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-17,* 573-593.

Mogford, R. H. (1990). *CHI design objectives and usability goals for the Canadian Automated Air Traffic System (CAATS).* Ottawa, ON, Canada: Thompson-Hickling Aviation.

Mogford, R. H., Murphy, E. D., and Guttman, J. A. (1993). The use of direct and indirect techniques to study airspace complexity in air traffic control. *Proceedings of the First Mid-Atlantic Human Factors Conference* (pp. 196-202). Norfolk, VA: Old Dominion University.

Mudd, S. A. (1961). *The Scaling and experimental investigation of four dimensions of pure tone and their use in an audio-visual monitoring problem.* Unpublished Ph.D. thesis, Lafayette, Indiana: Purdue University.

Munson, W. A. (1947). The growth of auditory sensitivity. *Journal of the Acoustical Society of America, 19,* 584.

Murch, G. M., and Hubar, J. (1982). Colour—the logical next step. *New Electronics, 15,* 31-32.

Narborough-Hall, C. S. (1985). Recommendations for applying colour coding to air traffic control displays. *Displays, 6,* 131-137.

Nielsen, J. (1989). The matters that really matter for hypertext usability. *Hypertext '89 Proceedings* (pp. 239-248). New York: Association for Computing Machinery (ACM).

Nielsen, J. (1990). *Hypertext and hypermedia.* New York: Academic Press.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88,* 1-15.

Norman, D. A. (1986). Cognitive engineering. In D. A. Norman and S. W. Draper (Eds.), *User-centered system design* (pp. 31-61). Hillsdale, NJ: Erlbaum.

Norman, D. A. (1983). Design rules based on analyses of human error. *Communications of the ACM, 26* (April), 254-258.

Norman, D. A. (1993). *The Design of everyday things.* New York: Doubleday/Currency.

Norman, K. L. (1991a). Models of the mind and machine: Information flow and control between humans and computers. In M. C. Yovits (Ed.), *Advances in computers* (Vol. 32, pp. 201-253). New York: Academic Press.

Norman, K. L. (1991b). *The Psychology of menu selection: Designing cognitive control at the human/computer interface.* Norwood, NJ: Ablex.

Norman, K. L., Weldon, L. J., and Shneiderman, B. (1986). Cognitive layouts of windows and multiple screens for user interfaces. *International Journal of Man-Machine Studies, 25,* 229-248.

Oborne, D. J. (1987). *Ergonomics at work.* New York: Wiley.

Open Software Foundation (OSF). (1992). *OSF/Motif™ Style Guide,* Revision 1.2. Englewood Cliffs, NJ: Prentice-Hall.

Paap, K. P., and Roske-Hofstrand, R. J. (1988). Design of menus. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 205-235). New York: Elsevier/North-Holland.

Parish, R. N., Gates, J. L., Munger, S. J., Grimma, P. R., and Smith, L. T. (1982). Development of design guidelines and criteria for user/operator transactions with battlefield automated systems. (Phase II. Final report: Volume II. *Prototype and handbook for material developers*). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Parasuraman, R., Molloy, R. and Singh, I. L. (1993). Performance consequences of automation induced "complacency." *International Journal of Aviation Psychology, 3,* 1-23.

Patterson, R. D., and Milroy, R. (February 1979). *Existing and recommended levels for auditory warnings on civil aircraft.* London: MRC Applied Psychology Unit (Civil Aviation Authority Contract Number 7D/S/0142).

Paul, F., Sarlanis, K., and Buckley, E. P. (1965). A human factors comparison of two data entry keyboards. *Proceedings of the Sixth Annual Symposium of the Professional Group on Human Factors in Electronics.* New York: Institute of Electrical and Electronic Engineers.

Phillips, M. D., Bashinski, H. S., Ammerman, H. L., and Fligg, C. M., Jr. (1988). A task analytic approach to dialogue design. In M. Helander (Ed.), *Handbook of human-computer interaction,* (pp. 835-857). New York: Elsevier-North Holland.

Phillips, M. D., and Tisher, K. (1985). Operations concept formulation for next generation air traffic control systems. In B. Shackel (Ed.), *Human-computer interaction-INTERACT '84* (pp. 895-900). New York: Elsevier/North Holland.

Plaisant, C., and Sears, A. (1992). Touchscreen interfaces for alphanumeric data entry. *Proceedings of the Human Factors Society 36th Annual Meeting*, (pp. 293-297). Santa Monica, CA: Human Factors Society.

Plaisant, C., and Wallace, D. (1990). *Touchscreen toggle switches: push or slide? Design issues and usability study*, (CAR-TR-521; CS-TR-2557). College Park, MD: University of Maryland. (Reprinted in B. Shneiderman, Ed., 1993, *Sparks of innovation in human-computer interaction*, pp. 195-204. Norwood, NJ: Ablex).

Pollack, I. (1952). Comfortable listening levels for pure tones in quiet and noise. *Journal of the Acoustical Society of America, 24*, 158.

Potosnak, K. M. (1988). Keys and keyboards. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 475-494). New York: Elsevier-North Holland.

Potter, R. L., Weldon, L. J., & Shneiderman, B. (1988). Improving the accuracy of touchscreens: An experimental evaluation of three strategies. *Proceedings of the CHI '88 Conference on Human Factors in Computing Systems* (pp. 27-32). New York: ACM. (Reprinted in B. Shneiderman, Ed., 1993, *Sparks of innovation in human-computer interaction*, pp. 161-169, Norwood, NJ: Ablex.)

Rollins, R. A., and Hendricks, R. (1980). Processing of words presented simultaneously to eye and ear. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 99-109.

Rook, F. W., and Lehner, P. E. (1984). On the cognitive psychology of cooperative problem solving with intelligent machines (Paper presented at the AI Symposium, College Park, MD, University of Maryland).

Salvendy, G. (Ed.). (1987). *Cognitive engineering in the design of human-computer interaction*. New York: Elsevier.

Sanders, M. S., and McCormick, E. J. (1987). *Human factors in engineering and design* (6th ed.). New York: McGraw-Hill.

Sanders, M. S., and McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). New York: McGraw-Hill.

Sears, A. (1990) *Improving touchscreen keyboards: Design issues and a comparison with other devices* (CAR-TR-515; CS-TR-2536). College Park, MD: University of Maryland.

Sears, A., Plaisant, C., and Shneiderman, B. (1993). A new era for high precision touchscreens. In R. Hartson and D. Hix (Eds.), *Advances in human-computer interaction* (pp. 1-33). Norwood, NJ: Ablex.

Sears, A., and Shneiderman, B. (1991). High precision touchscreens: Design strategies and comparison with a mouse. *International Journal of Man-Machine Studies, 34,* 593-613. (Reprinted in B. Shneiderman, Ed., 1993, *Sparks of innovation in human-computer interaction*, pp. 171-186, Norwood, NJ: Ablex.)

Seminara, J. L., Gonzalez, W. R., and Parsons, S. O. (1977). *Human factors review of nuclear power plant control room design* (EPRI Report NP-309). Palo Alto, CA: Electric Power Research Institute.

Sherr, S. (Ed.) (1988). *Input devices.* New York: Academic Press.

Shneiderman, B. (1983a). Design of interactive systems. In C. M. Mitchell, P. M. Van Balen, and K. Moe (Eds.), *Human factors considerations in system design* (NASA Conference Publication 2246, pp. 47-87). Greenbelt, MD: NASA/Goddard Space Flight Center.

Shneiderman, B. (1986). *Designing the user interface.* Reading, PA: Addison-Wesley.

Shneiderman, B. (1992). *Designing the user interface* (2nd ed.). New York: Addison-Wesley.

Shneiderman, B. (1983b). Direct manipulation: A Step beyond programming languages. *IEEE Computer*, 16(8), 57-69. (Reprinted in B. Shneiderman, Ed., 1993, *Sparks of innovation in human-computer interaction*, pp. 13-62, Norwood, NJ: Ablex.)

Shneiderman, B. (1991). Touch screens now offer compelling uses. *IEEE Software, 8(2),* 93-94, 107. (Reprinted in B. Shneiderman, Ed., 1993, *Sparks of innovation in human-computer interaction*, pp. 187-194, Norwood, NJ: Ablex.)

Silverstein, L. D. (1987). Human factors for color display systems: Concepts, methods, and research. *Color and the computer*, pp. 27-61. New York: Academic Press.

Simon, H. A. (1981). *The Sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.

Simpson, S. A., and Navarro, T. (1984). Intelligibility of computer generated speech as a function of multiple factors. *Proceedings of the National Aerospace and Electronics Conference*, pp. 932-940. New York: IEEE.

Simpson, C. A., and Williams, D. H. (1980). Response time effects of alerting tone and semantic context for synthesized voice cockpit warnings. *Human Factors, 22*, 319-330.

Smith, B. (1992). The use of animation to analyze and present information about complex systems. *Beyond the vision: The Technology, research, and business of virtual reality* (Proceedings of Virtual Reality '91: The Second Annual Conference on Virtual Reality, Artificial Reality, and Cyberspace, pp. 190-199). Westport, CT: Meckler.

Smith, S. L., and Mosier, J. N. (1986). *Guidelines for designing user interface software* (ESD-TR-86-278). Bedford, MA: MITRE.

Sondheimer, N. K., and Relles, N. (1982). Human factors and user assistance in interactive computing systems: An introduction. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-12*(2), 102-107.

Stammers, R. C., and Bird, J. M. (1980). Controller evaluation of a touch input air traffic data system: An "indelicate" experiment. *Human Factors, 22*, 582-589.

Stevens, S.S., and Davis, H. (1938). *Hearing, its psychology and physiology.* New York: Wiley.

Stokes, A. F., Wickens, C. D., and Kite, K. (1990). *Display technology: Human factors concepts.* Warrendale, PA: Society for Automotive Engineers.

Streeter, L. A. (1988). Applying speech synthesis to user interfaces. In M. Helander (Ed.) *Handbook of Human-Computer Interaction* (pp. 321-343). New York: Elsevier.

Teichner, W. H. (1979). Color and information coding. *Proceedings of the Society for Information Display, 20*, 3-9.

Thimbleby, H. (1990). *User interface design.* New York: Addison-Wesley.

Thorell, L. G., and Smith, W. J. (1990). *Using computer color effectively.* Englewood Cliffs, NJ: Prentice-Hall.

Treisman, A., and Davies A. (1973). Divided attention to eye and ear. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 101-117). New York: Academic Press.

Tufte, E. R. (1990). *Envisioning information.* Cheshire, CT: Graphics Press.

Tullis, T. S. (1988). Screen design. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 377-411). New York: Elsevier-North Holland.

Van Cott, H. P., and Warrick, M. J. (1972). Man as a system component. In H. P. Van Cott and R. G. Kinkade (Eds.), *Human engineering guide to equipment design* (pp. 17-40). Washington, DC: American Institutes for Research.

Van Orden, K., DiVita, J., and Shim, M. J. (1993). Redundant use of luminance and flashing with shape and color as highlighting codes in symbolic displays. *Human Factors, 35,* 195-204.

Vortac, O. U. (Durso, F. T., Gronlund, S. D., Lewandowsky, S., and Gettys, C. F.). (1991). *Cognitive factors in the use of flight progress strips: Implications for automation.* Norman, OK: University of Oklahoma, Cognitive Processes Laboratory.

Vortac, O. U., Edwards, M. B., Jones, J. P., Manning, C. A., and Rotter, A. J. (1993). En route air traffic controller's use of flight progress strips: A graph-theoretic analysis. *International Journal of Aviation Psychology, 3,* 327-343.

Vortac, O. U., Edwards, M. B., Fuller, D. K., and Manning, C. A. (1993). Automation and cognition in air traffic control: An empirical investigation. *Applied Cognitive Psychology, 7,* 631-651.

Wenzel, E. M. (1992). Three-dimensional visual acoustic displays. In M. Blattner and R. Dannenberg (Eds.) *Multimedia interfaçe design* (pp. 257-288). New York: ACM Press.

Wickens, C. D. (1984). *Engineering psychology and human performance.* Columbus, OH: Charles E. Merrill.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

Wickens, C. D., Todd, S., and Seidler, K. (1989). *Three-dimensional displays: Perception, implementation, and applications* (CSERIAC SOAR-89-01). Wright-Patterson AFB, OH: Armstrong Aerospace Medical Research Laboratory.

Wickens, C. D., Vidulich, M., and Sandry-Garza, D. (1984). Principles of S-C-R compatibility with spatial and verbal tasks: The Role of display-control location and voice-interactive display-control interfacing. *Human Factors, 26*, 533-543.

Winston, P. H. (1984). *Artificial intelligence* (2nd ed.). Reading, MA: Addison-Wesley.

Woods, D. D. (1984). Visual momentum: A Concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21*, 229-244.

Woods, D. D, and Roth, E. M. (1988). Cognitive systems engineering. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 3-43). New York: Elsevier-North Holland.

Woods, D. D., and Sarter, N. B. (1992). Field experiments for assessing the impact of new technology on human performance. In J. A. Wise, V. D. Hopkin, and P. Stager (Eds.), *Verification and validation of complex and integrated human-machine systems* (Preprints of NATO Advanced Science Institute, pp. 77-88). Daytona Beach, FL: Embry-Riddle.

Woodson, W., Tillman, B., Tillman, P. (1992). *Human Factors Design Handbook* (2nd ed.) New York: McGraw-Hill.

Ziegler, J. E., and Fahnrick, K. P. (1988). Direct manipulation. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 123-133). New York: Elsevier-North Holland.

7.8 ENDNOTES

1. The distinction between "slips" and "mistakes" comes to us from Donald Norman (1981). His insightful and readable recent book (Norman, 1991) is highly recommended to both designers and evaluators. Analysis of possible and actual slips and mistakes can be useful in designing for the reduction of error (Lewis and Norman, 1986).

APPENDIX 7A. CHECKLIST ITEMS

This checklist is divided into the following sections:

I. Visual Displays
- A. General
- B. Visual Alerts

II. Auditory Alerts
- A. General
- B. Speech Messages

III. Data Entry Procedures
- A. General
- B. Commands and Command Execution
- C. Menus
- D. Error Messages and User Guidance

IV. Data Entry and Control Devices
- A. General
- B. Keyboards
- C. Touchscreens
- D. Trackballs
- E. Control Grip Devices
- F. Mice
- G. Graphics Tablets
- H. Pushbuttons (Actual and Virtual)
- I. Foot Switches and Pedals

Note: Checklist items marked with an "E" indicate items that must be assessed with equipment and/or by referring to the specifications documentation.

I. VISUAL DISPLAYS

A. General

1. Information does not disappear from the screen without being intentionally deleted or suppressed by the controller (7.1.20).

2. The computer responds quickly so that the controller is not kept waiting for information (i.e., in less than one second for simple, frequently performed tasks; within four seconds for more complex processing; and within 12 seconds for infrequent, highly complex processing) (5.1.2).

3. Essential ATC information is never blocked or obstructed by other information (7.2.17).

4. All information that a controller needs to accomplish a task that is essential and time-critical is located on a single page or in a single window (7.2.17).

5. Visual displays provide necessary information in a usable form when it is needed (5.3.2, 7.2).

6. Display clutter is not a problem (7.2.20).

7. The meaning of each icon is immediately apparent to the controller or it is labelled (7.2.1).

8. Symbols chosen for the display are intuitive so that the controller can interpret them quickly and accurately (7.2.9, 5.1.5).

9. Controllers can change the amount of task-related detail that is presented (7.2.20).

10. When the meaning of the color is critical, color is used redundantly with another type of visual cue, such as shape, text, or size. For example, all yellow objects have a triangular shape (7.2.12, 3.2.3).

11. The controller is able to recognize and differentiate between color codes under all anticipated lighting conditions (9.6.1).

12. The controller will not need to identify more than five colors (to interpret the meaning of the color when it stands alone) (7.2.13, 3.2.4).

13. Color displays are readable and adequately bright under all anticipated lighting conditions (9.6.1).

14. When the controller must distinguish between the color of characters and symbols, small blue characters and symbols are not used (7.2.11, 3.2.3).

15. Saturated (i.e., vivid) colors are used only for critical or temporary information (7.2.14).

16. Saturated (i.e., vivid) colors are not used for small objects or areas (7.2.14).

17. Saturated (i.e., vivid) red and blue are never presented next to each other (7.2.14, 3.1.8; also see Figure 3-11).

18. Colors are far enough apart in perceptual terms that they are not confusable even when "washed out" by sunlight, if applicable (3.2.3).

19. Characters and symbols can be read easily under all anticipated lighting conditions (e.g., from dim light to direct sunlight, if applicable) (9.3.4, 9.6.1, 9.6.2, 7.2.8).

20. Computer displays and controls are clearly visible and easy to use under all anticipated lighting conditions (e.g., from dim light to direct sunlight, if applicable) (9.3.4, 9.6.1, 9.6.3, 7.2.8).

21. To acquire needed information, the controller only needs to look at a single, localized display (6.1.2).

22. The position and form of displayed objects appear the same to the controller seated directly in front of the object and to team members, when they view the display from other anticipated viewing angles (7.2.7).

23. If windows are used, the controller can scroll the underlying data set (7.2.17).

24. If windows are used, the controller can move windows (7.2.17).

25. If windows are used, the controller can resize windows (7.2.17).

26. If windows are used, the controller can iconify display pages (7.2.17).

27. If windows are used, the controller can open and close windows (7.2.17).

28. The active window is highlighted to distinguish it from inactive windows (7.2.17).

29. The relationship between different windows is clear to the user (7.2.17).

30. All information that a controller needs to accomplish a given task is located in a single window or within a small number of related windows (7.2.17).

31. Abnormal data are emphasized effectively so that it attracts the controller's attention (7.2.11).

32. Updated data are emphasized effectively so that it attracts the controller's attention (7.2.11).

33. Acronyms in the new display system have the same meanings as in the previous system (7.2.2).

34. Terms in the new display system have the same meanings as in the previous system (7.2.2).

35. Symbols in the new display system have the same meanings as in the previous system (7.2.3).

36. Symbol size can be adjusted by the controller (7.2.10).

37. Visual displays and their labels are sufficiently visible under all anticipated lighting conditions (9.3.4).

38. If size coding is used, it is limited to two widely different sizes (7.2.11).

39. Graphic displays are used only to present information that is naturally pictorial and to present dynamic data (7.2.5).

40. Placement of standard data fields is consistent from one display to another (7.2.18).

41. Formats used within data fields are consistent from one display to another (7.2.18).

42. Labels, terms, and abbreviations are used consistently across the display set (7.2.15).

43. Only one abbreviation is used for each word or item and abbreviations are used consistently on all visual displays (7.2.15).

44. Punctuation is used conservatively and consistently (7.2.15).

45. Continuous text is presented in mixed upper-and-lower case (7.2.15).

46. Computer printouts (in upper and lower case) are available for lengthy text (7.2.15).

47. Visual displays maintain good image quality even at the dimmest possible setting (7.2.8).

E 48. According to the display monitor manufacturer's report, the display refreshes at a rate of 65 cycles (or more) per second so that the display does not appear to flicker (7.2.6, 7.2.22, 3.1.6).

E 49. According to the display monitor manufacturer's report, a displayed object moves no more than .0002 times the viewing distance (in inches) in one second so that no display jitter can be detected (7.2.22, 7.2.23).

E 50. The heights and widths of characters appearing at the center and the four corners of the displays do not vary by more than 10 percent (7.2.22).

E 51. When the center of the display is compared to an edge, brightness uniformity does not vary by more than 50 percent (7.2.22).

E 52. The luminance of dynamic text and symbols are eight times that of the static background (7.2.22).

E 53. All colors are 8 times brighter than the static background symbology (7.2.14).

E 54. When the controller must distinguish between the color of characters, character height is at least 21 minutes of arc (7.2.10, 3.2.4).

B. Visual Alerts

1. Information that the controller must read and understand quickly, such as alarms or critical error messages, never blinks or flashes rapidly (greater than 3 Hz) (5.1.3, 5.1.5, 7.2.11).

2. High-priority alerts and other critical information are located within the central display area (i.e., the central 15 degrees of the area where the controller normally looks, given the normal viewing position) (7.2.11, 9.3.4).

3. Highlighting and blinking are used sparingly (7.2.11).

4. The color red is used only for warning/danger (7.2.12).

5. Yellow is used to indicate caution (7.2.12).

6. Green is used to indicate for normal/ready status (7.2.12).

7. The same color coding strategy is applied to every display used by the same controller (7.2.12).

8. No more than two levels of blinking are used (7.2.11).

9. If blinking is used, it is cancelable by the controller (7.2.11).

10. For a time-critical warning system (such as a conflict detection or resolution advisory), the controller response time has been measured and is within acceptable limits (5.1.5).

11. This design effectively directs the controller's attention by means of alerting, coding, and emphasis techniques (5.3.1, 7.2.11).

12. If size coding is used it is limited to two widely different sizes (7.2.11).

13. Alerts have a low incidence of false alarms (7.2.11).

E 14. Information that is blinking, has an "on" period that is at least as long as the "off" period (5.1.5).

E 15. If blinking is used, the blink rate is between 2 and 3 Hz (5.1.5, 7.2.11).

II. AUDITORY ALERTS

A. General

1. Auditory alerts are used only when necessary (7.3.2, 7.3.3).

2. The number of auditory alerts is sufficient, that is, auditory alerts are included wherever they are needed (7.3.2).

3. The meanings of auditory alerts are readily apparent (7.3.7).

4. All proposed auditory alerts have been tested and evaluated in a realistic environment by a representative set of controllers (7.4.3).

5. The auditory alert does not nag, or otherwise annoy, the controller (7.3.5, 7.3.6).

6. Auditory signals (and speech messages) are not masked by other auditory alerts or background noise (7.3.7, 7.3.8, 4.1.3).

7. For any situation, it is impossible for more than a few auditory alerts to be presented simultaneously (7.3.7).

8. The number of auditory signals (e.g., warnings, alerts) that the controller may need to identify is fewer than five (7.3.8).

9. Auditory alerts are easily discernible from other signals or noise (7.3.7, 7.3.8, 4.1.3).

10. Auditory alerts do not provide more information than is necessary (7.3.7).

11. The same auditory signal always indicates the same information (7.3.7).

12. Auditory alerts are consistently implemented throughout the system (7.3.7).

13. The information contained in an auditory alert is also displayed visually (7.3.8).

14. Auditory alerts are only used when immediate action is required (7.3.8).

15. Auditory alerts are cancelable by the controller (7.3.6, 7.3.7).

E 16. A modulated signal emits from one to eight beeps per second (7.3.8).

E 17. A warbling sound varies from one to three times per second (7.3.8).

E 18. The frequency of all auditory signals is between 500 and 3000 Hz so that they are well within the band of frequencies that humans are most sensitive to (7.3.8).

E 19. In general, auditory alerts sound for at least a 0.5 second duration (7.3.8).

E 20. The pause between repetitions of an auditory signal is less than or equal to three seconds (7.3.8).

E 21. Auditory alerts are at least 10 dB above ambient noise or have been demonstrated to be sufficiently intense for a specific working environment (7.3.7, 7.3.8).

B. Speech Messages

1. A detection signal display (for example the sound of static on the line) precedes a voice warning, unless a distinctive synthesized voice is used (7.3.3).

2. Speech messages are short enough to be easily remembered. (7.3.4).

3. Brief speech messages are available to the controller when there is the need to explain the specific nature of alarm and warning signals (7.3.4).

4. Speech displays are distinct from and not easily confused with other voices in the control room (7.3.4).

5. The controller does not need to remember more than one or two speech messages at a time in order to accomplish any of his or her ATC tasks (7.3.5, 7.3.7).

6. Speech messages are not masked by other auditory alerts or background noise (7.3.7, 4.1.3).

E 7. If important messages are produced by synthetic speech, they are at least 8 Db above the surrounding noise (4.2.5).

III. DATA ENTRY PROCEDURES

A. General

1. The number of keystrokes (or other control actions) necessary to input data is kept to a minimum and the amount and complexity of data entry is about the same as was required in the previous system (6.2.2).

2. With this system, data-entry errors can be caught and corrected before they propagate through the system (6.3.3).

3. The design assists the controller in detecting and correcting errors in data entry (5.2.5, 6.3.3).

4. This system makes it easy to recover from data-entry errors (6.3.3).

5. Keystrokes or other data-entry actions are echoed immediately on the screen, that is, there is no delay in providing a legible representation of what has been entered (7.5.3).

6. The data entry method helps to minimize errors and provides for quick, simple data editing and correction (7.5.3).

7. This user interface system queries the controller at critical choice points, e.g., "Are you sure you want to delete this flight plan?" (6.3.3).

8. A particular data item, such as assigned altitude, must be entered only once; the computer can retain this value and enter it in other fields, as appropriate (7.5.3).

9. The controller receives appropriate feedback on data acceptance or rejection (7.5.3).

10. The computer does not erase all or part of any erroneous data entry (7.5.3).

11. The controller controls the pace of data entry, that is, the computer does not impose time limits or time outs (7.5.3).

12. The computer does not restrict the order in which data items are entered (7.5.3).

13. The computer prompts the controller for data that have been deferred for entry (7.5.3).

14. Data processing is initiated only after an explicit command from the controller (7.5.3).

15. Boundaries indicate where to enter the data and show maximum field length (7.2.17).

16. A cursor appears to indicate data-entry mode and location (7.5.3).

17. The controller can edit all or part of a data field (7.5.3).

18. The controller is not required to enter leading zeroes for numeric entries (7.5.3).

19. When delimiters, such as punctuation, are required to partition long entries, the computer provides the required format and prompts for the order of data entry (7.5.3).

20. Field labels use accepted ATC terminology and are used consistently (7.5.4).

B. Commands and Command Execution

1. Command execution requires minimal controller action (7.5.2).

2. The consequences of destructive commands (e.g., "Delete") are explained (7.5.2).

3. Destructive commands (e.g., "Delete") require controller confirmation of intention before they are executed (7.5.2).

4. Command execution always occurs by explicit controller action, never as a by-product of another action (7.5.2).

5. The controller can suspend/interrupt or cancel/undo a transaction in progress (7.5.3).

6. Command ordering is consistent from screen to screen/window to window (7.5.2).

7. Command labels use accepted ATC terminology and are used consistently (7.5.4).

8. The relevant command set is displayed to show the controller which commands are currently available (7.5.2).

9. Commands are consistent in their placement across multiple screens, panels, or windows; in their wording; and in their method of activation (7.5.2).

10. The computer indicates the current operational mode (7.5.2).

11. Entry of long sequences of command parameters is not required (7.5.2).

12. Upper- and lower-case letters are accepted as equivalent when the controller is entering a command or command parameter (7.5.2).

13. Feedback is always given to indicate that the computer has initiated a command (7.5.2).

14. Commands should be stated in the affirmative; that is, they should tell the controller wha to do, rather than what not to do (7.2.15).

C. Menus[1]

1. Menu options are phrased to reflect the action executed and worded in user vocabulary (7.2.18).

2. Options that perform opposing actions are not placed adjacent to each other (7.2.18).

3. The number of menu options is between three and ten (five to six options is optimal) (7.2.18).

4. If an option, or set of options, is never available to the user, the option(s) is not in the menu (7.2.18).

5. If an option is temporarily unavailable, it is displayed in the menu, but dimmed (7.2.18).

6. Menu options are organized in logical or functional groupings with clear titles (7.2.18).

7. If not in logical groups, order is by frequency of usage, with most frequently used options at the top (7.2.18).

8. If not in logical groups or by frequency, options are in alphabetical or numerical order (7.2.18).

9. Less frequently executed options and destructive commands are at the bottom of the menu (7.2.18).

10. If similar options are in different menus, the options are ordered in a consistent manner (7.2.18).

[1] from User Interface Specifications for the Joint Maritime Command Information Systems, Version 1.3, by Kathleen Fernandes, November 1993.

11. Each word in the menu is presented in upper and lower case with the first letter capitalized (7.2.18).

12. Cascading submenus appear to the right of the parent menu (below, if space to the right is limited) (7.2.18).

13. When a menu is displayed, the location cursor is in the first available option (7.2.18).

14. When a pop-up menu appears, it appears near the element with which it is associated (7.2.18).

15. A window containing a pop-up menu provides an indication that the menu is available (7.2.18).

16. If they are presented in a vertical list, menu options are left justified (7.2.18).

17. Menu organization supports specific controller tasks (7.2.18).

18. Graphical or textual aids are provided to assist controllers in navigating through menu structures (7.2.18).

19. The controller is required to traverse no more than four levels in a menu structure (7.2.18).

20. When a trade-off is required between menu breadth (i.e., number of options at a level) and menu depth (i.e., number of levels), the design increases breadth rather than depth (7.2.18).

D. Error Messages and User Guidance

1. Error messages are provided whenever needed (7.5.5).

2. Each error message briefly summarizes the specific problem and proposes a specific solution (7.5.5, 7.5.6).

3. Error messages are direct and precise (7.2.14, 7.5.6).

4. Error messages are presented immediately after an error's occurrence (7.5.7).

5. Error messages are not redundant (7.5.6).

6. Guidance messages are presented in mixed upper and lower case (7.5.7).

7. Messages about limits not met or exceeded specify the appropriate range for data entry (7.5.8).

8. Questionable data entries elicit cautionary messages (7.5.8).

9. Feedback regarding processing delays specifies the process, the length of the delay, and completion of the process (7.5.8).

IV. DATA ENTRY AND CONTROL DEVICES

A. General

1. Input devices work in ways that are compatible and consistent with the controller's tasks (7.4.1).

2. The overall design of input devices does not require frequent switching between devices (7.4.2).

3. The input device(s) is/are appropriate for performing the necessary functions (e.g., alphanumeric data entry; selection of displayed objects; cursor positioning) (7.4.3).

4. Input devices have been compared not only for speed and accuracy, but also for factors such as induced fatigue, resolution capability, and space requirements (7.4.4).

5. Controls and their labels are sufficiently visible under dim lighting conditions (9.3.4).

E 6. Input devices meet operational requirements for accuracy, force, feedback, precision, and manipulation (7.4.1, 7.4.3).

B. Keyboards

1. Alphanumeric keys are arranged consistently on all keyboards that the controller will use (7.4.3). (The preferred arrangement is the QWERTY layout.)

2. Keyboards are readable under all operating conditions and backlit, if necessary (7.4.3).

3. If a numeric keypad is provided, it is visually separated from the main keyboard and arranged in a 3 X 3 + 1 matrix (7.4.3).

4. Function keys are provided for frequently used commands (7.4.3).

5. Function keys are clearly labeled to indicate their function (7.4.3).

6. The functions invoked by the function keys are consistent throughout the system (7.4.3).

7. Keys on keyboards and keypads have no more than two functions (9.2.2).

8. Nonactive keys are left blank (i.e., not labeled) (7.4.3).

9. The key used to initiate a command is clearly labeled "Enter" (7.4.3).

10. Keyed data are displayed quickly (echoed) on the screen (7.4.3).

11. Tactile and auditory feedback are provided in response to keystrokes (7.4.3).

12. The main keyboard is located directly in front of and below the associated visual display at a comfortable distance from the seated controller's position (7.4.3).

13. Forearm and wrist supports are provided (7.4.3).

14. Alphanumeric keys meet standards for dimensions, displacement, and separation (7.4.3).

15. Keyboard design includes guards to reduce inadvertent key activation (7.4.3).

16. If alternative keyboards are featured, they have been tested for usability and operational suitability (7.4.3).

E 17. The slope of the keyboard is adjustable between 15 and 25 degrees from the horizontal (7.4.3).

E 18. Keyboard height is adjustable between 23 and 32 inches (7.4.3).

C. Touchscreens

1. If a touchscreen is used, it is suitable for the task(s) to be performed by the controller (7.4.3).

2. Controllers can achieve sufficient touch accuracy with the touchscreen (7.4.3).

3. Touchscreen displays can be read easily under all anticipated lighting conditions (7.4.3)

4. The touch input strategy (e.g., land-on; first contact; lift-off) is compatible with the controller's task objectives (7.4.3).

E 5. Touchscreen displays meet standards for required finger pressure (displacement), separation of touch areas, and resistance (7.4.3).

D. Trackballs

1. The trackball can move the cursor in any direction without causing cursor movement in the opposite direction (7.4.3).

2. The trackball allows the controller to move the cursor quickly across relatively large distances and also to precisely position the cursor within a small area (7.4.3).

3. The trackball meets standards for physical dimensions, resistance, and clearance (7.4.3).

E. Control Grip Devices

1. Any input device meant to be held and operated by a standing controller can be held comfortably for a period of three to four hours (7.4.3).

F. Mice

1. If a mouse is part of the design, it can be used compatibly with all of the tasks the controller is supposed to perform (7.4.3).

2. Controllers can easily and smoothly position the cursor with the mouse (7.4.3).

3. Movement of the mouse produces cursor movement in the same direction on the display. For example, if the mouse is moved to the left, the cursor moves to the left on the display (7.4.3).

4. The mouse is equally usable with the left or right hand (7.4.3).

E 5. The mouse has no sharp edges and meets standards for width (1.5 to 3 in.), length (3 to 5 in.), and thickness (1 to 2 in.) (7.4.3).

G. Graphics Tablets

1. Movement of the stylus in any direction on the tablet surface produces smooth movement of the cursor in the same direction (7.4.3).

2. When the stylus is placed at any point on the tablet, the cursor appears at the associated coordinates on the display screen and maintains that position until the stylus is moved (7.4.3).

3. If the stylus and tablet are to be used for free-hand drawing, the device generates a continuous line as the stylus is moved (7.4.3).

4. If a graphics tablet is used, frequent switching to the keyboard is not necessary (7.4.3).

5. The graphics tablet can be located on the workstation within a comfortable distance from the controller (7.4.3).

H. Pushbuttons (Actual and Virtual)

1. Mechanical pushbuttons are sized and spaced to support activation but to prevent accidental activation (7.4.3).

2. The surfaces of "hard" pushbuttons are rough or concave (7.4.3).

3. Labeling of virtual pushbuttons is consistent (7.4.3).

4. The active and inactive states of virtual pushbuttons are visually distinct (7.4.3).

5. The on-off status of software-generated togglebuttons is made clear through the use of labels and graphic indicators (7.4.3).

E 6. Mechanical pushbutton resistance is in the range recommended for single-finger operations (10-40 oz.) (7.4.3).

I. Foot Switches and Pedals

1. Positive feedback is provided to indicate activation of the foot switch (7.4.3).

2. The controller is not required to operate more than one switch or pedal with the same foot (7.4.3).

3. Foot switches are positioned for operation by the toe or ball of the foot (7.4.3).

E 4. Foot switches/pedals meet requirements for dimensions, resistance, and displacement (7.4.3).

CHAPTER 8. WORKLOAD AND PERFORMANCE MEASUREMENT IN THE ATC ENVIRONMENT

Elizabeth D. Murphy

> "The human is most reliable under moderate levels of workload that do not change suddenly and unpredictably."
>
> Kantowitz and Casper (1988)

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 8. WORKLOAD AND PERFORMANCE MEASUREMENT

Why do we want to measure workload and performance? One answer is that we do not want the new system to overload the controller's natural human capabilities. Another answer is that we do not want to risk long periods when the controller's capabilities are underloaded. Both underload and overload have serious implications for human performance. Because the relationship between workload and performance is rather complex and not well understood, simply reducing workload does not guarantee improved performance or productivity. It may, in fact, have the opposite effect. An objective of workload and performance measurement is to provide data that system designers can use as a basis for identifying and redesigning embedded sources of overload and underload.

8.1 INTRODUCTION

For years, researchers have wrestled with workload measurement. At the request of system designers and evaluators, they have searched for the mythical workload thermometer—a magical measuring device equipped with a red "do not exceed" line. Unfortunately, workload measurement will never be that simple. A difficult task for one person is accomplished with ease by the next. The level of workload depends as much on the skill of the operator as it depends on the task. **The only unchanging truth is that workload measurement is as vital to an evaluation of new equipment and procedures as it is difficult.**

This chapter gives an overview of the issues associated with evaluating design products for their effects on controller workload and performance. New technologies should neither overload nor underload the controller, but should support efficient and effective controller job performance. Evaluating design products for effects on controller workload is one of the team's primary mandates.

After defining terms and presenting basic concepts, the chapter reviews some techniques to evaluate the effects of design products on ATC outcomes.

8.1.1 How should "workload" be defined as a basis for workload assessment?

Perils and Pitfalls in Assessing ATC Workload
The specific difficulties of workload assessment in ATC are summarized as follows (Hopkin, 1979): The variety and complexity of the controller's actions make it difficult to express all aspects of the controller's job in any single measurable dimension (p. 382). A practical problem is to measure the consequences for mental workload of a proposed change in system design. The main applications of any successful measure of mental workload would be in evaluations of proposed future systems to demonstrate that the changes they entail do not imply excessive workload, and that workload would not be reduced so much that boredom and job dissatisfaction would result (pp. 383-384).

In defining controller workload, it is important to recognize the differences between system workload and human workload. **System workload** is typically defined in terms of the number of inputs and outputs a computer system can handle in a unit of time. Some observable **human workload** comes from entering data into the computer system and dealing with the information generated by the system.

Typically, computer systems are expected to have very high levels of reliability so that high input-output levels can be maintained over long periods of time. People, however, cannot function like computers. Human operators can sustain high levels of input and output only for short periods of time. Just as constantly revving an engine will shorten its life, requiring a person to work constantly at full capacity will adversely affect productivity over time.

Some attempts to objectify the definition of workload equate observable workload with time pressure, that is, the ratio of time available to time required for task completion (e.g., Geer, 1981). Such a ratio can account for workload in observable tasks, such as entering data, turning knobs, and relaying messages. One drawback to the ratio approach, however, is that it typically cannot account for workload in purely mental tasks, such as planning and problem solving. Although time pressure is an important aspect of workload, it does not tell the whole story. Other factors that need to be considered include effort exerted, timing of tasks, operational strategies, and perceived stress.

For purposes of human factors research, **there is no single, agreed-upon definition of workload**. Different definitions are used depending on the particular purpose or research question. In the ATC context, workload is often used in the sense of task demands, or what the controller has to do in relation to each controlled aircraft. Many of these demands are observable inputs, and they can be quantified. For example, numbers of aircraft, complexity of traffic mix, complexity of flight paths, amount of communication and coordination required, and amount of data entry are all quantifiable contributors to ATC workload. Time pressure can be quantified in terms of time available versus time required. Of course, time required may vary as a function of skill and experience.

Any assessment of ATC workload must consider both observable (objective) and perceived (subjective) aspects of demand on the controller. A key point in workload evaluation is that there is no such thing as absolute workload independent of skill and experience. **It is imperative that system designers, manufacturers, and evaluators define their use of the term "workload."**

8.1.2 What is the difference between observable and perceived workload?

Observable workload comes from sources in both the external ATC situation and the ATC-facility environment. The ATC situation includes observable elements, such as number of aircraft, mix of aircraft types, number of intersecting flight paths, weather, other facilities, and so forth. The ATC facility presents the controller with numerous sources of observable workload: displays and controls at the ATC console, procedural requirements for manual data entry, communications equipment and procedures, paper documentation, requirements for coordination with other controllers and supervisors, and so forth. Observable workload can be summed up as those verbal or manual behaviors that the controller has to perform.

Perceived workload is the individual controller's personal experience or subjective perception of the demands imposed by the ATC environment and the ATC job (Stein, 1985). In this sense, workload represents what the controller has to "pay" in terms of effort or resources invested, to achieve a particular level of performance in a particular situation (Jorna, 1992). Ratings of perceived workload are often highly correlated with number of aircraft in the sector (Leighbody, Beck, and Amato, 1992; Stein, 1985).

The same level of observable workload may be perceived differently by different controllers, depending on their training, experience, skill, fatigue, or other factors (e.g., Hart and Wickens, 1990; Hopkin, 1988; Kantowitz and Casper, 1988; Moray, 1982). For example, a trainee (developmental) may have only five aircraft in the sector on a clear day (low observable workload), but the trainee's perceived workload may be high. For an full performance level (FPL) controller with 10 years of experience, this would probably be a low workload situation, barring unusual circumstances.

Perceived (or subjective) workload is an important issue for several reasons:

- Reducing observable task demands does not necessarily reduce perceived workload.
- Increases in this kind of workload may threaten safety and capacity.
- Any increase in perceived workload will threaten workforce acceptance of the new system.
- Severe decreases in perceived workload will induce boredom and monotony, which can adversely impact controller performance and job satisfaction.

The primary objective of workload assessment is to ensure that system demands neither overload nor underload the controller's information-processing capabilities for prolonged periods of time (Hopkin, 1979, 1988; Kantowitz and Casper, 1988). Both overload and underload can impact job performance. Another objective is to ensure that the controller always has some information-processing capacity in reserve to deal with the unexpected. To meet these objectives, a program of workload assessment must investigate both observable and perceived workload.

8.1.3 What is the objective of workload measurement?

Since moderate workload is associated with optimal human performance, workload evaluation should identify a design's potential for underloading as well as overloading the controller. Each extreme risks increased error. A regimen of moderate, balanced workload is generally recommended (e.g., Jorna, 1992).

Deciding what is moderate, however, may be difficult because of the variations in perceived workload from one person to the next. What is acceptable workload for one person may overload or underload another:

> "The ranges of acceptable and unacceptable workload simply are not known. Identifying these points or ranges has proven to be difficult because sources of workload vary among tasks, different people respond to the same objective demands by adopting different strategies and exerting different effort, and individuals [differ in their] abilities to cope with excessively low or high workload" (Hart and Wickens, 1990, p. 286).

This is why participants in workload assessments should represent the range of experience, skills, and abilities that are present in the controller workforce. Experienced controllers cannot be expected to anticipate all the problems that developmentals might have with a new system.

Absolute statements about workload should be viewed with skepticism, even if they are based on so-called "objective" measures. Even the most objective measures of perceived workload do not yield information that is comparable to measures of physical quantities. This is because workload is an inherently subjective construct. It is best to assess workload in relative terms (i.e., ratings for a new or enhanced system compared to ratings on the same measure for a reference system). Workload measures need to be put in context and compared to valid, reliable findings for prior systems (Hart and Wickens, 1992).

Workload can be evaluated for individual ATC subsystems, but it cannot be assumed that workload for one subsystem is independent of workload for another subsystem. Nor can it be assumed that workload levels for subsystems can be added to get the workload level imposed by the total system. It is important to assess the workload generated by interactions between and among subsystems. Controller workload may be acceptable for individual subsystems but unacceptable when they are combined.

8.1.4 How is "performance" defined?

For the purposes of system design and evaluation, job performance refers to the *common patterns of information processing and controller actions* that emerge when controllers actually use design products. (In this context, the term "performance" emphatically does NOT refer to individual performance evaluation or assessment for reporting purposes.)

Human-factors evaluations look for common trends in the ways that design products aid or impair the controller in performing the required tasks. These evaluations look at system efficiency and effectiveness across individuals. (Efficiency refers to speed, getting the job done in a timely manner; effectiveness refers to accuracy and overall quality, getting the job done in a safe and orderly manner.)

To a large degree, the controller's performance of information-processing tasks is crucial to effective, efficient control action. Therefore, it is important to look at how well the design supports information-processing tasks. An evaluation that looks only at observable controller actions may find acceptable levels of speed and accuracy, but the question remains, "At what cost to the controller?"

Controllers are known for their ability to make almost anything work. But should they have to do this? The technology-centered approach says, "Build it to do whatever it can and put it in the field." The user-centered approach asks, "What do controllers really need? Let's find out, build what is needed, and make sure it is suitable." **A design product is not what controllers need if controllers must invest more time and effort to make it work than they are investing in their current way of doing the job.**

Because it is critical for the controller to maintain high-performance standards with new ATC systems, it is essential to investigate the effects of workload on controller information-processing and observable actions. For example, it is important to identify opportunities for design-induced error that may not be immediately obvious but may correlate with certain workload conditions (e.g., Senders and Moray, 1991).

8.1.5 How are the terms "error" and "design-induced error" used in the context of workload and performance measurement?

The argument has been made that many errors can be identified only from hindsight (e.g., Norman, 1983). At the time that an action is taken or not taken, the person thinks s/he is doing the right thing. Typically, in ATC, there are many "right" ways to reach the same goal.

In this context then, **error means any shortcoming or confusion of intention or action that can be identified from its consequences.** An error may or may not result in an operational or system error, that is, a violation of separation minima. In many cases, errors will be caught and corrected by the controller before they propagate through the system; in other cases, the consequences of error may range from minor to severe.

Although the causes of error are many and varied, **a new design often introduces new sources of error.** These design-induced errors are traceable to specific features in the design product. **Evaluators should be on the lookout for design flaws that can induce mistakes.** For example, one of the contributing factors to a highly-publicized rash of automobile accidents (due to sudden acceleration) was a design flaw that made the brake and accelerator pedals almost indistinguishable without looking at them (Pollard and Sussman, 1989). Brake and gas pedals should be spatially separated, at different heights from the floor, and "feel" differently (i.e., have different force-deflection characteristics). Without these cues, drivers can step on the accelerator instead of, or at the same time as, stepping on the brake - sometimes with tragic consequences.

Evaluators should also look for error-tolerance in a new system design, that is, a capability to trap errors that are not caught and corrected by the controller. (Of course, designers will not usually provide an error-tolerant capability unless it is specified in the requirements document.)

In connection with error, the *concept of latent error* in automated systems becomes highly important:

> "All man-made systems contain potentially destructive agencies, like the pathogens within the human body. At any one time, each complex system will have within it a certain number of latent failures, whose effects are not immediately apparent but that can serve both to promote unsafe acts and to weaken its defense mechanisms. For the most part, they are tolerated, detected, and corrected, or kept in check by protective measures...But every now and then, a set of external circumstances...arises that combines with these resident pathogens in subtle and often unlikely ways to thwart the system's defenses and to bring about its catastrophic breakdown" (Reason, 1990. p. 197).

Although controllers are known for their ability and willingness to overcome design flaws, any system or procedure that requires controllers to invent a work around constitutes a potential source of delay, error, or additional self-imposed workload. Latent errors in the design are like "bugs" that have not been caught in software testing. They lie in wait for the combination of circumstances that will produce design-induced error.

8.2 ISSUES IN WORKLOAD AND PERFORMANCE MEASUREMENT

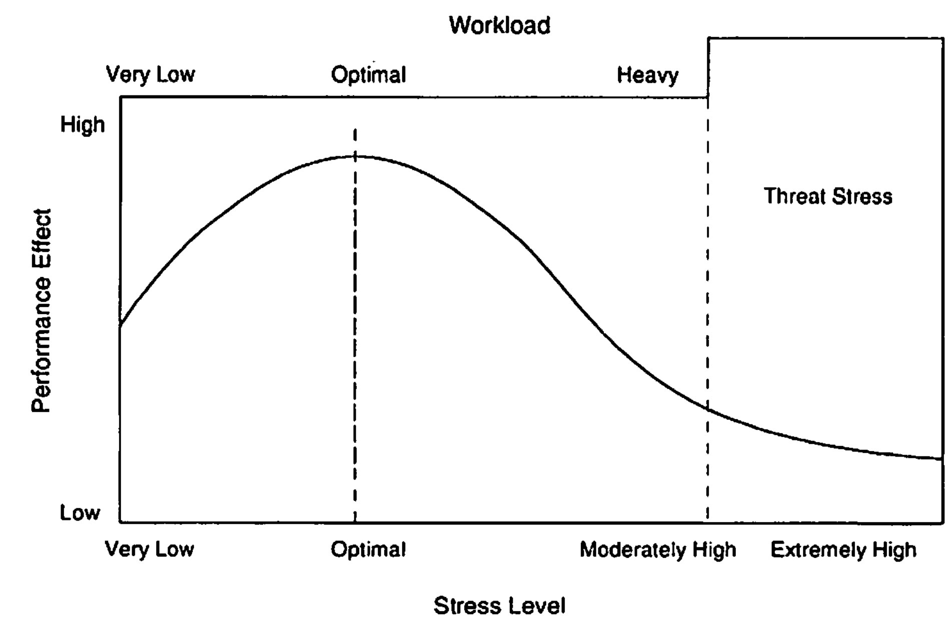
The research literature is bubbling over with controversy in the areas of workload and performance measurement. Much of the controversy focuses on research design issues, such as which measurement methods to use for particular purposes. In this section, we review selected issues that are of special concern in an ATC system design-and-development context.

Performance is most effective (i.e., the probability of error is lowest) at an optimal level of workload. The difficulty is that what is "optimal" varies from person to person. An optimal level of workload is "characterized by an active interaction between the person and his [or her] environment—talking with others, reading displays, adjusting controls, making decisions, etc., at a pace that the person can manage comfortably." Swain and Guttmann, 1983, p. 17-6

8.2.1 What is the relation between workload and job performance?

Based on intuition, we might assume that performance suffers as workload increases. This is only partially correct. As indicated in Figure 8-1, as workload increases from low to moderate, performance improves to an optimal level. When workload is low, boredom is a problem.

Figure 8-1. Relationship Between Workload and Job Performance
From *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications* (NUREG/CR-1278, SAND80-0200) by A. D. Swain and H. E. Guttman, 1983. By the U.S. Nuclear Regulatory Commission.



As effort increases, performance improves but does not continue improving in a straight line. If workload continues to increase, requiring increased effort, the relative improvement in performance declines. Notice that performance is typically best at moderate workload levels, and worst at both very low and very high levels of workload. The relationships depicted in the figure represent human responses in time-pressured operational environments where control actions taken or not taken can have serious consequences.

In the human factors literature, workload is thought to be a catalyst for, rather than a direct cause of, error (e.g., Morris and Rouse, 1988). That is, it is unlikely that workload alone will cause an error. It is more likely that workload will be a contributing factor (Reason, 1990, 1992; Stern, 1987) along with other factors, such as fatigue and false expectations or assumptions. Very low workload is generally associated with boredom and decreased effectiveness in monitoring performance, while very high workload is associated with information overload that results in decreased efficiency of information processing, which can lead to loss of situational awareness (e.g., Hopkin, 1979; Thackray, 1980).

Results of an evaluation need to be interpreted in terms of the linkage between workload and performance. Recommendations can be derived from the principle of conservation of resources. **Just as some system capacity needs to be kept in reserve, the controller cannot be expected to work at full capacity all the time.** Over time, people are more productive at moderate levels than at high levels of workload. The design must provide for some capacity to be kept in reserve for the unexpected.

8.2.2 What is the role of workload assessment in the design-and-development process?

"If the ATC demands of a particular job are excessive for nearly everyone doing that job, the demands must be reduced by redesigning tasks and reallocating responsibilities."
ICAO, 1993, p. 34

The primary role of workload assessment in the system design-and-development process is to identify the potential for information processing overload and underload so that breakdowns in operational performance can be avoided (Eggemeier, Shingledecker, and Crabtree, 1985). **Workload assessment should occur as early as possible in the design process and be repeated at regular intervals.**

During the design phase, workload assessment identifies opportunities for improving the user interface (Fadden, 1993). For example, cumbersome data-entry procedures can increase perceived workload and produce negative impacts on performance. Early identification of such problems will lead to more acceptable resolutions than would be possible if they were discovered later in the development process when changes are costlier and more difficult to make.

During the development process, workload assessment provides a structured method for examining various issues (Fadden, 1992). In new automated systems for ATC, for example, an important workload issue is the amount of data entry required of controllers. In concepts for even more advanced ATC automation, a significant workload issue centers on the controller's need to evaluate machine-generated conflict resolutions. Workload issues such as these should be examined again and again during the software development process (in parallel with new software releases), not left until it is too late to make the changes needed to keep perceived workload in the moderate range.

Useful predictions can be made about potential workload levels prior to Operational Test and Evaluation (OT&E). Once the task analysis has been completed and information requirements identified, scenarios can be built for purposes of predicting workload. As discussed in Section 8.3, various predictive methods and tools can be used to support workload prediction.

8.2.3 Why is there a need for baseline data on workload and performance?

Results of a workload (or performance) analysis will be difficult to interpret unless there is an established baseline for comparison. For example, if the workload effects of a new system design are of interest, valid data for workload in the current system are needed for the sake of comparison. Without a basis for comparison or a baseline criterion, it will be difficult to identify the workload effects of the new design. It will be possible to measure workload but not to say whether workload has decreased, stayed the same, or increased. When evaluating design products for their effects on workload, it is important to focus on changes in workload both across systems and over time.

8.2.4 Why should demands on memory and attention be measured?

In general, when given a choice, people tend to prefer options that reduce loading on short-term memory and that require lower investments in terms of attention and effort. That is, people want to conserve their mental resources and will choose less "costly" alternatives over more effortful ones. This means that ATC procedures that generate less workload than others tend to be used more frequently as task demands increase (Sperandio, 1978).

For example, in situations where several options are available, controllers typically prefer not to vector one aircraft around another because of the need to re-establish the vectored aircraft back on course as well as determine whether or not the pilot has complied with instructions. Other control actions are considered to be less costly than vectoring, in terms of the follow-up workload required.

Boredom may occur when there is little activity: the remedy is to provide more work. Boredom may occur when there is substantial activity but it has all become routine, requiring little effort and devoid of challenge and interest: the remedy is to maintain direct and active involvement in the control loop. Boredom tends to increase as skill and experience increase: the remedy is to design tasks with a hierarchy of required skills, since opportunities to exercise high-level skills can help to prevent boredom."
ICAO, 1993, p. 34

The design should be assessed to determine how it helps controllers remember what they have to by asking them to recognize—not recall—the information. Total dependence on unaided recall of information can increase error rates and task completion times.

Evaluation plans should investigate the burden on short-term, working memory as well as the other "costs" associated with meeting task demands in a new ATC environment. Since error is associated with what a controller is likely to forget, evaluators should try to identify instances of forgetting that might be induced by the new design. These costs can be compared with similar measures taken in the previous environment. When the old and new systems are compared on memory-and-forgetting costs, the difference will be an indicator of the new system's operational acceptability. Alternative designs for the new system can be compared in a similar way.

Research conducted at the Civil Aeromedical Institute (CAMI) indicates that combining "repetitive, monotonous work with requirements for high alertness, continuous and rapid decisions, and various penalties for any error that occurs, may very well represent a combination that is quite stressful." (Thackray, 1980, p. 10). Melton (1982) points out, however, that ATC work is no more stressful than work in many other occupations. Low workload can lead to monotony, boredom, complacency, and the strain of not having enough to do.

If required to work under these low workload conditions for more than an hour (Thackray and Touchstone, 1982), the highly trained and highly motivated controller's attention can wander and be captured by some extraneous element in the environment (Jorna, 1992). The fact that this happens is not the controller's "fault." It is the outcome to be expected from knowledge of human limitations under low workload. Giving "controllers as much freedom as possible to control and schedule their own workload" is one way to address the problem of boredom (ICAO, 1993, p. 35).

Ideally, evaluations should include low-workload conditions that are in effect continuously for 90 minutes or more. Evaluation plans should also investigate required levels of alertness, monotony, and penalties for error, not only singly, but in combination. In the area of attention, the objective is to design tasks and functions for the controller that will stimulate a fairly constant, moderate level of alertness. Long periods when little is required are just as potentially harmful as are long periods that require total, intensive attention.

One challenge to achieving the objective of moderate workload for the controller is that demands for ATC services tend to bunch together. Peaks and valleys in traffic are the reality. What this implies is that requirements for workload-smoothing tools and techniques need to be specified and met by designers. Various aiding approaches can lower the peaks, not by decreasing the number of aircraft lined up for ATC services, but by decreasing the amount of time and effort required per aircraft and by shifting some tasks to the computer. Other aiding techniques can raise the valleys and alleviate the boredom by either shifting tasks from the computer to the controller or by providing on-line training or other meaningful tasks.

8.2.5 Why should demands for manual data entry be measured?

A specific issue, related to both workload and performance, is the possibility of increased requirements for data entry by the controller. Members of FAA-chartered controller teams have voiced this concern in their pleas for no increase in keyboard data entry. Data entry imposes both objective, physical and subjective mental demands on the controller.

Physical demands result from the number of keystrokes or other specific inputs that must be made. Subjective demands arise from requirements to recall the appropriate form, content, and sequencing of data entries. In a time-pressured situation, recall of infrequently used commands can be expected to be slower and less accurate, with resulting slower and less accurate job performance.

Manual data entry takes time, energy, and attention away from the ATC situation and the controller's major task, the separation of aircraft. In the tower, manual data entry also requires the controller's focus to change (accommodate) from far-away aircraft to the very close keyboard. Such changes in accommodation and shifts in attention are costly. Each can increase the time needed to process important information. Even though this usually presents no problem, such costs cannot be ignored. In fact, the controller team for increased tower automation recommended a portable hand-held input device to allow the tower controller to maintain far accommodation while entering data without having to look down.

Data entry is essentially a secondary task that must be timeshared with primary tasks. **A design that imposes heavy demands for manual data entry will likely require extensive re-design to achieve operational suitability.** If manual data entry is necessary, it should promote the controller's interest, that is, serve to support a primary task, not make the controller a slave to the computer.

Another word of caution is that the amount of data entry should not be the sole focus of a design evaluation. It is just one of many factors that need to be assessed. **Data entry should be considered within the total context of all demands imposed on the controller.** Focusing on data entry alone is likely to mean that other critical design flaws will be overlooked. The evaluation plan should be balanced so that results make realistic estimates of overall task demands and performance requirements.

8.3 EVALUATING CHANGES INTRODUCED BY DESIGN PRODUCTS

The human-factors consequences of technological innovations need to be evaluated in order to determine their usefulness for ATC and to resolve any problems (ICAO, 1993). This section poses and answers key questions that can guide the evaluation of design products, such as mock-ups, prototypes, and other demonstration systems.

8.3.1 How should workload and performance be addressed in a plan to evaluate a proposed design or prototype?

If designers and evaluators consider only observable task demands and ignore the wide range of possible variation in perceived mental workload, the final design products (displays, controls, etc.) may be based on false assumptions about acceptable workload. **Evaluation plans should include assessments of information-processing workload as well as assessments of observable task demand.** It is important to identify concurrent workload, that is, not just on a task by task basis, but considering overlapping tasks. Individual tasks may be manageable, but when they must be performed together, workload may become excessive.

The evaluation plan should address the effects of proposed designs on the performance of information-processing tasks and the overall quality of planning and decision making. Support for planning and designing evaluations should be provided by human factors personnel who are familiar with evaluation procedures and available measures. There are many different procedures and measures available for measuring workload and performance. Since the use of inappropriate measures will give meaningless results, the procedures and measures selected need to be chosen very carefully. This chapter provides information that will be useful in determining the most appropriate measures for the investigation at hand.

8.3.2 How do design changes affect the nature of the controller's job? Are the changes beneficial?

Research conducted in various complex, automated environments indicates that new technology generally shifts the focus of workload from manual tasks to information-processing tasks, such as evaluating and understanding the status of automated processes (e.g., Goodstein, 1981; Parasuraman, 1987; Parasuraman, Molloy, and Singh, 1993; Wickens and Kessel, 1981). Certain increases in automation would not change the nature of the controller's task. For example, automation could be used to predict when sectors should be combined or decombined, or when additional assistance (e.g., a D-side controller or air traffic coordinator) should be called in. However, other increases in automation may distance the controller from the airspace, causing their situational awareness to deteriorate. In time, skills that go unused, do to the automation performing the task, will also deteriorate. For example, continued reliance upon computer-generated conflict resolutions may result in controllers being less ready to respond when a conflict cannot be solved by the computer. (These and other issues are discussed in detail in Chapter 6, *Issues in ATC Automation.*)

As automation increases, the controller becomes a system monitor or system manager. This means that the controller in a highly automated system will intervene only when the automation breaks down or is unable to compute satisfactory "answers" to problems. Under routine conditions at this level of automation, the controller essentially supervises the computer.

Although designed to reduce physical workload, increased automation often introduces new sources of mental workload and new sources of error. Accordingly, the controller's role and operational activities can be expected to change with the introduction of any new technology, especially one that increases the level of automation.

The nature of such changes can be determined by comparing the allocation of functions in the old system to the allocation of functions in the new system. This comparison should be made to determine whether an active, involved role remains for the controller. Changes may be beneficial in relieving the controller of repetitive, tedious tasks that are better allocated to the computer. They may be detrimental, however, if tasks allocated to the controller largely require monitoring of automated functions and acceptance or rejection of resolutions offered by the computer.

Although not currently implemented or planned for ATC, an approach known as **dynamic allocation** might benefit the controller by re-assigning some tasks to the computer when the controller is busy, and by re-assigning other tasks to the controller when workload is light. This approach assumes some level of collaboration between the controller and the computer. In the future, various collaborative, dynamic-allocation schemes may be considered. **A key point is that the plan for re-assigning tasks should be under human control, not under computer control.**

Determining the relative benefit of system changes requires testing and evaluation. It cannot simply be assumed that benefits will occur. Further, testing on a subsystem-by-subsystem basis is unsatisfactory. When subsystems are combined, their effects on the controller's role may be quite different from their separate effects. **Testing must be designed to identify new sources of mental workload, new sources of error, and unacceptable levels of demand for the fully-integrated system.**

8.3.3 How is workload measured in an evaluation?

As we have noted, it is important to ensure that the demands imposed by new ATC technology do not exceed the natural human capabilities of air traffic controllers. Likewise, it is the team's responsibility to ensure that controllers' abilities are not underused or underloaded as a result of system changes. **Both overload and underload can impact job performance and system effectiveness.**

Human factors specialists who support the acquisition team should be well acquainted with methods and measures for evaluating workload. They should also be familiar with software tools that are available to support selection and application of workload measures (e.g., Casper, Shively, and Hart, 1987; Harris, Hill, and Lysaght, 1989).

Selecting measures of workload requires knowledge of the available measures and the aspects of workload they aim to measure. Decisions that require workload estimations should not be made on the basis of results from one workload measurement technique alone. **Investigators should use a set**

of workload measures designed to tap into workload from different perspectives. One recommendation is to use task analytic methods to identify the "major components of tasks, followed by a battery of performance-based measures designed to evaluate the load on each [task] component" (Gopher and Donchin, 1986, pp. 41-43).

- *Measures of Task Demand or Objective Workload.*

  Several objective measures of ATC workload are available. Taking a simple count of the number of controlled aircraft is one indication of objective workload. However, airspace characteristics are also critical (Redding, Cannon, and Seamster, 1992).

  Some airspace characteristics that are important to consider when estimating workload are:

  - sector size
  - number of altitudes assigned to the sector
  - sector operations or functions (e.g., high altitude overflights vs. altitude transitions or terminal operations)
  - procedural constraints (e.g., traffic management).

  An analysis of recent data on sector complexity provide empirical support for the notion that key airspace characteristics also include the relative frequency of complex aircraft routings in the sector, the need for sequencing and spacing arrivals and departures, and the amount of radio frequency congestion during peak traffic periods (Mogford, Murphy, and Guttman, 1993).

  During simulations, objective workload can be tracked automatically by internal software. These include key strokes, number of aircraft, transition times of aircraft passing through sectors, and number and duration of ground-air communications.

  Observational techniques can also be used to document objective workload. Trained observers can tabulate the occurrences of various controller actions during defined time periods. Video and audiotaping are further methods that can provide records of controller activities for analysis. Each method has benefits and limitations that can be investigated by human factors support personnel.

It should be noted that there is no direct, simple relationship between number of errors (or error rate) and workload. Errors occur just as frequently under low workload as they do under high workload (Fadden, 1992; Kantowitz and Casper, 1988; Morris and Rouse, 1988; Office of Aviation Medicine, 1986; Swain and Guttman, 1983). Therefore, it is never wise to use error rate as the only indicator of workload.

Secondary tasks in ATC include updating flight progress strips, managing the display of data blocks, and some optional communication/coordination tasks. Performance on these secondary tasks provides an indirect measure of the attentional and information-processing demands of the primary task. Using this approach, test participants perform additional tasks in parallel with their primary task. (In ATC, primary tasks are those associated with maintaining aircraft separation.) A decline in performance on the secondary task is thought to indicate increased demands for resources by the primary task, that is, an increase in workload. Use of secondary task measures in design evaluation can help evaluators to estimate how much workload controllers will experience with a particular design.

Interpreting the results of secondary task measures can be a little tricky. The straightforward approach is that no decline in secondary task performance means low workload on the primary task. This may not always be the case. A finding of no decline in secondary task performance cannot be interpreted as fool-proof evidence of low workload on the primary task. Continued good performance on the secondary task may simply mean that performance of the particular primary task chosen did not deplete the resources required by the secondary task used. **When secondary tasks are used to measure workload, several different secondary tasks should be used.** Each secondary task should demand different combinations of resources so that the origins and levels of primary task workload can be determined more accurately (Hart and Wickens, 1990).

Physiological measures were used in a series of studies of controller stress conducted during the 1970s at CAMI (e.g., Melton, Plis, Hoffmann, and Saldivar, 1973). These studies are reviewed by Melton (1982) and listed at the end of this chapter under the suggestions for further reading. The measures used included heart rate, biochemical changes, and galvanic skin response. Other commonly used measures include heart rate variability, blood pressure, brain activity, breathing rate, and various optical measures (e.g., pupil diameter, eye movements, and eye blinks); they do not measure workload directly, rather, they measure the body's response to the stress induced by the workload. These responses are very individual. For example, while one person may respond to a high level of workload with an increase in heart rate, another person may show no such change.

In general, when physiological measures are used, they should be used within a battery of measures, to supplement the information obtained from performance measures and subjective ratings (Hart and Wickens, 1990). A comprehensive approach is one that involves comparisons across primary task performance measures, secondary tasks that are a natural part of the job, subjective measures, and physiological measures.

For example, researchers at CAMI are currently conducting an investigation on the effects of shiftwork and age, using a battery of measures including performance measures (response time and accuracy), heart rate, temperature, biochemical measures, subjective ratings of mood and fatigue, and a workload rating scale (Della Rocco, 1991). Participants are also keeping logs of their activities, such as eating and sleeping. During study sessions, participants perform job tasks resembling those performed by pilots and air traffic controllers, including mental arithmetic, problem solving, pattern discrimination, and critical tracking. A secondary monitoring task provides another index of workload. Performance is measured under varying demand conditions.

While such an exhaustive approach to workload measurement is not feasible for most system evaluations, it is important to understand the advantages and limitations of available workload measures. Workload and performance measures used in an evaluation need to be carefully chosen. The results of these tests need to be analyzed and interpreted with an understanding of the validity and limitations of the measures used.

There is some disagreement in the research literature about the applicability of specific physiological measures for different workload levels and for short-term versus long-term assessment. Before using any of these measures, human factors support personnel should review the specific research methods and findings associated with the measures being considered.

A practical consideration is that physiological measures vary in their intrusiveness, that is, the extent to which any equipment needed to collect the measures interferes with task performance. (Equipment worn for this purpose may include optical measurement devices or electrodes attached to the skin.) **The less intrusive the measurement technique, the better will the results reflect true workload values.**

- *Measures of Perceived Workload.*

Performance indicators, such as speed and accuracy, are sometimes used as measures of perceived workload, that is, the assumption is made that changes in task performance reflect changes in workload. It might be presumed, for example, that quick and accurate performance indicates low perceived workload. To some extent, however, speed and accuracy depend on other factors, such as motivation, effort exerted, and fatigue (Meister, 1985). Other influences on task performance include task difficulty, time available, and operational strategies. High performance scores may occur under conditions of low workload, but they may also occur when test participants have exerted extreme effort under high-workload conditions (Hart and Wickens, 1990).

In ATC, primary tasks include monitoring for separation, formulating and issuing clearances to resolve conflicts, and sequencing traffic according to flow restrictions. On these tasks, highly motivated controllers will put a great deal of effort into producing fast and accurate results. But, as pointed out previously, it is not a good idea to require 100 percent effort over time. **Acceptable designs will allow the controller to complete tasks and sub-tasks comfortably, while keeping some resources in reserve to handle the unexpected.**

Three basic techniques are commonly used for measuring perceived workload:

- Rating scales
- Secondary task-performance measures
- Measures of internal, physiological activity.

Each of these techniques has its own set of benefits and limitations.

Rating scales are the most direct, the most practical, and the most widely used measures of perceived workload. The FAA Technical Center has developed a workload rating scale for studying air traffic controller workload during operational system testing (Leighbody, Beck, and Amato, 1992). Based on the Pilot Objective/Subjective Workload Assessment Technique (POSWAT), this scale is known as the Air Traffic Workload Input Technique (ATWIT).

In an evaluation of ATWIT, 18 air traffic controllers used modified POSWAT rating descriptions to describe how hard they were working in real-time, during single-position, single-sector en route simulation runs. ATWIT scale values from one (little workload) to nine (excessive workload) were defined according to contributions by four factors:

- The ease of performing ATC tasks
- The need for ATC task prioritization
- The chance for error or omission
- The need for data position coordination.

The ATWIT study found that different controllers had different perceptions of ATC task difficulty but had similar responses to increased traffic volume. Workload ratings increased as traffic volume increased. The authors of this workload study conclude that "use of real-time subjective workload estimates is a useful workload measurement tool in operational system testing" (Leighbody, Beck, and Amato, 1992). Further evaluations of ATWIT are planned.

Rating scales can be administered on-line (like ATWIT) or off-line, following completion of a test session. On-line scales elicit instantaneous responses but are intrusive in that they interrupt the flow of performance. Off-line scales are non-intrusive but rely on recall of workload after the fact.

A major limitation of subjective ratings of perceived ATC workload is the high variability of ratings given by different controllers. High between-rater variability lowers the reliability of the results. If, for example, some controllers rate a particular unit of work as low in perceived workload and others rate it high, these results may be related to skills, experience, or other subjective rater biases not reflective of the effort required for the controller to perform the unit of work. **Whenever ratings are collected, they should be analyzed for inter-rater agreement. Another constraint is that**

many controllers are reluctant to rate any level of workload as too high and as something they can't handle. One way around this problem is to consider supplementing the use of subjective ratings with more objective measures. For example, ATWIT might be used for initial workload screening, with follow-up based on other techniques, such as secondary task performance measures and physiological measures.

Further detail on all types of workload measures is available in several thorough treatments of this topic: Corwin, Sandry-Garza, Biferno, Boucek, Logan, Jonsson, and Tetalis, 1989; Gawron, Schiflett, and Miller, 1989; Hart and Wickens, 1990; Hockey, 1986; Lysaght, Hill, Dick, Plamondon, Linton, Wierwille, Zaklad, Bittner, and Wherry, 1989; Meister, 1985; O'Donnell and Eggemeier, 1986; and Roscoe, 1987. Each of these sources provides extensive references to the research literature on workload assessment. They also provide technical discussions of the advantages and disadvantages of the various measures.

8.3.4 What are some key design issues to address through performance measurement?

Observable job performance is the outcome of the controller's information-processing activities, such as planning, monitoring, and decision making. Therefore, it is extremely important to evaluate a design's effects on these activities. These effects are influenced by the way information is selected, represented, and presented to the controller. If the design withholds or hides necessary information, these crucial information-processing activities may be impaired.

Directly measurable aspects of job performance are speed (e.g., time spent per task), accuracy, and number of job elements completed in a given unit of time. Qualitative aspects of ATC performance are how well safety and order were maintained and how expeditiously traffic was transitioned through the controller's sector. (Note that there is a trade-off relationship between these criteria. For example, safety and order may be somewhat at risk if the sole focus is on increased capacity.) Because maintenance of safety is the primary standard of ATC performance, a design should be evaluated from the safety perspective. The number of operational/system errors is often assumed to be a measure of safety, although minor violations may have no impact on safety. Other measures of safety need to be identified or developed.

Controllers' **situational awareness** is key to their performance, since the perceived situation guides problem identification and problem resolution. **A crucial aspect of design evaluation is assessment of the support the design provides for the controller's situational awareness.** Building and updating the mental picture and using the picture as the basis for projections about the future are thought to be crucial information-processing tasks in maintaining situational awareness (Ammerman, Bergen, Davies, Hostetler, Inman, and Jones, 1988; Endsley, 1988).

Evaluation of how well a prototype system helps controllers maintain situational awareness can be conducted using various techniques. One technique is to ask controllers questions designed to assess their situational awareness. Another is to ask the controller to turn away from the visual display and sketch the current and projected ATC situation (e.g., Means, Mumaw, Roth, Schlager, McWilliams, Gagne, Rice, Rosenthal, and Heon, 1988; Schlager, Means, and Roth, 1990).

Timesharing of several tasks is common in air traffic control. The controller may be checking altitude information displayed in data blocks, listening to a pilot's request for an altitude change, and manually entering a flight plan amendment. Timesharing is possible because of the controller's ability to process different kinds of information in parallel. At some point, timesharing comes up against the limits of parallel processing. As timesharing demands increase, the controller may find it necessary to perform only the most critical tasks in parallel and put less critical tasks "on hold," at least for brief periods.

Evaluations of designs for their timesharing requirements need to consider the relationships among workload, effort, and performance. **Evaluation teams should be skeptical of any design that requires continuous timesharing of more than a few tasks of low-to-moderate difficulty.** In general, timeshared tasks should not draw upon the same resources. For example, timesharing visual, auditory, and manual tasks is usually preferable to timesharing two visual tasks or two auditory tasks and a manual task (Wickens, 1984, 1989). Timesharing of tasks that draw on the same resources or capacities risks overload and degraded performance.

Under high time pressure, the controller may be able to focus only on the immediate problem and simply "not see" problems developing elsewhere. If timesharing demands are heavy, a controller may forget his/her intention to issue a clearance to descend an aircraft because s/he has become busy with another task. Both examples are performance outcomes that can be predicted from what we know about cognitive narrowing or tunnel vision under time pressure (Sheridan, 1981).

Task sequencing should be considered a design issue in its own right. Different approaches to sequencing tasks and task types, one after another, can have different effects on workload and performance. If a long period of monitoring (i.e., greater than 30 minutes) is followed by a conflict resolution task, for example, the controller's level of alertness will have to shoot up suddenly. If the monitoring tasks are timeshared with other kinds of tasks, the initial alertness level will be higher.

"Many forms of automated assistance in ATC may have the unintended effect of increasing boredom."
ICAO, 1993, p. 35

An important consideration in task sequencing is variety. Having a variety of tasks to perform is especially necessary during periods of low traffic, to maintain the controller's readiness to respond. Giving the controller some control over task sequencing, through dynamic task allocation or other techniques, may be an approach to investigate.

Task sequencing and timesharing should not occur as side-effects of a design but should be considered

proactively by system designers. Both are design issues in the same sense that software functionality is a design issue. The following guidelines can help to achieve moderate task loading in relation to time (Fadden, 1993):

- Make any procedural task sequence interruptable at any point to accomplish time- or event-driven actions.
- Avoid abrupt changes in normal task loading.
- Minimize the need for precisely timed tasks.
- Where task start-time constraints are necessary, relax task completion-time requirements.
- Similarly, where task completion-time constraints exist, make the start-time requirements flexible.

The objective is to minimize opportunities for error arising from poorly designed requirements for task sequencing and timesharing.

When performance for different groups is being compared, it is important to show that perceived workload does not differ from group to group. We know that controllers can put forth extra effort (i.e., sustain heavy information-processing workload) and get the job done efficiently, even under degraded system conditions, but system design should not require them to work at that level as a matter of course.

To support interpretation of performance outcomes within an overall context, evaluators should determine the relationship between performance on concurrent ATC primary tasks and other workload measures (e.g., subjective ratings; secondary task measures). Looking at performance on concurrent or overlapping tasks is important from the perspective of adding realism and credibility to the evaluation and resulting recommendations for design changes. (Detailed guidance on performance measurement is provided in Chapter 10, Human Factors Testing and Evaluation.)

8.3.5 How will design elements affect job satisfaction?

Even when performance standards can be maintained, job satisfaction may be adversely affected by a design. We know, however, that highly motivated and self-disciplined employees, such as controllers, can maintain performance standards under highly adverse conditions that affect job satisfaction. These adverse, workload-inducing conditions may range all the way from glare and reflections on the display screen to incompatible display-control relationships. **The problem is that, over time, adverse environmental, workstation, and computer-human interface elements can have subtle negative effects on workforce performance, health, and well-being.**

Therefore, system designs and their proposed implementation should be evaluated for their potential adverse effects on job satisfaction over time.

8.3.6 How will this design affect on-site ATC team interactions?

The whole area of ATC team interactions in the field is one that has been left largely unexplored by any formal investigation. There is not even a clear definition of an "ATC team," much less a definition of ATC-team workload or ATC-team performance. To some, the team concept is limited, for example, to just the radar and data controllers working an en-route sector or to the ground controller and local controller in a tower. The concept might apply, however, to an entire area with one controller to a sector or to a facility, with all the facility personnel as a team made up of sub-teams. The entire airspace system can be envisioned as a network of such ATC teams. The entire aviation-system team includes both pilots and controllers, as well as Airway Facilities (AF) personnel and management. **The ways in which team members interact and communicate with each other can be affected by design concepts and their implementation.**

> "An incidental consequence of various forms of computer assistance can be a reduction in team roles and functions."
>
> ICAO, 1993, p. 38

Despite the lack of attention to ATC teams *as teams*, few would dispute the statement that on-the-job teamwork is critical to ATC success. Because design elements can support or impair ATC teamwork, they should be evaluated from this perspective. In particular, the design of coordination tasks and procedures can impact the workload and performance of operational controller teams.

The design should be evaluated for its ability to meet the job needs of controller teams in the following areas (FAA, 1990):

- Information about the ATC situation
- Presentation of information
- Transfer of information between team members
- Procedural efficiency and effectiveness
- Balance of workload among team members
- Support for team decision making
- Changes in team roles due to automation.

The objectives are to identify potential design flaws and to meet the job needs of both controller teams and individual controllers. Re-design of tasks, jobs, and system components during early phases of system development can help to ensure system acceptability and operational suitability.

8.4 REFERENCES AND SUGGESTIONS FOR FURTHER READING

Aldrich, T. B., and Szabo, S. M. (1986). A methodology for predicting crew workload in new weapon systems. *Proceedings of the Human Factors Society* (pp. 653-657). Santa Monica, CA: Human Factors Society.

Ammerman, H. L., Bergen, L. J., Davies, D. K., Hostetler, C. M., Inman, E. E., and Jones, G. W. (1988). *FAA air traffic control operations concepts* (DOT/FAA/AP-87-01, Vols. I - VIII, e.g., Vol. VI: ARTCC/HOST En Route Controllers, Change 1). Washington, DC: Department of Transportation/Federal Aviation Administration.

Bainbridge, E. L. (1987). Ironies of automation. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.), *New technology and human error* (pp. 271-283). New York: Wiley.

Bikson, T. K. (1987). Cognitive press in computer-mediated work. In G. Salvendy, S. L. Sauter, and J. J. Hurrell, Jr. (Eds.), *Social, ergonomic, and stress aspects of work with computers* (pp. 353-364). Amsterdam: Elsevier.

Boff, K. R., and Lincoln, J. E. (1988). *Engineering data compendium* (Vol. II, Human perception and performance, Section 7.7, Workload characteristics). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Bruce, K., Murphy, E. D., and Stewart, L. J. (1987). Cognitive task analysis: Capturing expertise for rule-base development. *Abridged Proceedings of the Second International Conference on Human-Computer Interaction.* Amsterdam: Elsevier.

Buckley, E. P., DeBaryshe, B. D., Hitchner, N., and Kohn, P. (1983). *Methods and measurements in real-time air traffic control system simulation* (DOT/FAA/CT-83/26). Atlantic City Airport, NJ: Department of Transportation/Federal Aviation Administration Technical Center.

Buckley, E. P., O'Connor, W. F., and Beebe, T. (1969). *A Comparative analysis of individual and system performance indices for the air traffic control system* (Final Report). Atlantic City, NJ: Department of Transportation/Federal Aviation Administration National Aviation Facilities Experimental Center.

Cain, L. G. (1989). *Initial Sector Suite System (ISSS) interim workload scenario description* (MTR-89W00006). McLean, VA: MITRE.

Casali, J. G., and Wierwille, W. W. (1983). A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human Factors, 25*, 623-641.

Casper, P. A., Shively, R. J., and Hart, S. G. (1987). Decision support for workload assessment: Introducing WC FIELDE. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1145-1149). Santa Monica, CA: Human Factors Society.

Chiles, W. D. (1982). Workload, task, and situational factors as modifiers of complex human performance. In E. A. Alluisi, and E. A. Fleishman (Eds.), *Human performance and productivity: Stress and performance effectiveness* (pp. 1-56). Hillsdale, NJ: Erlbaum.

Corwin, W. H. (1992). In-flight and postflight assessment of pilot workload in commercial transport aircraft using the Subjective Workload Assessment Technique. *International Journal of Aviation Psychology, 2*, 77-94.

Corwin, W. H., Sandry-Garza, D. L., Biferno, M. H., Boucek, G. P., Jr., Logan, A. L., Jonsson, J. E., and Metalis, S. A. (1989). *Assessment of crew workload measurement methods, techniques and procedures* (Vol. 1 - Process, methods, and results; WRDC-TR-89-7006). Wright-Patterson Air Force Base, OH: Air Force Systems Command, Wright Research and Development Center, Cockpit Integration Directorate.

Costa, G. (1991). Shiftwork and circadian variations of vigilance and performance. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and system issues in air traffic control* (NATO Advanced Science Institute Series F: Computer and Systems Sciences, Vol. 73, pp. 267-280). New York: Springer-Verlag.

CTA INCORPORATED. (1991). *Analysis of corresponding ISSS and NAS commands for keystrokes (phase I) and quick entry actions (phase II)* (Report Number T104). Englewood, CO: CTA.

Danaher, J. W. (1980). Human error in ATC system operations. *Human Factors, 22*, 535-545.

Della Rocco, P. (1991). *Fatigue and performance: Shiftwork in controllers of varying age* (Dissertation Prospectus).

Derrick, W. (1988). Dimensions of operator workload. *Human Factors, 30*, 95-110.

Dorner, D. (1987). On the difficulties people have in dealing with complexity. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.), *New technology and human error*. New York: John Wiley.

Eggemeier, F. T. (1981). Current issues in subjective assessment of workload. *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 513-517). Santa Monica, CA: Human Factors Society.

Eggemeier, F. T., Shingledecker, C. A., and Crabtree, M. S. (1985). Workload measurement in system design and evaluation. *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 215-219). Santa Monica, CA: Human Factors Society.

Empson, J. (1987). Error auditing in air traffic control. In J. A. Wise and A. Debons (Eds.), *Information systems: Failure analysis* (NATO Advanced Science Series, Vol. F32, pp. 191-198). New York: Springer-Verlag.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1990). Predictive validity of an objective measure of situation awareness. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 41-45). Santa Monica, CA: Human Factors Society.

Ericsson, K. A., and Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Federal Aviation Administration. (1990). *National plan for aviation human factors* (Draft). Washington, DC: Department of Transportation/Federal Aviation Administration.

Fadden, D. M. (1993) Workload assessment. In K. Cardosi (Ed.), *Human factors for operational personnel* (Report No. DOT/FAA/RD-93/5). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Finkelman, J. M., and Kirschner, C. (1980). An information- processing interpretation of air traffic control stress. *Human Factors, 22*, 561-568.

Gawron, V. J., Schiflett, S. G., and Miller, J. C. (1989). Measures of in-flight workload. In R. S. Jensen (Ed.), *Aviation psychology* (pp. 240-287). Brookfield, VT: Gower Technical.

Geer, C. W. (1981). *Human engineering procedures guide* (AFAMRL-TR-81-35). Wright-Patterson Air Force Base, OH: Armstrong Aeromedical Research Laboratory.

Goettl, B. P. (1991). Attention, workload, and automation. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and system issues in air traffic control* (NATO Advanced Science Institute Series F: Computer and Systems Sciences, Vol. 73, pp. 293-298). New York: Springer-Verlag.

Goodstein, L. P. (1981). Discriminative display support for process operators. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp. 433-449). New York: Plenum Press.

Gopher, D., and Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors, 26*, 519-532.

Gopher, D., and Donchin, E. (1986). Workload—An examination of the concept. In K. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 41-1 - 41-49). New York: Wiley.

Guralnik, D. B. (Ed.) (1984). *Webster's new world dictionary* (2nd college edition). New York: Simon and Schuster.

Hancock, P. A., and Chignell, M. H. (1988). Mental workload dynamics in adaptive interface design. *IEEE Transactions on Systems, Man, and Cybernetics, 18*, 647-658.

Hancock, P. A., and Meshkati, N. (Eds.). (1988). *Human mental workload.* Amsterdam: North Holland.

Harris, R. M., Hill, S. G., and Lysaght, R. J. (1989). OWLKNEST: An expert system to provide operator workload guidance. *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1486-1490). Santa Monica, CA: Human Factors Society.

Hart, S. G. (1987). *Research papers and publications (1981-1987): Workload research program* (NASA Technical Memorandum 100016). Moffett Field, CA: NASA-Ames Research Center.

Hart, S. G. (1986). Theory and measurement of human workload. In J. Zeider (Ed.), *Human productivity enhancement: Training and human factors in systems design* (Vol. I, pp. 396-456). New York: Praeger.

Hart, S. G., and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North Holland.

Hart, S. G., and Wickens, C. D. (1990). Workload assessment and prediction. In H. R. Booher (Ed.), *MANPRINT: An approach to systems integration* (pp. 257-296). New York: Van Nostrand Reinhold.

Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner, A. C., Jr., Zaklad, A. L., and Christ, R. E. (1992). Comparison of four subjective workload rating scales. *Human Factors, 34,* 429-440.

Holt, R. W. (1982). *Theory, evidence, inference: A handbook on the scientific method.* Washington, DC: University Press of America.

Hopkin, V. D. (1988). Air traffic control. In Wiener, E. L., and Nagel, D. C. (Eds.), *Human factors in aviation* (pp. 639-664). New York: Academic Press.

Hopkin, V. D. (1982). *Human factors in air traffic control* (AGARD-AG-275). Neuilly Sur Seine, France: NATO, Advisory Group for Aerospace Research and Development (AGARD).

Hopkin, V. D. (1979). Mental workload measurement in air traffic control. In N. Moray (Ed.), *Mental workload: Its theory and measurement* (pp. 381-385). New York: Plenum Press.

Hopkin, V. D. (1992). Verification and validation: Concepts, issues, and applications. *Verification and validation of complex and integrated human-machine systems* (Preprints of NATO Advanced Study Institute). Daytona Beach, FL: Embry-Riddle Aeronautical University.

Human Technology, Inc. (1990). *Cognitive task analysis of prioritization in air traffic control* (Vol. 1). McLean, VA: Author.

Jahns, D. W. (1973). *A concept of operator workload in manual vehicle operations* (Report No. 14). Meckenheim, Germany: Forschungsinstitut fur Anthropotechnik.

Janis, I. L. (1982). Decision making under stress. In L. Goldberger and S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects*. New York: Macmillan's Free Press.

Johannsen, G., Rjinsdorp, J. E., and Sage, A. P. (1983). Human system interface concerns in support system design. *Automatica, 19*, 595-603.

Jorna, P. G. A. M. (1992). Operator workload as a limiting factor in complex systems. *Verification and validation of complex and integrated human-machine systems* (Preprints of NATO Advanced Study Institute). Daytona Beach, FL: Embry-Riddle Aeronautical University.

Kahn, M. J., Tan, K. C., and Beaton, R. J. (1990). Reduction of cognitive workload through information chunking. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 1509-1513). Santa Monica, CA: Human Factors Society.

Kantowitz, B. H., and Casper, P. A. (1988). In E. L. Weiner and D. C. Nagel (Eds.), *Human factors in aviation* (pp. 157-187). New York: Academic Press.

Keppel, G., and Saufley, W. H., Jr. (1980). *Introduction to design and analysis.* New York: W. H. Freeman.

Kerr, B. (1973). Processing demands during mental operations. *Memory and Cognition, 1*, 401-412.

Kinney, G. C. (1977). *The Human element in air traffic control: Observations and analyses of the performance of controllers and supervisors in providing ATC services* (MITRE Technical Report MTR-7655). McLean, VA: MITRE.

Kirk, R. E. (1968). *Experimental design: Procedures for the behavioral sciences.* Monterey, CA: Brooks/Cole.

Kitchin, J. B., and Graham, A. (1974). Mental loading of process operators: An attempt to devise a method of analysis and assessment. In E. Edwards and F. P. Lees (Eds.), *The Human operator in process control* (pp. 105-119). London: Taylor & Francis.

Langan-Fox, C. P., and Empson, J. A. C. (1985). 'Actions not as planned' in military air-traffic control. *Ergonomics, 28*, 1509-1521.

Leighbody, G., Beck., J., and Amato, T. (1992). An operational evaluation of air traffic controller workload in a simulated en route environment. *37th Annual Air Traffic Control Association Conference Proceedings* (pp. 122-130). Arlington, VA: Air Traffic Control Association.

Leplat, J. (1988). Task complexity in work situations. In L. P. Goodstein, H. B. Anderson, and S. E. Olsen (Eds.), *Tasks, errors and mental models*. New York: Taylor & Francis.

Loftus, G. R., Dark, V. J., and Williams, D. (1979). Short-term memory factors in ground controller/pilot communication. *Human Factors, 21*, 169-181.

Lysaght, R. J., Hill, S. G., Dick, A. O., Piamondon, B. D., Linton, P. M., Wierwille, W. W., Zaklad, A. L., Bittner, A. C., Jr., and Wherry, R. J. (1989). *Operator workload: Comprehensive review and evaluation of operator workload methodologies* (Technical Report 851). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Mandler, G. (1982). Stress and thought processes. In L. Goldberger and S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (pp. 88-104). New York: Macmillan's Free Press.

McKinley, J. B., and Jago, R. J. (1985). Potential controller error modes in Automated En Route ATC (AERA). *Proceedings of the Third Symposium on Aviation Psychology* (pp. 189-197). Columbus, OH: Ohio State University.

Means, B., Mumaw, R., Roth, C., Schlager, M., McWilliams, E., Gagne, E., Rice, V., Rosenthal, D., and Heon, S. (1988). *ATC training analysis study: Design of the next-generation ATC training system* (OPM Work Order 342-036). Richmond, VA: Human Resources Research Organization (HumRRO).

Meister, D. (1985). *Behavioral analysis and measurement methods*. New York: Wiley.

Melton, C. E. (1982). *Physiological stress in air traffic controllers: A review* (FAA-AM-82-17). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Melton, C. E., and Bartanowicz, R. S. (1986). *Biological rhythms and rotating shift work: Some considerations for air traffic controllers and managers* (DOT/FAA/AM-86-2). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

Melton, C. E., Jr., McKenzie, J. M., Polis, B. D., Funkhouser, G. E., and Iampietro, P. F. (1971). *Physiological responses in air traffic control personnel: O'Hare Tower* (FAA-AM-71-2). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Melton, C. E., McKenzie, J. M., Polis, B. D., Hoffmann, M., Saldivar, and J. T., Jr. (1973). *Physiological responses in air traffic control personnel: Houston Intercontinental Tower* (FAA-AM-73-21). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Melton, C. E., McKenzie, J. M., Saldivar, J. T., Jr., and Hoffmann, S. M. (1974). *Comparison of Opa Locka Tower with other ATC facilities by means of a biochemical stress index* (FAA-AM-74-11). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Melton, C. E., McKenzie, J. M., Smith, R. C., Polis, B. D., Higgins, E. A., Hoffmann, S. M., Funkhouser, G. E., and Saldivar, J. T. (1973). *Physiological, biochemical, and psychological responses in air traffic control personnel: Comparison of the 5-day and 2-2-1 shift rotation patterns* (FAA-AM-73-22). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Melton, C. E., Smith, R. C., McKenzie, J. M., Hoffmann, S. M., and Saldivar, J. T. (1976). *Stress in air traffic controllers: Effects of ARTS-III* (FAA-AM-76-13). Washington, DC: Department of Transportation/Federal Aviation Administration Office of Aviation Medicine, Civil Aeromedical Institute; also appeared in *Aviation, Space, and Environmental Medicine* (Sept. 1976, pp. 925-930).

Melton, C. E., Smith, R. C., McKenzie, J. M., Wicks, S. M., and Saldivar, J. T. (1977). *Stress in air traffic personnel: Low-density towers and flight service stations* (FAA-AM-77-23). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Mertens, H. W., Higgins, E. A., and McKenzie, J. M. (1983). *Age, altitude, and workload effects on complex performance* (FAA-AM-83-15). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Mogford, R. H., and Guttman, J. (1992). Unpublished data analysis.

Mogford, R. H., Harwood, K., Murphy, E. D., and Roske-Hofstrand, R. J. (1994). *Review and evaluation of applied research techniques for documenting cognitive processes in air traffic control* (Technical Note DOT/FAA/CT-TN93/39). Atlantic City International Airport, NJ: Department of Transportation/Federal Aviation Administration Technical Center.

Mogford, R. H., Murphy, E. D., and Guttman, J. A. (1993). The Use of direct and indirect techniques to study airspace complexity factors in air traffic control. *Proceedings of the First Mid-Atlantic Human Factors Conference* (pp. 196-202). Norfolk, VA: Old Dominion University.

Moray, N. (Ed.). (1979). *Mental workload: Its theory and measurement* (NATO Conference Series, Vol. 8). New York: Plenum Press.

Moray, N. (1982). Subjective mental workload. *Human Factors, 24*, 25-40.

Moray, N., Dessouky, M. I., Kijowski, B. A., and Adapathya, R. (1991). Strategic behavior, workload, and performance in task scheduling. *Human Factors, 33*, 607-629.

Morris, N. M., and Rouse, W. B. (1988). *Human operator response to error-likely situations in complex engineering systems* (NASA CR-177484). Washington, DC: National Aeronautics and Space Administration.

Moursand, J. P. (1973). *Evaluation: An introduction to research design.* Monterey, CA: Brooks/Cole.

Muckler, F. A., and Seven, S. A. (1992). Selecting performance measures: "Objective" versus "subjective" measurement. *Human Factors, 34*, 441-456.

Murphy, E. D., Reaux, R. A., Stewart, L. J., Coleman, W. D., and Bruce, K. (1989). Where's the workload in air traffic control? *Proceedings of the Fifth International Symposium on Aviation Psychology*. Columbus, OH: Ohio State University.

Myers, J. L. (1979). *Fundamentals of experimental design* (3rd. ed.). Boston: Allyn and Bacon.

Norman, D. A. (1983). *Comments on mistakes in diagnosis: With emphasis on resulting failures to detect the misdiagnosis* (Summary of a talk presented at GA Technologies, Inc.). La Jolla, CA: University of California, Institute for Cognitive Science.

Norman, D., and Bobrow, D. (1975). Data limited and resource limited processing. *Cognitive Psychology, 7*, 44-60.

North, R. A. (1987). A workload index for iterative crewstation evaluation. In H. M. Fiedler (Chair), *Workload assessment techniques and tools* (Symposium proceedings). Dayton, OH: The Workload Sub-Technical Advisory Group of the Department of Defense Human Factors Engineering Technical Advisory Group.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors, 33*, 17-34.

O'Donnell, R. D., and Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman, and J. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 42-1 - 42-49). New York: Wiley.

Office of Aviation Medicine. (1986). *Staff study: ATC operational errors.* Washington, DC: Department of Transportation/Federal Aviation Administration Office of Aviation Medicine.

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors, 29*, 695-706.

Parasuraman, R., Molloy, R., and Singh, I. L. (1993). Performance consequences of automation-induced "complacency." *International Journal of Aviation Psychology, 3*, 1-24.

Pasmooij, C. K., Opmeer, C. H. J. M., and Hyndman, B. W. (1976). Workload in air traffic control. In T. B. Sheridan and G. Johannsen (Eds.), *Monitoring behavior and supervisory control* (pp. 107-118). New York: Plenum Press.

Pollard, J. and Sussman, E. D. (1989). *An examination of sudden acceleration.* (Final Report) DOT-HS-807-367. National Highway Traffic Safety Administration, Office of Defects Investigation, Washington, D. C.

Plutchik, R. (1974). *Foundations of experimental research* (2nd ed.). New York: Harper & Row.

Reason, J. (1990). *Human error.* New York: Cambridge University Press.

Reason, J. (1979). Actions not as planned: The price of automatization. In G. Underwood and R. Stevens (Eds.), *Aspects of consciousness* (Vol. 1: Psychological issues, pp. 67-89). London: Academic Press.

Reaux, R. A., Murphy, E. D., Stewart, L. J., Gresh, J. L., and Bruce, K. (1989). Building a modeling and simulation analysis tool to predict air traffic controller workload and performance. *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 52-56). Santa Monica, CA: Human Factors Society.

Redding, R. E. (1992). Analysis of operational errors and workload in air traffic control. *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 1321-1325). Santa Monica, CA: Human Factors Society.

Redding, R. E., Cannon, J. R., Lierman, B. C., Ryder, J. M., Seamster, T. L., and Purcell, J. A. (1990). *Cognitive task analysis of prioritization in air traffic control* (Report to the Federal Aviation Administration). McLean, VA: Human Technology, Inc.

Redding, R. E., Cannon, J. R., and Seamster, T. L. (1992). Expertise in air traffic control (ATC): What is it, and how can we train for it? *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 1326-1330). Santa Monica, CA: Human Factors Society.

Reid, G. B., and Colle, H. A. (1988). Critical SWAT values for predicting operator overload. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1414-1418). Santa Monica, CA: Human Factors Society.

Reid, G. B., and Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A scaling procedure for measuring mental workload. In P. A. Hancock and N. Meshkati (Eds.), *Human mental workload* (pp. 185-213). Amsterdam: North Holland.

Robertson, A., Grossberg, M., and Richards, J. (1979). *Validation of air traffic control workload models* (Report FAA-RD-79-83). Cambridge, MA: U. S. Department of Transportation, Transportation Systems Center.

Roscoe, A. H. (1987). In-flight assessment of workload using pilot ratings and heart rate. *Proceedings of the Advisory Group for Aerospace Research and Development* (AGARDograph No. 282). Neuilly sur Seine, France: AGARD.

Rosenberg, B. L. (1988). *Workload and performance measurement at the FAA Technical Center: Past, present, and future* (Unpublished working paper, WP-BR-88.005). Atlantic City Airport, NJ: FAA Technical Center.

Roske-Hofstrand, R. J. (1989). Video in applied cognitive research for human-centered design. *SIGCHI Bulletin, 21(2)*, 75-87.

Roske-Hofstrand, R. J., and Armstrong, H. B. (1989). If it weren't for the airplanes, this job would be easy: Air traffic controllers and aerodynamics. *34th Annual Air Traffic Control Association Conference Proceedings* (pp. 21-28). Arlington, VA: Air Traffic Control Association.

Roth, E. M., and Mumaw, R. J. (1994). Cognitive-engineering. In J. J. Marciniak (Ed.), *Encyclopedia of software engineering* (Vol. 1, pp. 110-123). New York: Wiley.

Roth, E. M., and Woods, D. D. (1990). Analyzing the cognitive demands of problem-solving environments: An approach to cognitive task analysis. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 1314-1317). Santa Monica, CA: Human Factors Society.

Runkel, P. J., and McGrath, J. E. (1972). *Research on human behavior: A systematic guide to method.* New York: Rinehart and Winston.

Schlager, M. S., Means, B., and Roth, C. (1990). Cognitive task analysis for the real(-time) world. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 1309-1313). Santa Monica, CA: Human Factors Society.

Selcon, S. J., Taylor, R. M., and Koritsas, E. (1991). Workload or situational awareness?: TLX vs. SART for aerospace systems design evaluation. *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 62-66). Santa Monica, CA: Human Factors Society.

Senders, J. W. (1970). The Estimation of operator workload in complex systems. In K. B. DeGreene (Ed.), *Systems psychology* (pp. 207-216). New York: McGraw-Hill.

Senders, J. W., and Moray, N. P. (1991). *Human error: Cause, prediction, and reduction.* Hillsdale, NJ: Erlbaum.

Shaffer, M. (1991). The use of video for analysis of air traffic controller workload. *Journal of Air Traffic Control, January - March*, 35-37, 40.

Sheridan, T. B. (1981). Understanding human error and aiding human diagnostic behavior in nuclear power plants. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp.19-36). New York: Plenum Press.

Smith, R. C., and Melton, C. E., Jr. (1974). Susceptibility to anxiety and shift difficulty as determinant of state anxiety in air traffic controllers. *Aerospace Medicine, June*, 599-601.

Soede, M., and Coeterier, J. F. (1971). Time analysis of the tasks of approach controllers in ATC. *Ergonomics, 14*, 591-601.

Sperandio, J. C. (1971). Variation of operator's strategies and regulating effects on workload. *Ergonomics, 14*, 571-577.

Sperandio, J. C. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics, 21*, 193-202.

Stager, P., Hameluck, D., and Jubis, R. (1989). Underlying factors in air traffic control incidents. *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 43-46). Santa Monica, CA: Human Factors Society.

Stein, E. S. (1985). *Air traffic controller workload: An examination of workload probe* (DOT-FAA-CT-TN 84/24). Atlantic City Airport, NJ: Department of Transportation/ Federal Aviation Administration Technical Center.

Stein, E. S. (1984). Controller workload: Past–present–future. *29th Annual Air Traffic Control Association Conference Proceedings* (pp. 290-293). Arlington, VA: Air Traffic Control Association.

Stein, E. S. (1988). *Air traffic controller scanning and eye movements - A literature review* (DOT/FAA/CT-TN 84/24). Atlantic City Airport, NJ: Department of Transportation/Federal Aviation Administration Technical Center.

Stern, J. A. (1987). Workload: A psychophysiological approach. In H. Fiedler (Chair), *Workload assessment techniques and tools* (Symposium proceedings). Dayton, OH: The Workload Sub-Technical Advisory Group of the Department of Defense Human Factors Engineering Technical Advisory Group.

Stern, J. A., and Skelly, J. J. (1984). The eye blink and work load considerations. *Proceedings of the Human Factors Society 28th Annual Meeting* (pp. 942-944). Santa Monica, CA: Human Factors Society.

Strasser, H. (1981). Physiological measures of mental load. In E. N. Corlett and J. Richardson (Eds.), *Stress, work design, and productivity* (pp. 77-89). New York: John Wiley.

Swain, A. D., and Guttmann, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications* (NUREG/CR-1278, SAND80-0200). Washington, DC: U. S. Nuclear Regulatory Commission.

Tattersall, A. J., Farmer, E. W., and Belyavin, A. J. (1991). Stress and workload management in air traffic control. In J. A. Wise, V. D. Hopkin, and M. L. Smith (Eds.), *Automation and system issues in air traffic control* (NATO Advanced Science Institute Series F: Computer and Systems Sciences, Vol. 73, pp. 255-266). New York: Springer-Verlag.

Technology Branch. (1989). *Controller visual scanning behavior - Manual data collection analysis report* (ACD-340-90-1). Atlantic City Airport, NJ: Department of Transportation/Federal Aviation Administration Technical Center, Advanced Concepts Division.

Thackray, R. I. (1980). *Boredom and monotony as a consequence of automation: A consideration of the evidence relating boredom and monotony to stress* (FAA-AM-1). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Thackray, R. I., Bailey, J. P., and Touchstone, R. M. (1979). The effect of increased monitoring load on vigilance performance using a simulated radar display. *Aviation, Space, and Environmental Medicine, 49*, 1215-1218.

Thackray, R. I., and Touchstone, R. M. (1979). *Effects of noise exposure on performance of a simulated radar task* (FAA-AM-79-24). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Thackray, R. I., and Touchstone, R. M. (1982). *Performance of air traffic control specialists (ATCSs) on a laboratory radar monitoring task: An exploratory study of complacency and a comparison of ATCS and non-ATCS performance* (FAA-AM-82-1). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute.

Thorell, L. G., and Smith, W. J. (1990). *Using computer color effectively*. Englewood Cliffs, NJ: Prentice-Hall.

Vidulich, M. A., Ward, G. F., Schueren, J. (1991). Using the Subjective Workload Dominance (SWORD) technique for projective workload assessment. *Human Factors, 33*, 677-692.

Vortac, O. U., Edwards, M., Jones, J., Manning, C., and Rotter, A. J. (1992). *Analysis of flight progress strip use by en route air traffic controllers* (Project report -- Observational Study 1.2, Version 1.1). Norman, OK: University of Oklahoma, Department of Psychology.

Weiner, E. L. (1989). *Human factors of advanced technology ('glass cockpit') transport aircraft* (NASA Technical Report 117528). Washington, DC: National Aeronautics and Space Administration.

Whitfield, D., and Stammers, R. B. (1978). The Air-traffic controller. In W. T. Singleton (Ed.), *The Analysis of practical skills* (Vol. 1, The Study of real skills, pp. 209-235). Lancaster, UK: MTP Press.

Wickens, C. D., and Kessel, C. (1981). Failure detection in dynamic systems. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp. 155-169). New York: Plenum Press.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

Wickens, C. D. (1984). *Engineering psychology and human performance.* Columbus, OH: Charles E. Merrill.

Wickens, C. D. (1989). Resource management and time sharing. In J. Elkind, S. Card, J. Hochberg, and B. Huey (Eds.), *Human performance models for computer aided engineering* (pp. 180-202). Washington, DC: National Academy Press.

Wickens, C. D. (1993). Timesharing, workload, and human error. In K. M. Cardosi and M. S. Huntley (Eds.), *Human factors for flight deck certification personnel* (DOT/FAA/RD-93/5). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Wickens, C. D., and Kessel, C. (1981). Failure detection in dynamic systems. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp. 155-168). New York: Plenum Press.

Wickens, C. D., and Yeh, Y. Y. (1988). Dissociation of performance and subjective measures of workload. *Human Factors, 30*, 111-120.

Wierwille, W. W., and Connor, S. A. (1983). Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. *Human Factors, 25*, 1-16.

Wierville, W. W., and Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors, 35*, 263-282.

Wierwille, W. W., Rahimi, M., and Casali, J. G. (1985). Evaluation of sixteen measures of mental workload using a simulated flight task emphasizing mediational activities. *Human Factors, 27*, 489-502.

Woods, D. D. (1985). Cognitive technologies: The design of joint human-machine cognitive systems. *The AI Magazine, Fall,* 86-92.

Woods, D. D. (1988). Coping with complexity: The psychology of human behaviour in complex systems. In L. P. Goodstein, H. B. Anderson, and S. E. Olsen (Eds.), *Tasks, errors and mental models* (pp. 128-148). New York: Taylor & Francis.

APPENDIX 8A. CHECKLIST ITEMS

1. Workload evaluations have considered both observable and perceived effects of task demand on the controller (8.1.1).

2. System demands do not overload or underload the controller for prolonged periods of time (8.1.2, 8.3.3)

3. Participants in workload assessments represent the range of experience, skills, and abilities that are present in the controller workforce (8.1.3).

4. Workload has been assessed with an appropriate battery of measures (8.1.3, 8.3.3).

5. Controllers will be able to make this design work without having to invent ways around design flaws (8.1.5).

6. This design fosters an active yet comfortably manageable role for the controller (8.2.1).

7. In comparison to the established baseline, controller workload stays about the same with this design (8.2.3).

8. The design requires little or no unaided recall of information from memory (8.2.4).

9. This design does not place heavy demands on short-term, working memory (8.2.4).

10. This design does not have greater memory demands than the previous system did (8.2.4).

11. This design does not increase the amount of data entry for controller tasks (8.2.5).

12. This design does not require the controller to recall infrequently used data-entry commands (8.2.5).

13. When tasks are performed together, workload remains manageable with this system (8.3.1).

14. This design does not contribute to increased information-processing workload (8.3.2).

15. If the design includes dynamic allocation of tasks, the plan for re-assigning tasks is under human control (8.3.2).

16. With this design, the controller is able to maintain the highest standards of safety, without having to exert extreme effort (8.3.4).

17. The design supports the controller in building and maintaining situational awareness (8.3.4).

18. The design supports the controller in making projections about the near-future traffic situation (8.3.4).

19. This design does not require timesharing of many moderately difficult tasks (8.3.4).

20. When timesharing is necessary, the tasks to be timeshared are spread across the controller's resources, that is, visual, auditory, and manual capacities, instead of loading up on just one or two capacities (8.3.4).

21. When timesharing demands are heavy, this design helps the controller remember to execute intended actions (8.3.4).

22. This design's strategy for task sequencing helps the controller remain alert (8.3.4).

23. This design provides variety in the kinds of tasks the controller needs to perform (8.3.4).

24. Procedural task sequences are interruptable at any point (8.3.4).

25. With this design, the controller does not experience abrupt changes in normal task loading (8.3.4).

26. With this design, timing of tasks can be flexible (8.3.4).

27. When certain tasks must be completed at specific times, their initiation is at the controller's discretion (8.3.4).

28. Use of this design over time will not have a negative effect on job satisfaction (8.3.5).

29. This design will have positive effects on the ways in which ATC team members interact and communicate with each other (8.3.6).

30. This design provides appropriate information to all members of ATC teams (8.3.6).

CHAPTER 9. WORKSTATION AND FACILITY DESIGN AND EVALUATION

Jeremy A. Guttman and Elizabeth D. Murphy

> "The workspaces of individual controllers are grouped into suites, according to the jobs and tasks. Suite design includes environment, software and hardware features. Each work position must contain all the facilities needed for the whole range of duties at that position, including information displays, data input devices and communications, and these must meet all the ergonomic requirements of reach and viewing distances and accessibility."
>
> ICAO, 1993

TABLE OF CONTENTS

CHAPTER 9. WORKSTATION AND FACILITY DESIGN AND EVALUATION

LIST OF FIGURES AND TABLES

CHAPTER 9. WORKSTATION AND FACILITY DESIGN AND EVALUATION

The first known application of psychological principles to an equipment design problem occurred during World War II when pilots of B-17s, B-25s, P47s, and other aircraft were retracting the landing gear (instead of the flaps) after landing (Roscoe, 1992). The psychologist who investigated this problem, Alphonse Chapanis, noted that the flap and gear controls were easily confused. The controls were, in most cases, identical toggle switches (or nearly identical levers) and were located side-by-side. After these controls were modified so that they could easily be differentiated by feel (the landing gear control had a small wheel on the end and the flap control had a wedge-shaped end), pilots stopped retracting the aircrafts' wheels after landing.

This is one example of "design-induced error" that needs to be avoided in the design of work environments. In the ATC environment, the work environment is the physical space that the controller works in and the elements of the workstation. The controller's "workstation" refers to console, work surfaces, related equipment (such as headset and microphone), and furniture.

9.1 INTRODUCTION

An ATC workstation includes all the items located within the controller's work space: the main display and control console, auxiliary displays, communications equipment, work surfaces, seating, and storage. If designed properly, ATC workstations and facility environments can promote the controller's safety, health, job performance, and job satisfaction. Proper design results from an understanding of the operational realities and the application of basic human factors principles.

This chapter examines the following key human factors issues in workstation and facility design:

- Derivation of design requirements for controller workstations.

- Design and layout of workstation components.
- Design of control room consoles, seating, and communications equipment.
- Environmental design in ATC facilities.

The discussion offers evaluation criteria that can be applied in many of these areas and refers the interested reader to the fairly vast literature on workstation and environmental design.

9.2 DESIGN REQUIREMENTS AND CONTROLLER PERFORMANCE

Workstation design is a key element of system design because controllers need workstations and facility designs that support their performance of functions and tasks in the new system. Any proposed arrangement of workstation components should be tested for usability, operational suitability, and workforce acceptance.

9.2.1 How are ATC operational requirements restated as workstation design requirements?

The process of translating operational requirements into workstation design requirements may be the most difficult task facing an engineering design team. Formal processes for providing this information are still being developed. Whether formally or informally stated, the following steps are usually followed to formulate workstation design requirements:

- Controllers provide input in terms of current or projected operational needs, scenarios and situations in which they will be using the new workstation. These are summarized as operational requirements.
- Controllers, along with system, design, and human factors engineers, describe the types of functions needed to fulfill the operational requirements. For example, if a conflict resolution advisory function is defined as an operational requirement, then the team would define how the resolution should be displayed to the controller.

- Technology is assessed to determine which functions can be implemented in the time frame proposed for the development effort. Functions that require advancements in technology are proposed as pre-planned improvements. At this point, a strawman design is usually proposed. A prototype can be produced for preliminary evaluation. This prototype could be a full scale mock-up of a preliminary workstation design.

- Task lists are generated based on the operational requirements. The functions proposed to meet these requirements and the proposed design or prototype of this design are also developed. These lists describe the number and type of tasks needed to perform the proposed functions. Task categories could include planning, data entry, information receipt, verbal coordination, and decision making. It is also useful to map controller tasks to the information that is displayed to the controller as these tasks are performed. This will help to determine whether sufficient information is displayed to the controller to perform the tasks.

- These tasks are matched against the operational scenarios in the sequence in which they would be performed. These task sequences can be analyzed in terms of the time needed to perform them compared to the time available. Estimates of controller workload can also be generated using the task information by examining the extent to which sensory channels such as vision or audition are being under or overused. If simulation facilities are available, actual task data can be gathered to provide workload and operational suitability estimates. (Workload evaluation is discussed in Chapter 8, Workload and Performance in the ATC Environment.)

- Based on analysis and simulation testing, final decisions are made concerning which functions can be allocated to the controller (e.g., use of input devices and communication devices) and which functions can be allocated to the machine (e.g., data reduction and display).

- If the preceding steps are followed, system workstation design requirements should naturally fall out of the function allocation process. For example, if a function is assigned to the controller, the workstation must provide a means for the controller to carry out that function.

9.2.2 How can the design of workstations and facilities affect the performance of ATC tasks?

The design of workstations can have a dramatic effect on the ease of use of system functions. Specifically, the design of a workstation can influence the speed and accuracy of the performance of ATC tasks. For that reason, proper workstation design is critical to the overall ATC system performance.

Software design has allowed many more functions to be placed on a workstation display than air traffic controllers are physically capable of operating with hard switches and knobs. Some guidelines for the design of workstations are as follows (DoD, 1989):

- **Avoid excessive layering of functions on keyboards and keypads.** Keys with three or four different functions could be confused by controllers, especially under stressful situations. This confusion will increase controller error rates for the functions associated with these keys.

- **Ensure that labels are understandable.** Use standard or familiar label abbreviations and place switch labels so that they are clearly visible. Unclear labeling can cause errors in control inputs as well as increased search time.

- **Ensure consistent backlighting of keyboards and workstation controls.** Controls that are illuminated at a lower level than keyboards or indicators will be hard to identify and use under dim conditions.

- **Place frequently used controls in locations which are easy to reach and see.** Poor placement of controls and control indications could cause delays when action is needed. Air traffic controllers must be involved in the prioritization of ATC control functions for purposes of workstation layout.

9.3 USER-CENTERED WORKSTATION DESIGN

The primary objective is to fit the workstation to the user, not the user to the workstation. This is a tall order, because of the wide range of individual differences, but it is achievable.

9.3.1 What are the human constraints on the design of ATC workstations?

Workplace and workstation design should be compatible with controller expectations and controller capabilities. Differences between people in physical dimensions and psychological attributes should be considered in the design of ATC workstations.

Dimensional Factors. The science of establishing proper sizes of equipment taking into account human dimensions is known as **anthropometry**. (The word comes from the Greek, meaning measure of man, or woman.) People vary with respect to anthropometric dimensions, such as sitting height, leg length, and reach. If these physical characteristics are ignored, the extremes of the controller population may not fit into the workstation or may not be able to reach workstation controls. Workstation dimensions should be compatible with the anthropometric characteristics of the controllers using them.

An evaluation can be accomplished by creating a full-scale mock-up of the workstation. Real people representing the extremes of the controller population can assess the workstation for clearance, vision, and reach. Mock-up evaluations can uncover problem areas that may not be evident from scaled drawings of the workstation. In addition, design solutions can be reached to take into account these differences. Examples of such solutions include adjustable workstation displays and seating. Dimensions that should be adjustable include keyboard height and angle, screen position and angle, and seat-pan/backrest angle.

Psychological Factors. Controller acceptance of a new or upgraded workstation design is critical. If a workstation is well organized, has convenient features, and is attractive, it is more likely to be accepted than if it does not have these features. Deficiencies in these workstation design factors could result in more serious problems, such as eye strain, fatigue, or even sickness. The controller's sensory, cognitive, and psychomotor attributes should be taken into account when making workstation design decisions.

9.3.2 What human factors issues need to be considered in determining the amount of space needed for an ATC workstation?

The amount of space needed in an ATC facility for workstations and groups of workstations is governed by several factors, including the type of airspace environment to be controlled (e.g., tower, terminal, en route) and the physical limitations of these facilities. The following are major human factors considerations for determining workstation spacing:

- Crew size, such as one-, two-, or three-person controller teams.

- Visual requirements of these persons or teams.

- Equipment accessibility for maintainer and support personnel.

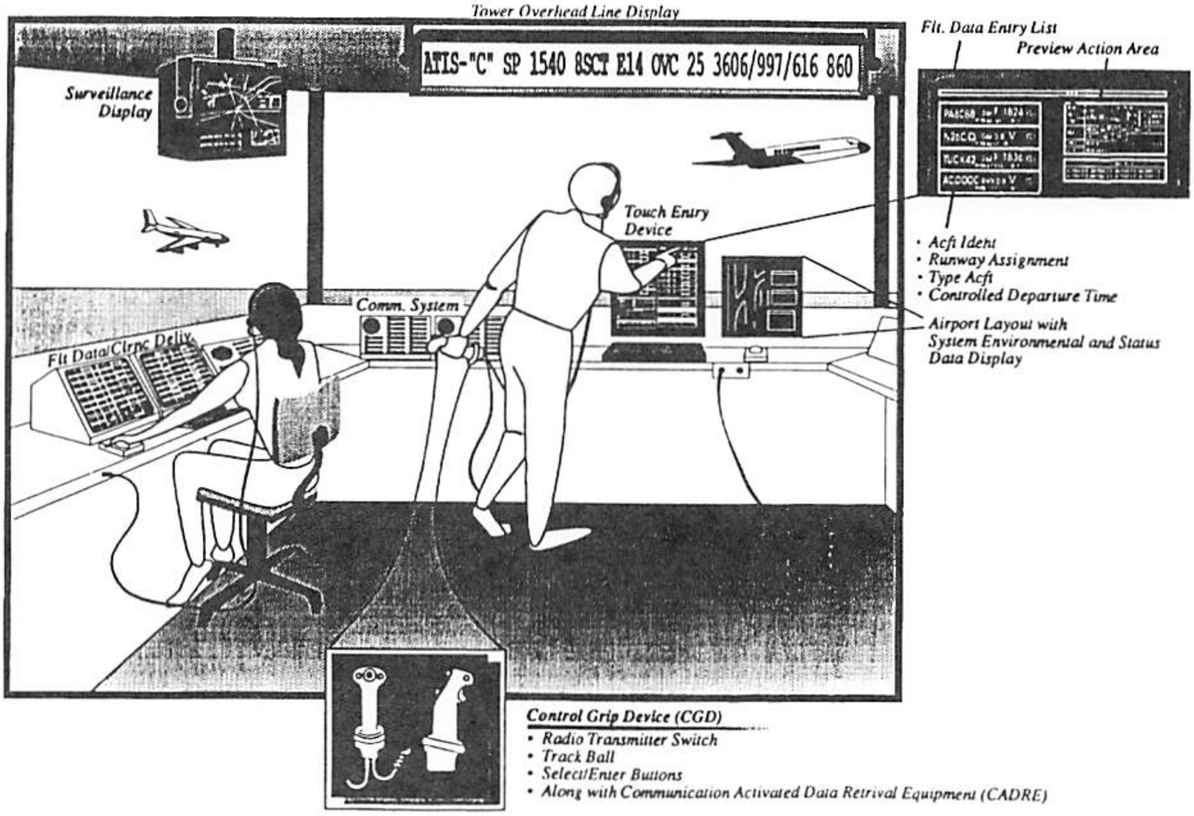
Crew Size. Space requirements will differ depending on whether the workstation will be used by only one individual or shared with others. As crew size increases, unaided, face-to-face communication becomes more difficult and members must sit or stand closer together for acceptable speech intelligibility. Space can be saved if two or more persons can share the same workstation controls and displays. If two or more persons frequently need to pass simultaneously between rows of workstations, then clearance between rows of workstations needs to be sufficient.

Visual Requirements. Spacing and clearance for workstation displays also depend on visual requirements for controllers and controller teams. Spacing and placement of workstations will differ depending on whether the controller is seated or standing or whether there will be a combination of seated and standing personnel. It is possible that a controller may have to stand behind another controller to see a display. This situation would require attention to human dimensions for workstation placement and spacing.

Figure 9-1 illustrates the human factors issues that have been considered in designing new workstations and equipment for ATC tower controllers. This design concept was developed to address controllers' objections to an earlier design concept, which had tower controllers sitting in "head-down" positions, when they often need to be standing and always need to be able to see out the window. The watchwords of this design are **visibility** and **flexibility** for the controller. This design concept illustrates the need for operational expertise on the design team or, at least, an understanding of operational realities. It is an interesting design problem that tower displays need to be devised for controllers who must constantly look out the window.

Equipment Access. Although operational considerations are primary, workstations should be arranged and spaced to allow accessibility to system components for easy removal and replacement by maintainers and support technicians. Brackets, support structures, or other obstacles should not interfere with the opening or removal of equipment covers or racks. In addition, workstations should be located so that other equipment or workstations do not interfere with accessibility to workstation system components.

Figure 9-1. Design Concept for Resolution of Human Factors Issues in Air Traffic Control Tower (ATCT) Workstation Design
Federal Aviation Administration (Air Traffic Requirements Branch, ATR-320). (June, 1991.) Tower Control Computer Complex (TCCC) Redesign. (Briefing to the Administrator by the Advanced Automation Program Office.)



9.3.3 What are the human factors issues involved in arranging workstation controls?

The speed and accuracy with which procedures can be accomplished depends on the logical arrangement of controls. This also includes controls that are software generated. Controls should be designed and arranged to support the controller's natural sequence of actions for accomplishing operational procedures. Arrangement decisions should be based on the task elements or procedures necessary to perform the required functions.

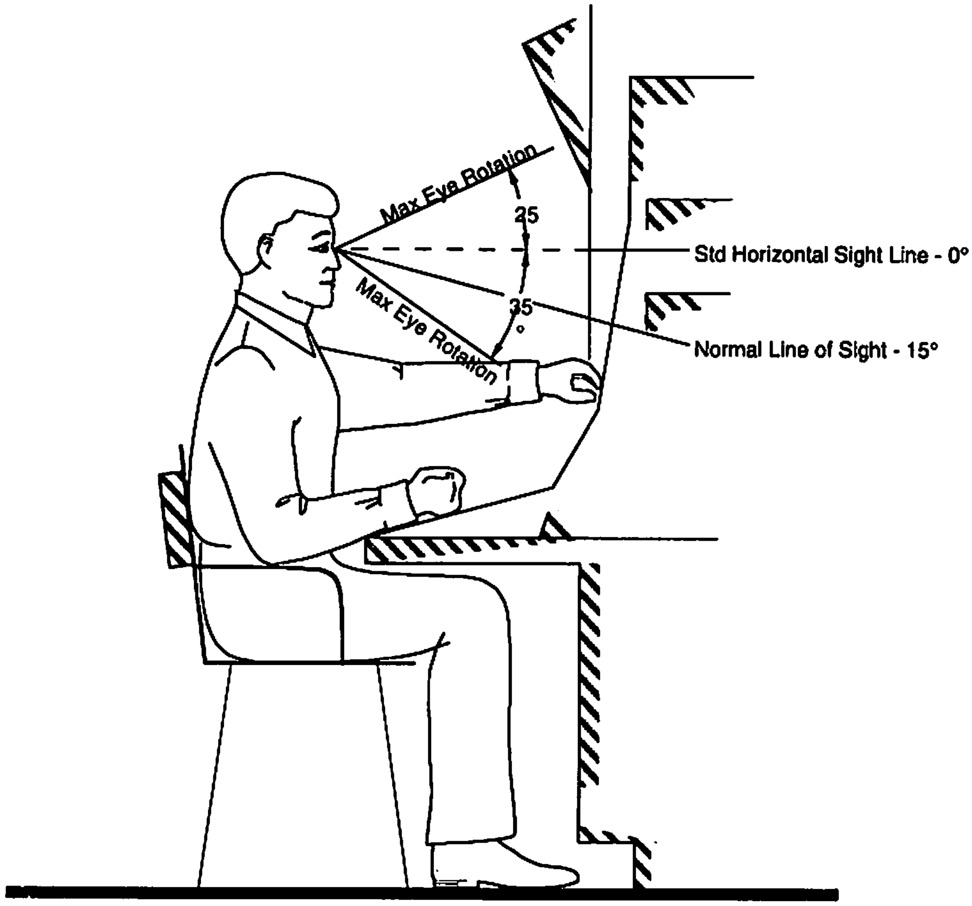
The following are major human factors concerns:

- Visibility of controls
- Clearance between controls
- Perceptual factors, including similarity and physical proximity of controls
- Expectations of the user population
- Health and safety concerns
- Standardization
- Flexibility

Visibility. Visibility of a control is usually necessary for efficient operation of an ATC workstation. Optimal visibility of controls is attained by arranging the controls with respect to the position of the controller's eyes. Factors such as whether the controller will be sitting or standing will influence the eye position. This geometry enables the designer to locate controls and displays for optimal visibility as well as to control glare and reflections. Figure 9-2 illustrates the visibility issues for a seated air traffic controller.

Clearance Between Controls. Controls should be arranged so that there is enough space between them to operate them easily. If controls such as toggle switches need to be grasped, then more clearance space is required. Controls that are spaced too closely together can increase the likelihood of accidentally activating a control or make it difficult to activate a desired control. Minimum separation distances depend on the type of control in question and the effects of other factors. Minimum distances between controls, as recommended by DoD (1989), range from 10 mm (0.5 in.) to 50 mm (2.0 in.) for bare-handed operators.Factors that influence the exact clearance dimension between controls include the size and breadth of the hand, which varies throughout the controller population. Controllers with larger hands will need more clearance to easily actuate hard controls.

Figure 9-2. Visibility of Controls and Displays for a Seated Air Traffic Controller
Adapted from *Human Engineering Guide to Equipment Design*, edited by H. Van Cott and R. Kincade, 1972, Figure 9-6, p. 393. Copyright 1972 by the American Institutes for Research. Also adapted from MIL-STD-1472D, DoD, 1989.



Factors which affect clearance between software-generated controls include the sensitivity of the input device, such as a trackball or touchscreen, and the resolution of the legends or icons used for control labels or symbols. Safety factors may also require additional clearance between a critical control and less important ones.

Perceptual Factors. Perceptual factors also play a role in the arrangement of controls. Controls that are logically linked, tend to be used in sequence, or are similar in some other fashion should be arranged in close proximity. In addition, controls that control various displays should be placed close to the display, if feasible. Providing physical boundaries, such as etched lines, can improve the identification and use of controls.

Expectations of the User Population. Controls should operate in ways that are consistent with the past experience of the user population. In the United States, we expect to turn a switch on by flicking it up and to turn it off by flicking it down. (In the United Kingdom, expectations are reversed: down means on, and up means off.) Likewise, we in the U.S. expect that turning a knob clockwise will cause an increase in the flow of energy (e.g., power, brightness, pressure). Design that violates such expectations runs the risk of contributing to incidents and accidents.

Health and Safety. The controller's health and safety should take a high priority in workstation design. A hazard analysis should be completed when developing a new or upgraded workstation or workplace design. Safety hazards can range from a sharp corner or projection on a work surface, to a control and display layout that could increase the likelihood of a controller error. Deficiencies in these workstation design factors could result in problems such as eye strain, fatigue, or even sickness. An example of an illness induced by poor workstation design is carpal tunnel syndrome. This can be induced by a workstation that promotes repetitive bent-wrist motions. This problem is caused by damage to the median nerve that passes through the wrist. Symptoms include numbness, loss of grip, and even loss of hand function. Design can reduce the likelihood of a controller's developing carpel tunnel syndrome. Additionally, electromagnetic radiation from video display hardware should be minimized to reduce potential health risks (Black, 1992; DoD, 1989).

Standardization. Standardization among systems provides several important benefits, including reduction in training time for the new system and a lower likelihood of controller error allowing transfer from one ATC facility to another. Standardization also allows cost savings from the development of common hardware as well as reduced logistic support. The designers of a new or upgraded system should be careful, however, to recognize the danger of carrying over a poor design concept for the sake of standardization. For example, mirror-image control layouts may be ideal from a logistics and standardization perspective, but may lead to confusion and increased errors from an operational perspective. Standardization of space requirements across facilities may be in conflict with the needs of individual facilities.

Flexibility. Flexibility can be incorporated into the design by providing monitor adjustments. High-mounted display monitors, in particular, should be tiltable. Control and display adjustability supports controller performance and comfort. Given a suitably functional arrangement, flexibility is another key design objective for ATC workstation design. For example, control arrangements should be usable by both left- and right-handed controllers.

9.3.4 What are the physical criteria for evaluating the suitability of workstation controls and displays?

The design and layout of workstation controls and displays usually requires some trade-offs since not all controls, displays, and indicators can be optimally located. Therefore, some evaluation and prioritization is necessary when deciding on control and display locations. Normally, high-priority controls should be placed as close as possible to the air traffic controller, and high-priority displays and indications should be centrally located (DoD, 1989). For evaluation of the physical (as opposed to the cognitive) aspects of controls and displays, there are three main criteria:

- *Visibility* of the controls, displays, or indicators and their associated labeling.

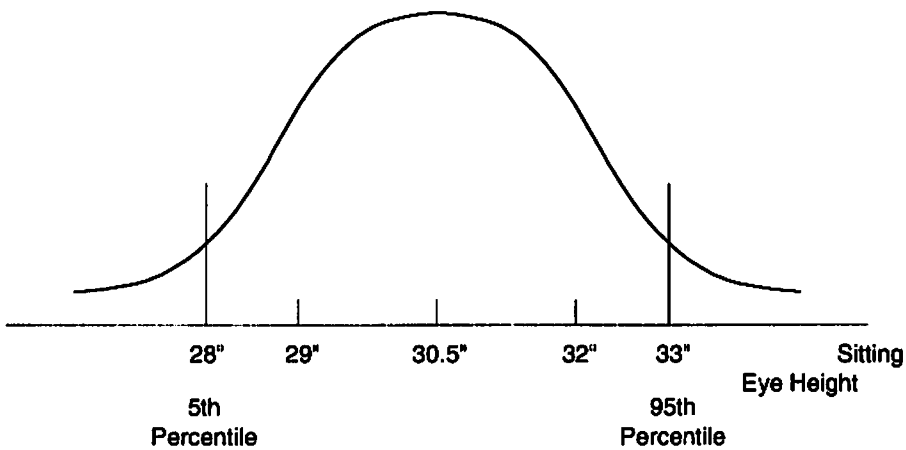
- *Clearance* between controls so that there is enough space between them to permit adequate grasping and manipulation and to prevent accidental activation.

- *Reach* to controls without excessive shoulder movement and or bending of the back.

To assess these factors, human measurements must be used to obtain a proper physical fit between the workstation dimensions and the different sizes of controllers who will be using the workstation. For example, eye position for a seated controller will vary depending on the length of the controller's torso and neck. Fortunately, surveys have been made on hundreds and, in some cases, thousands of people on the dimension of sitting eye height as well as other physical dimensions (DoD, 1989). The variation between controllers on this dimension can be assessed from forming a distribution of all these individual measurements. The average sitting eye height as well as the extremes can be determined from this distribution. Figure 9-3 illustrates the distribution of the sitting eye height for the population of male adults. The designer can use these dimensions for control and display layouts. People with these dimensions can be identified to evaluate the workstation. A typical design goal is to accommodate the 90 percent of the user population from the 5th percentile female through the 95th percentile male (DoD, 1989).

For an evaluation of workstation design, persons representing the extremes of the air traffic control population should be selected. If a full-scale mock-up is available, these people can systematically assess the visibility of controls and displays, and the labeling associated with them. In addition, reach and clearance issues can also be addressed by these persons. A full-scale mockup can also be evaluated by air traffic controllers for operational suitability of controls and displays.

Figure 9-3. Distribution of Sitting Eye Height for Male Adults
Adapted from MIL-STD-1472D, 1989, DoD, Table XIV.



9.3.5 What are the key human factors issues associated with shared visual displays that are located at some distance from any individual workstation?

Large screen displays are useful when a group of controllers must frequently refer to the same information. In addition, space constraints may preclude the use of individual displays for the display of commonly used information. Several key human factors issues are associated with the design of shared visual displays:

- Viewing angle
- Viewing distance
- Image luminance
- Control of displayed information.

Viewing Angle: Large screen displays should only be used when the geometry of the display allows all intended controllers to see the information displayed without visual obstructions. This constraint includes people walking in front of the display and obscuring information due to placement of the display in front of a busy walk way.

Viewing Distance: The display should not be placed further from an observer than the resolution of the display will permit. For large screen displays, character and symbol sizes based on recommended visual angle will have to be increased from standard display sizes due to longer viewing distances. In addition, the display should not be placed *closer* to any observer than one-half the length of the longest dimension of the display (DoD, 1989).

Image Luminance: The luminance of the image should be uniform across the display. However, some non-uniformity can be tolerated by the human eye. If the brightest part of the display is more than three times brighter than the dimmest part of the display, then the image will be unacceptable to shared observers. If a projection display is being used, the maximum luminance of the screen as a function of viewing angle should not vary more than four times that of the minimum luminance (DoD, 1989).

Control of Displayed Information: The control of displayed information should be designed so that critical information cannot be inadvertently modified or deleted. Control of changes to a shared display should be under the control of designated personnel who operate according to pre-established procedures. The content of displayed information should be evident to a trained observer without reference to the display control settings. For example, on a large screen display that contains information pertaining to the arrival and departure times of several aircraft, the legend that indicates that a flight will be arriving later than normal should be evident from the screen. The controller viewing the information should not have to look at the control panel settings to figure out that a flight will be arriving later than expected.

9.4 DESIGN OF CONTROL ROOM CONSOLES AND SEATING

Design of both consoles and seating should be based on controllers' needs for workspace, access to equipment, and comfort. These needs are reflected in various kinds of data that are available to designers. Control room consoles and seating should be thoroughly evaluated from the operational perspective.

9.4.1 What kinds of data can be useful in designing control room displays?

The data required for design of consoles fall into two major categories:

- *Anthropometric data*, which describe the differences in size and length of the operator's hands, limbs, torso, and other physical characteristics for the user population.

- *Task data*, which describe the types, numbers of tasks, and logical task sequences needed to perform ATC operations. Once the tasks are defined, ATC consoles can be designed and arranged to provide the most logical sequence of tasks for ATC operations.

Anthropometric Data: Designing for the Extremes. The minimum goal of any anthropometric design process in ATC environments is to ensure the air traffic controller is given adequate room, comfort, and access to workstation controls and displays.

The first step involves defining operational requirements for the ATC system as a whole and the workstation as a subset of that system. These operational requirements will dictate the functions needed on the workstation, which will in turn present design options for the types and number of knobs, switches, buttons, and equipment to be located at the workstation.

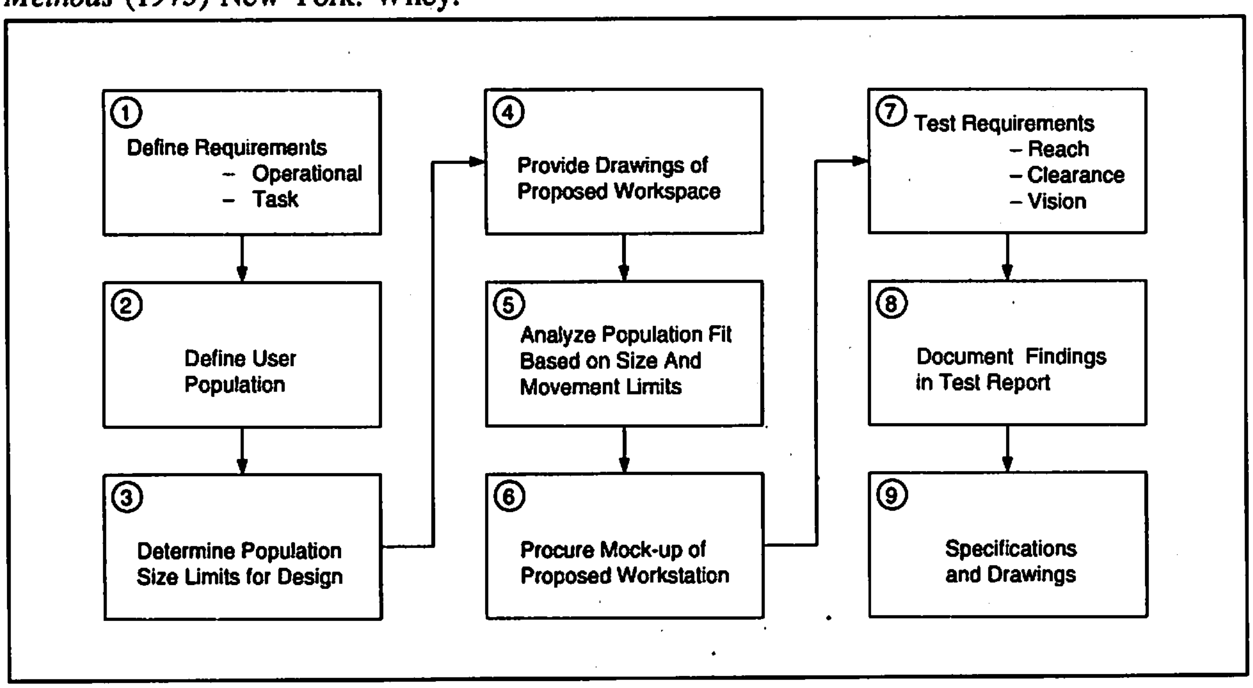
Once the requirements have been determined, the user population is defined. This definition is normally accomplished by surveys of the workforce on critical physical dimensions. Fortunately, surveys have been done on a number of user populations, so it is possible to extrapolate based on existing surveys (e.g., DoD, 1989).

Once the user population has been defined, design limits are imposed to define the percent of the workforce that can be expected to be comfortable at the workstation. As previously noted, typical design limits are the 5th percentile female and the 95th percentile male (DoD, 1989).

After design limits have been agreed upon, preliminary drawings of the proposed workstation and workspace are prepared. Designers and human factors engineers can then analyze workforce fit with the workstation, often using computer programs specifically designed to test reach and clearance for newly designed workspaces.

For major upgrades, full-scale mock-ups are often produced. These mock-ups are physical representations of the workstations, often incorporating the proposed lighting system, but they can be as simple as foam-core mock-ups. Once mock-ups are constructed, human factors testing for reach clearance and vision can be accomplished. Figure 9-4 illustrates these steps for the workstation design process.

Figure 9-4. Flow Diagram For ATC Workstation Design Procedure
Roebuck, J. A., Droemer, K. A. E., and Thomson, W. G. *Engineering Anthropometry Methods* (1975) New York: Wiley.



Task Data: Designing for Operational Suitability. Task data gathered from currently implemented ATC consoles can help to build operational suitability into new systems. Task types, times, and sequencing provide a basis for determining controller workload by allowing the designer to forecast where ATC tasks will "pile up." (Methods of evaluating workload are discussed in Chapter 8, Workload and Performance Measurement in the ATC Environment.)

Task data can be gathered through various means, including video tapes of ATC simulation exercises and interviews with air traffic controllers. The data collection method needs to be chosen carefully in order to provide the appropriate data. If the console is a completely new design, initial task data may have to be inferred in part from drawings, plans, or procedures. Controllers can provide estimates based on their operational experience and familiarity with a new design. When prototypes are available, more precise measures can be taken.

The following kinds of task data can provide useful guidance to designers:

- The frequency of a task associated with a particular control, display, or indicator.
- The criticality of a task associated with a particular control, display, or indicator.
- Estimated times for performing control or display operations with a particular control, display, or indicator.
- Relationships or links between tasks.

The criticality and frequency of an ATC task may not be perfectly correlated. For example, an alert indicating that an aircraft is being hijacked is expected to be a rare occurrence. However, this is critical information that must be distinctive and easily recognized. It is useful, therefore, to collect these task measures independently and, when appropriate, to combine frequency and criticality ratings to determine the most critical ATC tasks. Frequency and criticality ratings can be additively combined and weighing factors can be used if the design team feels that either frequency or criticality of the task is more important. Different methods can result in different rankings of tasks (Sanders and McCormick, 1987). In any case, the more design is guided systematically by realistic task data, the more likely it is to produce an operationally suitable system.

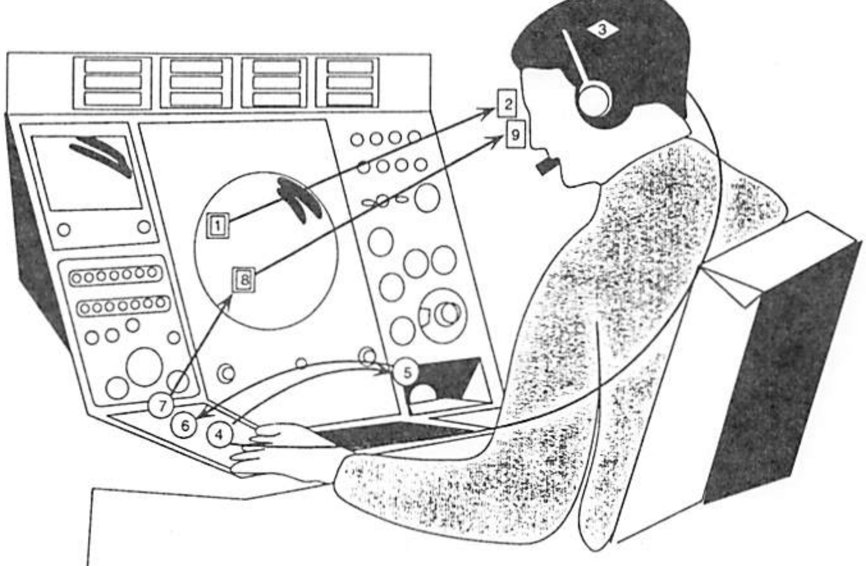
Once an inventory of tasks is established, an analysis of how the tasks are interrelated can be performed. Relationships between people and console components can be defined as links. The following are typical link types:

- Communication Links
  - Visual
  - Auditory/Verbal

- Movement Links
  - Eye movements
  - Hand movements
  - Foot movements

Once links are established, they can be graphically depicted using spatial operational sequence diagrams (Geer, 1981; Sanders and McCormick, 1987, 1993). Figure 9-5 shows an example for an air traffic controller tracking an aircraft. These diagrams allow the designer to arrange controls and displays to produce the most efficient set of task relationships.

Figure 9-5. Operational Sequence Diagram for En Route ATC
From *Human Factors in Engineering and Design* (6th Edition) by M.S. Sanders and E. J. McCormick, 1987, Figure 13-4, page 371. Copyright 1987 by McGraw-Hill. And from *Human Engineering Procedure Guide* (AMRL-TR-81-35) by C. W. Geer. Copyright 1981 by Armstrong Aeromedical Research Laboratory, Wright-Patterson Air Force Base, OH. Adapted by permission.



9.4.2 What features should control room seating have in order to meet guidelines for seat design?

Many air traffic controllers spend a good deal of the time sitting. The seating used by air traffic controllers in the United States was evaluated some time ago by a survey of controller likes and dislikes. Some highlights of the results were as follows (Kleeman and Prunier, 1979):

- The seats were heavily used. Nearly 70 percent of the controllers surveyed spent 5 to 8 hours sitting per shift.

- Sixty two percent of the controllers surveyed felt comfortable sitting in the then-current chair from 1 to 5 hours.

- Seats used at the time of the survey were rated as favorable by 68 percent of the controllers surveyed.

- Seventy four percent wanted the seat back upholstered with woven fabric.

- Fit and comfort were ranked highest of the most wanted features in a seat.

- Seat height adjustment was ranked as the most important adjustment in a seat.Seating research has produced the following general guidelines for seat design (ANSI/HFS, 1988; Grandjean, 1988):

- The buttocks should support most of the body weight.

- The thighs should exert as little pressure as possible on the seat.

- The lower back **must** be supported.

- It should be possible to place the feet on the floor.

- The seated person should be able to change posture.

The following design features meet guidelines for seating (DoD, 1989):

- *Compatibility*. The seat provides an adequate supporting framework for the body.

- *Vertical Adjustment*. The design provides an adjustable range of 15 to 21 inches in seat height with adjustable increments of no greater than 1 inch.

- *Backrest.* The seat has a supporting backrest that reclines between 100 and 115 degrees. It adequately supports the lower and middle back. The backrest brings the torso into a position that places the controller's eyes as recommended with respect to visual displays, with no more than 3 inches of body movement.

- *Cushioning.* Both the backrest and the seat are cushioned with at least 1 inch of compressed material.

- *Armrests.* Arm rests are provided. The arm rests are 2 inches wide and 8 inches long.

Extensive research has been conducted on seating requirements for work environments. The interested reader is referred to Andersson, 1987; Bendix, 1986; and Grieco, 1986. Because there are conflicting recommendations in the existing standards on seating, it would be wise to test any proposed seating on a representative sample of the controller population. Providing two or three good seating design from which controllers can choose is recommended because "there is no one best chair for everyone" (Sanders and McCormick, 1993, p. 438).

9.5 DESIGN OF COMMUNICATION EQUIPMENT

Anyone who doubts that the success of the ATC system is tightly linked to the quality of its communications support need only review the evidence from the communications switching problem that occurred in New York City in 1991. Air traffic control was disrupted up and down the East Coast, and flights were delayed for hours (Roush, 1993). At the controller's end, in the ATC facility, the resolution of ergonomic issues can support or impair communication with pilots and other facilities.

9.5.1 What are the key human factors recommendations on the design of equipment for speech transmission?

Key issues include intelligibility and the effects of environmental noise. Because intelligibility depends on the portion of the speech spectrum between 200 to 6,100 Hz, DoD and NASA guidelines recommend that microphones and associated system-input devices be designed for optimal response to that range of frequencies (DoD, 1989; NASA, 1989). If system engineering constraints require a narrower range, 250 to 4,000 Hz is the minimally acceptable frequency range. Additionally, across the frequency response bandwidth, amplitude variation should not exceed plus or minus 3 dB (NASA, 1989). Microphones used with amplifiers should have a dynamic range wide enough to pick up variations of at least 50 dB in signal input (DoD, 1989; NASA, 1989).

Techniques used to minimize the effects of noise during speech transmission include increasing the level of the speech signal in relation to the noise signal, and peak clipping of the speech signal, i. e., clipping off the extremes in amplitude. In cases where increasing the level of the speech signal may not be feasible, peak clipping is one alternative.

The purpose of peak clipping is to increase the intensity levels of consonants in relation to the intensity levels of vowels. This is desirable because the lower-amplitude waveforms associated with consonants are more susceptible to noise interference or masking as compared to the higher-amplitude waveforms associated with vowels. If the speech sound is passed through a peak-clipper and the resulting signal is reamplified to the available peak power level, the less noise-resistant components of speech will be protected from the effects of noise.

Other techniques can improve the intelligibility of transmitted speech. These include exclusion of noise at the microphone by using a noise shield, a throat microphone, or a pressure-gradient microphone. Each of these technologies has advantages and disadvantages (See Boff and Lincoln, 1988; Hawkins and Presson, 1986).

Calibration of microphones and headphones requires the use of instruments such as voltmeters, sound level meters, acoustic couplers, and spectrum analyzers. The design should preclude squeal problems and echo effects by isolating feedback between headphones and microphones.

9.5.2 What are the key human factors recommendations on the design of speech-reception equipment?

The primary issue in the design of headphones and loudspeakers is the assurance of intelligibility (i.e., that listeners will understand the transmitted message). The same frequency ranges apply to headphones and loudspeakers as those specified in the previous discussion of microphones and transmission equipment (i.e., 200 to 6,100 Hz or a minimal range of 250 to 4,000 Hz).

A related issue is the response capability of multiple channels fed into headphones, for example, several speech channels that the controller needs to monitor simultaneously. Such channels should "respond uniformly (plus or minus 5 dB) over the range from 100 to 4,800 Hz" (DoD, 1989, p. 63). When loudspeakers are used for multi-channel monitoring, a filtering scheme may be necessary to aid listeners in differentiating among channels (DoD, 1989).

In cases where peak clipping of the speech signal is not used but the transmission equipment does use pre-emphasis, the recommendation is for reception equipment to use frequency de-emphasis of signal characteristics that complement the characteristics of pre-emphasis only if such de-emphasis will improve intelligibility (DoD, 1989).

Positive feedback noise should be controlled so that it does not adversely affect normal voice communication (NASA, 1989). Likewise, delays due to satellite transmission need to be minimized so that the number of transmissions does not increase substantially and impair communication between pilots and controllers that are blocked (Nadler, DiSario, Mengert and Sussman, 1990).

Reproduction of the speaker's verbal input, as heard on the headset, should be in phase with the actual input. If this is not the case, the speaker's input will be disrupted by the out-of-phase side tone. The speaker's side tone should be received in the headset without any prior filtering or modification (DoD, 1989).

Binaural rather than monaural headsets are recommended if listeners will be working in high ambient noise (85 dB(A) or above) (which is typically not the case in ATC facilities but is in many aircraft cockpits). If binaural headsets are used, they should be wired so that sounds reach the two ears in opposing phases, unless operational requirements mandate otherwise (DoD, 1989; NASA, 1989). Research indicates that intelligibility is greatest for binaural speech and noise presentation "when the phase relations of speech and noise at the two ears are opposite (i. e., when speech is in phase and noise 180° out of phase, or the opposite)" (Boff and Lincoln, 1988, p. 1819).

Intelligibility is aided by any procedure that allows listeners to separate and lateralize speech and noise signals (Boff and Lincoln, 1988). Speech intelligibility can be measured in several ways. The following are commonly used methods (DoD, 1989; McCormick and Sanders, 1982). In each method, the following techniques are applied.

- The ANSI standard method of measuring phonetically balanced (PB) monosyllabic word intelligibility (ANSI, 1960). Listeners are scored for their accuracy in repeating words that proportionately sample everyday speech sounds.

- The modified rhyme test (MRT). Scoring is based on the number of correct responses to rhyming word pairs (e.g., coat-goat); after the word pairs are presented, participants indicate which word they heard.

- The articulation index (AI). This approach uses procedures for deriving indirect indicators of speech intelligibility (ANSI, 1969). The computation of AI is too complex to be discussed here. A sample procedure is detailed in a widely used human factors text (McCormick and Sanders, 1982).

The ANSI standard method should be used when test sensitivity and accuracy must be high; the MRT can be used if test requirements are less stringent or if the time and training requirements for the ANSI method cannot be met; AI calculations can be used to derive estimates, comparisons, and predictions of system intelligibility (DoD, 1989; McCormick and Sanders, 1982). The higher the articulation index of a system, the higher the percentage of the system's vocabulary would be correctly understood by the listener. For example, with an AI of .47, 75 percent of a 1,000 word vocabulary would be correctly understood; 90 percent of a 256 word vocabulary would be understood (Sanders and McCormick, 1993).

For an ATC environment, where exceptionally high intelligibility is necessary, the minimally-acceptable intelligibility scores should be as follows for the different methods (DoD, 1989, p. 65):

- Phonetically Balanced word intelligibility (PB) – 90%
- Modified Rhyme Test (MRT) – 97%
- Articulation Index (AI) – 0.7

Scores given in percentages represent the percentage of speech content that is correctly understood by the listener (McCormick and Sanders, 1982). The AI score can be thought of as "a summated index of the differences... between a typical speech spectrum and the spectrum of the background noise" (McCormick and Sanders, 1982, p. 160). A speech communication system with an AI under 0.3 is likely to produce unintelligible, garbled speech signals (Beranek, 1947; McCormick and Sanders, 1982).

9.5.3 What are some key recommendations on the design of headphones and audio control options?

An important piece of ATC equipment for presenting audio material to the controller is the currently available headset. Earphones for input to the controller and a microphone for output comprise the headset. The microphone portion of any headset must be designed to function optimally over the spectrum for intelligible human speech, ideally 200 to 6000

Hz, with a minimally acceptable range of 250 to 4000 Hz (DoD, 1989). Incoming audio signals should always be provided both through the earphone and outside the earphones. Dual channel earphones provide two advantages in ATC. First, they can be used for alternating a signal from ear to ear which is a more effective means of alerting the controller than presenting the signal simultaneously to both ears. Second, dual channels can be used to prevent masking when two audio signals/messages arrive simultaneously. For example, voice communication from pilots can be presented to the controller through one ear, and an emergency alarm can be presented to the other ear.

To minimize the occurrence of simultaneous audio presentation there should be an alarm/message priority system and, where feasible, signal and speech messages should be presented one at a time. Also, related or redundant alarms/messages should be integrated, for example, to indicate a complete system failure.

Certain aspects of audio display should be placed under human control rather than or in addition to computer control. The controller should be able to switch off audio signals that are designed to continue as long as the related problem persists. A capability to easily disable critical alarms, however, could promote controller error. If a problem is not fixed after a reasonable amount of time, the computer should again remind controller of a serious condition. Once a signal terminates for whatever reason, the computer should automatically reset it so that the alarm will repeat should the problem reoccur. Turning off an auditory signal should not erase any associated visual information (Smith and Mosier, 1986).

Signals that repeat or are of unlimited duration should be used sparingly, for rare emergency conditions. Otherwise, they can annoy the controller and create a habit of immediately and casually turning off alarms.

The volume of incoming speech messages is another audio display element that the controller should continue to be able to vary since aging, exposure, and other factors may degrade an individual's auditory perception abilities. The amount of loudness control depends largely on the design. If intensity level holds with it an associated meaning, signal intensity may have to remain fixed to retain that meaning. The volume control mechanism should prevent the controller from reducing any audio display to an inaudible level. If signals are used to indicate a controller's computer input error, the controller should have control over disabling the "beeps" or switching to a visual indicator (Smith and Mosier, 1986).

9.5.4 What provisions should designers of communications equipment make for controller comfort and convenience?

Headphones or any communication equipment to be worn by the controller should be designed for comfort. The headset should be designed so that no bare metal parts come into contact with the controller's skin (DoD, 1989; NASA, 1989). Controllers who wear glasses should not experience discomfort because of the headphone. Design of microphones, headphones, and telephone headsets should allow hands-free operation under normal working conditions (DoD, 1989; NASA, 1989).

Telephone handsets should be readily accessible. If multiple handsets are needed, the most frequently used or most urgently needed handset should be the most readily accessible (DoD, 1989).

9.5.5 What kinds of operational controls should be provided with voice communication equipment?

Three kinds of controls are recommended (DoD, 1989):

- *Volume/gain controls.* Preferably, volume/gain controls should be separate from power (on-off) controls. These controls should be limited to an audible level so that the controller cannot inadvertently disable the system with the volume control, and they should be powerful enough to drive sound pressure level to at least 100 dB overall when two earphones are in use.

- *Squelch control.* Continuously monitored communication channels should include a signal-activated switching device (squelch control) that suppresses channel noise during inactive periods; but the controller should be able to manually deactivate the squelch using an on-off switch when weak signals are detected.

- *Foot/hand-operated controls.* Foot pedals should be provided for seated controllers to activate "talk-listen" or "send-receive" control switches; hand-operated controls should be provided for standing controllers and for use in emergencies.

Evaluation of alternative designs for foot controls can include collection of data on participants' reaction time, speed of operation, force required, and personal preferences. Use of foot controls in ATC may be somewhat controversial because they will tend to restrict the controller's posture, possibly contributing to fatigue. Human factors issues should be resolved before requiring controllers to use foot pedals. The chief of these issues may be workforce acceptance. Design issues include the location of the foot pedal's fulcrum (which directly affects activation time) and the placement of foot pedals in relation to other controls and in relation to the seated controller.

9.6 ENVIRONMENTAL DESIGN

The facility environment can have positive or negative effects on controllers' performance and job satisfaction. Although we know that a crucial aspect of the environment is the organizational climate, we limit this discussion to physical characteristics of the environment, such as lighting, temperature, and noise. Unacceptable levels of these and other environmental factors can have measurable effects on productivity.

9.6.1 What are some considerations in specifying lighting requirements for ATC workstations and facilities?

As noted in Chapter 3, Visual Perception, the perception of light is an interaction between the physical properties of light and the make-up of the human visual system. **Illuminance** is defined as the amount of light that strikes a spherical surface at a given distance from the light source. Consider a candle sitting in a candle holder. Light is emitted by candle is all directions. The farther one is from the candle, the fainter the illumination. The amount of illumination striking a surface from a point source of light follows an inverse square law where:

$$\textbf{Illumination} = \textbf{candlepower}/\mathbf{D}^2$$

where candlepower is the output level of the light source and D is the distance of the light measurement device or observer from the source. **Luminance** is defined as the amount of light reflected once it strikes a surface. Light that is emitted from a surface, as from a fluorescent bulb or a workstation display, can also be measured in terms of illuminance and luminance (Sanders and McCormick, 1987).

The condition of the visual system also influences how light is perceived. Factors such as age and other individual differences can influence what is considered the optimal illumination level for an individual controller.

As mentioned in Chapter 3, Visual Perception, the human visual system changes as one ages. The major change that occurs with age is a thickening of the lens of the eye, which in turn deteriorates one's near visual acuity and the amount of light that reaches the retina. There is a 50 percent reduction in the amount of light that reaches the retina at age 50, compared to age 20 (Weale, 1961). The light that does reach the retina is also more scattered, reducing the contrast of the image. All these changes in the visual system mean that the older individual must acquire more light to achieve the same level of visual acuity. Older controllers, therefore, require more direct illumination (through ambient lighting) than do younger controllers.

Age is only one of several factors that can influence the perception of light and what is considered optimal by a controller. Other factors include the following:

- Tasks the controller is required to perform.
- Lamps and the type of ambient lighting (e.g., direct sunlight, semi-direct or diffuse lighting).
- Level of illuminance.
- Distribution of light.
- Range of luminance from the workstation display and indicators.

Lighting and Controller Task Requirements. Lighting requirements typically have been documented in the form of standards and guidelines that often have no specific reference to the task a controller has to perform. Recently, the controller user community has been represented by the Sector Suite Requirements Validation Team (SSRVT) in the development of lighting requirements for the AAS. The SSRVT has defined the following facility lighting criteria related to the tasks the AAS controller will have to perform (Bashinski, Krois, Snyder, and Tobey, 1990):

- Usability of data entry device. All key labels on the keyboard, electronic display keypads, and trackball should be easy to recognize with an acceptable amount of effort.
- Usability of text, handwritten, and color-graphic material. Lighting must adequately support the writing of notes on the console shelf and the reading of materials placed on the shelf.
- Usability of controls. Lighting must be sufficient for the controller to readily locate switches, controls, headset jacks, connectors, handles, and display recess mechanisms. Control labels must be readable.

- Minimal amount of visual distraction. Lighting must minimize distracting shadows, glare, and reflections.

- Readability of color display and adequate display brightness. Color display washout must be minimized, and adequate contrast must be provided for textual and graphic information. The controller must be able to identify all color codes under all anticipated lighting conditions. For detailed information on the effect of sunlight on color displays, see section 4.1.2 of Appendix A, *Human Factors Considerations for Color Displays for ATC.*

- Adequate emergency and maintenance lighting. Illumination must be sufficient to perform ATC tasks under emergency conditions and to allow safe ingress/egress from operating positions. Lighting design must buffer maintenance lighting so that it does not interfere with ATC tasks at the workstation console.

These task requirements may require mock-up evaluation by air traffic controllers to determine if the workstation or facility lighting is adequate.

Lamps and Type of Lighting. Two main types of lamps are used in facilities: **incandescent** filament lamps, which produce light by electrical heating of a filament, and **gas-discharge** lamps, in which light is produced by the passage of an electric current through a gas. The most common kind of gas discharge lamp is the fluorescent lamp.

Either incandescent or gas-discharge lamps can produce light directly or in a diffuse manner. Direct lighting is useful for illuminating small areas, such as task lighting for reading or writing notes. Diffuse lighting is often used to illuminate large areas because it minimizes glare and reflections. One or more devices can be incorporated into a lamp assembly to control the distribution of light. These include lenses, diffusers, shielding, and reflectors (Sanders and McCormick, 1987).

Illuminance Level. The task of determining an operational illumination level is a major concern for both the users and designers of ATC systems and facilities. The major factor in determining illuminance level is the type of activity to be performed and the extent to which detailed visual acuity will be required. Generally, the more detailed vision that is required, the higher the ambient illumination needs to be. Air traffic control towers present special lighting problems. During the day, direct sunlight can "washout" a visual display. Nighttime interior cab lighting must balance the controller's needs to see out the window and to read text, for example, flightstrip text.

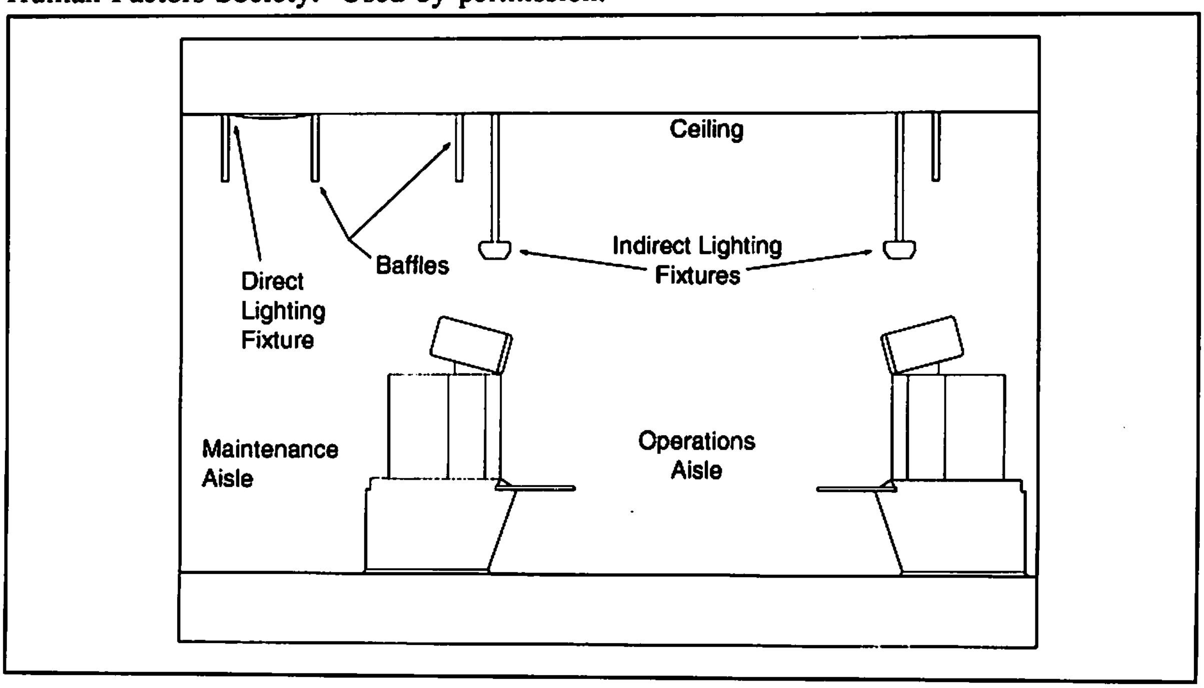
Designers and implementers of new ATC facilities and workstations must work closely with skilled lighting engineers and human factors engineers to accurately determine appropriate illumination levels.

Distribution of Light. In addition to the overall level of illumination, the distribution of light can affect the visual acuity, and performance of the air traffic controller. Differences in luminance between adjacent areas in the workplace can cause the eye to continuously adapt to these different levels as one shifts one's gaze from place to place (Sanders and McCormick, 1987). One key criterion is luminance ratio. The **luminance ratio** is the ratio of the luminances of two adjacent areas in the visual field. If the luminance ratios of two areas are too different, then acuity and comfort will be affected. For adjacent surroundings, the Illuminating Engineering Society recommends no greater than a three times difference between the brightest and darkest areas for adjacent surroundings in the workplace (Ref: IES ANSI/IES-1-1982).

Range of Luminance on the Workstation. Since displays on ATC workstations emit light, they do not need ambient illumination to be seen. General ambient illumination is required, however, to perform other tasks in conjunction with the display. For most tasks, the higher the ambient illumination, the better the controller's visual acuity will be. The opposite is true. however, for lighting for ATC workstation displays, in that the higher the ambient illumination level, the more difficult it will be to read the display. Therefore, the recommended illumination levels may have to be a compromise between the higher lighting levels needed for paper and pencil tasks and the lower levels desired for viewing CRT screens.

To avoid making extreme compromises one option is to use task lighting. This entails using a focused light source for use on paper and pencil tasks. Location and adjustment of the light source is critical to avoid glare and reflections on the workstation displays and controls. Alternatively, indirect lighting has been found satisfactory for ATC task performance, while avoiding unwanted glare or reflections (Krois, Lenorovitz, McKeon, Snyder, Tobey, and Bashinski, 1991). Figure 9-6 illustrates the suggested placement of direct and indirect lighting fixtures in an AAS ARTCC facility.

Figure 9-6. Recommended Lighting Scheme for an AAS ARTCC Facility
From "Air Traffic Control Facility Lighting" by P.A. Krois et al. (1991), *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 551-555). Copyright 1991 by the Human Factors Society. Used by permission.



9.6.2 What are the major sources of reflection and glare, and how can reflection and glare be reduced or eliminated?

Reflection and glare can be sources of distraction for the controller. They can also interfere with performance and cause physical symptoms, such as headaches and eye strain. Glare and reflections can interfere with the visual effectiveness of the display presentation causing delays in information transfer. The following is a summary of the kinds of problems caused by reflection and glare and some potential solutions that can be used to eliminate these problems.

Reflection refers to the amount of reflected light on a display screen compared to the amount of light emanating from the display (ANSI/HFS, 1988). The amount of light that falls on a work area or display screen is made up of two components: the amount of light directly striking the work area, and the amount of light that is reflected from walls, ceilings, and other surfaces (Sanders and McCormick, 1987). There are three main kinds of reflection:

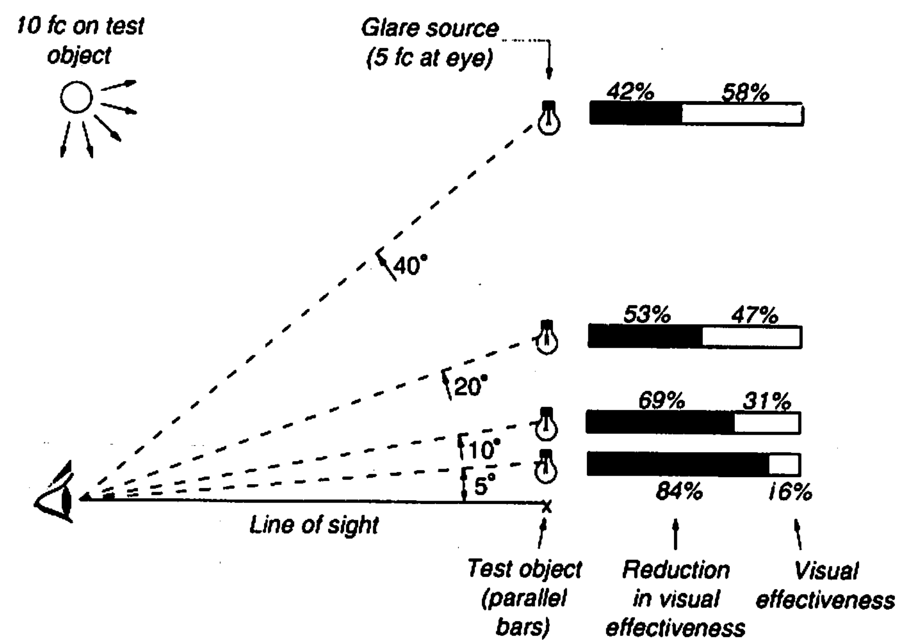
- *Specular reflection* in which light strikes a smooth, polished surface and is reflected in a mirror-like fashion.

- *Spread reflection* in which light strikes a rough unpolished surface, and the reflected light is spread and somewhat diffused.

- *Diffuse reflection* in which light strikes a flat painted or matte surface and is reflected equally in all directions.

Glare is scattered light that enters the retina. This scattered light reduces the contrast of the image or target being viewed. The effects of glare become greater as the source of the glare gets closer to the controller's line of sight. Figure 9-7 illustrates the effects of direct glare on visual effectiveness.

Figure 9-7. Effects of Direct Glare on Visual Effectiveness
From a study by M. Luckiesh and F. K. Moss as cited in *Human Factors in Engineering and Design* (6th Edition) by M. S. Sanders and E. J. McCormick, 1987, p. 415. Copyright 1987 by McGraw-Hill. Used by permission.



Reduction of Reflection and Glare. Several solutions are recommended to reduce the distracting effects of glare and reflections (Sanders and McCormick, 1987):

- Installing shields and filters around the light source and/or hoods around the display.
- Configuring the work areas so that reflected light will not be directed towards the eyes.
- Positioning the light source as far from the visual line-of-sight as possible.
- Reducing the intensity of the light source.

- Using diffuse or indirect light sources.

- Using surfaces that diffuse light such as flat paint on walls and non-glossy work surfaces.

Screen filters are often used to reduce glare on workstation displays. These filters attach to the face of the CRT. Light passes through the filter and is reflected off the CRT screen, then reflected back through the filter. Because the characters and graphics on the screen only pass through the filter once, the display contrast is improved. Unfortunately, when the display brightness is turned down, contrast and readability are reduced.

9.6.3 What are the optimal levels of temperature and humidity for operational environments?

Thermal comfort is an important aspect of the working environment in an ATC facility. The measurement of thermal comfort is difficult, however, because of the number of variables involved. In addition, there are individual differences in thermal comfort (in part, due to differences in metabolism). The following major factors affect thermal comfort:

- Air temperature
- Wall temperature or temperature of surrounding structure
- Humidity
- Ventilation (e.g., speed of recirculating air)
- Physical workload of the controller

Extreme Temperatures. One potential situation that may arise in an ATC facility is a temporary loss of heating or cooling that would force temperatures to unusual extremes. In the face of this situation, it would be useful to know the upper-most temperature limits at which a controller can safely control aircraft without a reduction in performance. Tasks requiring manual dexterity, such as data entry or target tracking, are affected by cold temperatures. The deterioration of performance is specifically associated with lower handskin temperature. Studies have shown decreases in performance with handskin temperatures between 55 degrees and 65 degrees fahrenheit (Riley and Cochran, 1984). Evidence about the effect of cold temperatures on mental activities is less clear, and no clear guideline has been established.

Humidity and Ventilation. Detailed technical guidance on humidity and ventilation is available in human engineering standards (e.g., DoD, 1989). The key human factors objective is to achieve a comfortable balance of temperature, humidity, and ventilation so that controllers are not subjected to environmental stressors. Such stressors can have adverse effects on physical and mental well-being and can detract from job performance.

Humidity. Relative humidity in the facility environment should vary with temperature (DoD, 1989). At 21 degrees Celsius (70 degrees F), the relative humidity should be about 45 percent. As temperature increases, the relative humidity should decrease. In any case, relative humidity should remain above 15 percent to prevent irritation and drying of the skin, eyes, and respiratory tract. Since the temperature in ATC facilities should remain fairly stable, relative humidity should not fluctuate significantly. For all ATC environments, and especially for tower environments, which can be affected by external heat and cold, provisions must be made to maintain temperature and humidity within comfortable levels.

Ventilation. Two-thirds of the air used for ventilating ATC facilities should be fresh, uncontaminated air (DoD, 1989). The amount of air required per minute varies as a function of the enclosure volume per person (See DoD, 1989, Figure 39). Air should be moved past operational personnel at a velocity that does not exceed 60 m (200 ft) per minute. To avoid drafts, controllers' chairs should not be in the direct line of air vents. The speed of the airflow must be adjusted so that papers or open manuals on work surfaces are not disturbed. Ventilation must be adequate to ensure that any fumes, vapors, dust or gases are kept within acceptable limits (DoD, 1989).

9.6.4 What are the acceptable levels of noise for operational environments and how are they measured?

DoD guidelines state that facilities requiring frequent telephone or radio use and direct communication of up to 5 feet shall have ambient noise levels which do not exceed 65 dB (DoD, 1989). Ambient noise is chiefly caused by vibration. Large capacity ventilation systems or other equipment often generate vibrations, but they can also come from the combined effects of conversations of people working in a facility.

Noise Control. Ambient noise can be controlled in one or more ways (Sanders and McCormick, 1993):

- At the source of the noise
- Along the path of the noise
- At the receiver of the noise.

Methods of noise control at the source include better design of equipment that will reduce the amount of noise generated. In addition, routine and proper maintenance of equipment in a facility can reduce noise. These actions include lubrication of moving parts and alignment, when necessary. Another method of reducing noise at the source is to isolate vibrating surfaces with flexible materials, such as rubber or elastomers.

Methods of noise control along the path of the noise include enclosing noisy equipment. High frequency vibrations are more directional than low frequency vibrations and are, therefore, more easily contained. Enclosing noisy equipment will reduce the amount of high frequency noise that will reach the air traffic controller. In addition, adding sound absorptive materials to walls, ceilings, and floors can reduce noise levels from 3 to 7 dB (Sanders and McCormick, 1987).

Methods of noise control at the receiving end chiefly consist of headphones or earplugs. Recent advances in technology have provided noise cancellation devices that use a technique called **Active Noise Reduction** (ANR). ANR systems sample the noise in the environment through an external microphone source. Once the noise is sampled, the system emits sound waves which are 180 degrees out of phase with the ambient noise. The effect is a cancellation of much of the noise at the controller's ear (Guttman, 1991). Unfortunately, this technology is not practical for most ATC applications for two reasons. First, a controller needs to be able to hear adjacent controllers. Second, ANR headsets require a higher signal-to-noise ratio than the passive headsets to maintain equal speech intelligibility (Gower and Casali, 1994).

9.7 REFERENCES AND SUGGESTIONS FOR FURTHER READING

Alpern, M. (1971). Effect of mechanisms in vision. In J. W. Kling and L. A. Riggs (Eds.) *Experimental psychology* (pp. 367-394) New York: Holt, Rinehart and Winston.

American National Standards Institute (ANSI)/Human Factors Society (HFS). (1988). *American national standard for human factors engineering of visual display terminal workstations* (ANSI/HFS 100-1988). Santa Monica, CA: Human Factors Society.

American National Standards Institute (ANSI). (1960). *Method of measuring phonetically balanced monosyllabic word intelligibility* (S3.2-1960). New York: Author.

American National Standards Institute (ANSI). (1969). *Methods for calculation of the articulation index* (Report 535). New York: Author.

Bashinski, H., Krois, P., Snyder, C., and Tobey, W. (1990). Lighting ergonomics evaluation issues for air traffic control facilities. In B. Das (Ed.), *Advances in industrial ergonomics and safety II* (pp. 947-954). New York: Taylor & Françis.

Begault, D. R., and Wenzel, E. M. (1993). Headphone locationization of speech. *Human Factors, 35*, 361-376.

Beranek, L. L. (1947). The design of speech communications systems. *Proceedings of the Institute of Radio Engineers* (New York), *35*, 880-890.

Black, P. (1992). PC radiation: How to cut your risk. *Business Week, November 30*, 116.

Boff, K. R., and Lincoln, J. E. (1988). *Engineering data compendium: Human perception and performance.* Wright-Patterson AFB, OH: Armstrong Aeromedical Research Laboratory (AAMRL).

Campbell, F. W., and Westheimer, G. (1960) Dynamics of accommodation responses of the human eye. *Journal of Physiology, 197*, 551-566.

Chapanis, A. (1965). Color names for color space. *American Scientist, 53*, 327-346.

Department of Defense. (1989). *Human engineering design criteria for military systems, equipment and facilities* (MIL-STD-1472D). Washington, DC: Author.

Ericsson, K. A., and Faivre, I. A. (1988). What's exceptional about exceptional abilities? In L. K. Obler and D. Fein (Eds.) *The exceptional brain: Neuropsychology of talent and special abilities* (pp. 436-473). New York: Guilford Press.

Geer, C. W. (1981). *Human engineering procedures guide* (AMRL-TR-81-35). Wright-Patterson Air Force Base, OH: Armstrong Aeromedical Research Laboratory.

Gower Jr., D. W. and Casali, J. G. Speech intelligibility and protective effectiveness of selected active noise reduction and conventional headsets. *Human Factors, 36* (2), 350-367.

Grandjean, E. (1988). Fitting the task to the man (4th ed.). New York: Taylor & Francis.

Guttman, J. A. (1991). *Speech recognition system technology assessment and lessons learned* (McDonnell Douglas Technical Report MDC 91Q0801). St. Louis, MO: McDonnell Douglas Aircraft.

Hawkins, H. L., and Presson, J. C. (1986). Auditory information processing. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. II. Cognitive processes and performance). New York: Wiley.

Helander, M. (1988). *Handbook of human-computer interaction.* New York: Elsevier/North Holland.

International Civil Aviation Organization. (1993). *Human factors in air traffic control.* (Human Factors Digest No. 8; Circular 241-AN/145). Montreal, Quebec, Canada: Author.

Jenkins, J. P. (1983). Human factors aspects of control room design. In C. M. Mitchell, P. M. Van Balen, and K. Moe (Eds.), *Human factors considerations in system design* (NASA Conference Publication 2246, pp. 25-46). Greenbelt, MD: National Aeronautics and Space Administration/Goddard Space Flight Center.

Kleeman, W., and Prunier, T. (1980). Evaluation of chairs used by air traffic controllers of the U. S. Federal Aviation Administration. In R. Easterby, K. Droemer, and D. Chaffin (Eds.), *Anthropometry and biomechanics: Theory and application* (pp. 235-239). London: Plenum Press.

Krois, P. A., Lenorovitz, D. R., McKean, P. S., Snyder, C. A., Tobey, W. K., and Bashinski, H. S. (1991). Air traffic control facility lighting. *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 551-555). Santa Monica, CA: Human Factors Society.

Luckiesh, M., and Moss, F. K. (1927-1932). The new science of seeing. *Interpreting the science of seeing into lighting practice* (Vol. 1). Cleveland: GE Company.
Lueder, R. (1986). Work station design. In R. Lueder (Ed.), *The Ergonomics payoff: Designing the electronic office*. Toronto, Ont., Canada: Holt, Rinehart & Winston.

McCormick, E. J., and Sanders, M. S. (1982). *Human factors in engineering and design* (5th ed.). New York: McGraw-Hill.

Mitchell, C. M., Bocast, A., and Stewart, L. J. (1983). Guidelines on ergonomic aspects of control rooms and highly automated environments. In C. M. Mitchell, P. M. Van Balen, and K. Moe (Eds.), *Human factors considerations in system design* (NASA Conference Publication 2246, pp. 127-202). Greenbelt, MD: NASA/Goddard Space Flight Center.

Murch, G. M., and Huber, J. (1982). Colour—the logical next step. *New Electronics, 15*, 31-32.

Nadler, E. DiSario, R. Mengert, P., and Sussman, E.D. (1990). *A simulation study of the effects of communication delay on air traffic control* (Report No. DOT/FAA/CT-90/6). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

National Aeronautics and Space Administration (NASA). (1989). *Man-systems integration standard* (NASA-STD-3000, Revision A). Houston, TX: NASA-Johnson Space Center.

Riley, M. W., and Cochrane, D. J., (1984). *Dexterity performance and reduced ambient temperature. Human Factors, 26*, 207-214.

Roscoe, S.N. (1992). From the roots to the branches of cockpit design: Problems, principles, products. *Human Factors Society Bulletin, 35* (12), 1-2.

Roush, W. (1993). Learning from technological disasters. *Technology Review, August/September*, 50-57.

Sanders, M. S., and McCormick, E. J. (1987). *Human factors in engineering and design* (6th ed.). New York: McGraw-Hill.

Sanders, M. S., and McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). New York: McGraw-Hill.

Schiffman, H. R. (1976). *Sensation and perception*. New York: Wiley.

Teichner, W. H. (1979). Color and visual information coding. *Proceedings of the Society for Information Display, 20*, 3-9.

Weale, R. A. (1961). Retinal illumination and age. *Transactions of the Illuminating Engineering Society, 56*, 95.

APPENDIX 9A. CHECKLIST ITEMS

Design Requirements and Controller Performance

1. Keys on keyboards and keypads have no more than two functions (9.2.2).

2. Workstation labeling is visible and understandable (9.2.2).

3. Keyboards and workstation controls are consistently backlit, if necessary (9.2.2).

4. Frequently used controls are easy to see and to reach (9.2.2).

User-Centered Workstation Design

5. Workstation dimensions are adequate for the extremes of the controller workforce (i.e., the 5th to the 95th percentile on applicable dimensions) (9.3.1, 9.4.1).

6. The workstation provides sufficient space for three-person teams (9.3.2).

7. Displays and controls are visible to the controller whether he or she is seated and standing (9.3.2, 9.3.3).

8. The workstation design concept considers the operational needs (e.g., space requirements) of the particular ATC facility for which it is intended (ATCT, TRACON, ARTCC) (9.3.2).

9. Workstations are arranged and spaced to allow ready access by AF personnel (9.3.2).

10. It is easy to open or remove equipment covers on racks (9.3.2).

11. Easy access is provided to workstation components (9.3.2).

12. Workstation controls are arranged in a logical order (9.3.3).

13. Controls are designed and arranged to be consistent with the controller's natural sequence of operational actions (9.3.3).

14. Controls that are logically linked, or used in sequence, are located near each other (9.3.3).

15. Controls used to adjust visual displays are located near the display set (9.3.3).

16. Adjacent controls are arranged so that sufficient space is between them to operate them easily and minimize the chances of accidental activation (9.3.3, 9.3.4).

17. Mirror-image layouts are avoided within and across ATC facilities (9.3.3).

18. The needs of individual facilities have been considered in determining space requirements (9.3.2).

19. Video display monitors are adjustable; they can be tilted in the vertical plane and swiveled in the horizontal plane (9.3.3).

20. Ceiling-mounted display monitors can be swiveled and tilted (9.3.3).

21. Controls are equally accessible and usable by left-and right-handed controllers (9.3.3).

22. High-priority controls are centrally located and placed as close as possible to the controller (9.3.4).

23. High-priority visual displays are centrally located (9.3.4).

24. Controls, indicators, displays and their labels are sufficiently visible under all anticipated lighting conditions (9.3.4, 9.6.1).

25. Controls can be reached without excessive shoulder movement or back bending/stretching (9.3.3, 9.3.4).

26. Control and display layouts accommodate the extremes of reach dimensions found in the workforce (9.3.4).

<u>Design of Control Room Seating</u>

27. Chairs provided for controller workstations are comfortable and support the lower back (9.4.2).

28. Seat height is adjustable within 15 to 21 inches (9.4.2).

29. The seat's backrest reclines between 100 and 115 degrees (9.4.2).

30. The seat and backrest are adequately cushioned (9.4.2).

31. Arm rests are 2 inches wide and 8 inches long (9.4.2).

Design of Communications Equipment

32. System-input devices are designed for optimal response to the range of frequencies between 200 and 6,100 Hz (9.5.1).

33. Across the frequency response bandwidth, amplitude variation is at or below plus or minus 3 dB (9.5.1).

34. Microphones used with amplifiers have a dynamic range that permits them to pick up variations of at least 50 dB in signal input (9.5.1).

35. Appropriate techniques are used to minimize the effects of noise during speech transmission (9.5.1).

36. There are no noticeable squeal problems or echo effects (9.5.1).

37. Appropriate provision has been made for the calibration of microphones and headphones (9.5.1).

38. Listeners can differentiate adequately between multiple channels fed into headphones (9.5.2).

39. Messages conveyed over loudspeakers are adequately intelligible (9.5.2).

40. Normal voice communication is unaffected by positive feedback noise (9.5.2).

41. Effective communication is not impaired by delays due to satellite transmission (9.5.2).

42. The speaker is not distracted by his/her own side tone (9.5.2).

43. Headset design helps to improve intelligibility (9.5.2).

44. If intelligibility testing is conducted, scores are at or above the following cutoff points for the various testing methods (ANSI phonetically balanced (PB) - 90%; Modified Rhyme Test - 97%; Articulation Index - 0.7) (9.5.2).

45. No bare metal parts of the headset come into contact with the controller's skin (9.5.4).

46. Controllers who wear glasses can comfortably wear headphones or other communication equipment (9.5.4).

47. Hands-free operation of communication equipment is possible under normal working conditions (9.5.4).

48. Telephone handsets are readily accessible (9.5.4).

49. For multiple telephone handsets, the most frequently used or the most urgently needed handset is the most readily accessible (9.5.4).

50. Volume/gain controls are separate from on-off controls (9.5.5).

51. Volume/gain controls are limited to an audible level (9.5.5).

52. When two earphones are in use, sound pressure level can be increased to at least 100 dB overall (9.5.5).

53. Squelch control is provided to suppress channel noise during inactive periods (9.5.5).

54. The controller can manually deactivate the squelch control (9.5.5).

55. Foot pedals are provided as alternatives to hand-activated microphone switches (9.5.5).

56. Human factors issues associated with foot pedals (e.g., location of the pedal's fulcrum, placement of the foot pedal in relation to other control, placement of the foot pedal in relation to the seated controller) have been investigated and resolved (9.5.5).

Environmental Design

57. Ambient lighting of the workstation is adequate and adjustable (9.6.1).

58. The controller can easily recognize all key labels on the keyboard, electronic display keypads, and trackball (9.6.1).

59. Lighting at the console shelf is adequate for reading and writing (9.6.1).

60. Under the proposed lighting conditions, the controller can readily locate switches, controls, headset jacks, connectors, handles, and display recess mechanisms (9.6.1).

61. The controller is not distracted by shadows, glare, or reflections (9.6.1, 9.6.2).

62. Color displays are readable and adequately bright under the proposed lighting conditions (9.6.1).

63. The controller is able to identify color codes under all anticipated lighting conditions (9.6.1).

64. Adequate display contrast is maintained for textual and graphic information (9.6.1).

65. Lighting is adequate for emergency and maintenance purposes (9.6.1).

66. Maintenance lighting does not interfere with controller tasks at the ATC console (9.6.1).

67. The brightest area in the workplace is no more than three times brighter than the darkest area (9.6.1).

68. Levels of temperature, humidity, and ventilation provide adequate thermal comfort (9.6.3).

69. Ambient noise is at or below 65 dB (9.6.4).

CHAPTER 10. HUMAN FACTORS TESTING AND EVALUATION

Kim M. Cardosi

"In the past, controllers have sometimes accepted and adapted to changes which have failed to fulfill their initial promise, and that controller has had to compensate for deficiencies in the system. Measures are therefore necessary to indicate whether such potential problems may arise in relation to any proposed change, and preferably to give guidance on how they might be overcome."

Hopkin, 1980

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 10. HUMAN FACTORS TESTING AND EVALUATION

10.1 INTRODUCTION

Many different types of questions are best answered with the results of a human factors test. Following are some of the most common human factors questions:

- Which of two or more proposed designs (of displays, controls, training programs, etc.) is best from a human factors standpoint?
- What performance benefits are achieved from a specified design change?
- Is a design of a new system or subsystem viable from a human factors perspective?
- What changes, if any, need to be made to a prototype system to minimize operator error?
- Is a proposed training program (e.g., for new equipment) adequate?
- How long will it take for a controller to perform a task, or part of a task, with a new system?

Human factors specialists, working with operations specialists, can often anticipate human factors problems by examining specifications documents, proposed designs, and prototypes of new systems and subsystems. Still, human factors tests are often required to identify problems that are not self-evident or to be able to quantify the impact of new systems on line operations. Formal evaluations are always needed to ensure that the new system or procedure is ready for implementation.

This chapter will address the following questions.

- When is a human factors test warranted?

- How is operator performance measured?
- What method of testing should be employed?
- How should test results be analyzed and interpreted?

Understanding the principles and philosophy behind human factors testing is useful even to people who never conduct human factors tests because it helps operations specialists critique tests conducted by manufacturers, universities or industrial labs and determine the validity of their conclusions.

10.2 TO TEST OR NOT TO TEST

It isn't always easy to predict all of the ways in which an operator will use or misuse a new system or a new component of an existing system. Nor is it always evident what types of errors operators are likely to make.

One example of a faulty display design that should never have made it to implementation is the case of a major air carrier that wanted to give flight attendants a cue as to when sterile cockpit rules were in effect (i.e., when only safety-critical communications are permitted in the cockpit and between the cockpit and cabin crews). The airline installed a small indicator light above the cockpit door that was to be illuminated when sterile cockpit rules were in effect. Problems arose because the light that was chosen was green. In most cultures, the color green is *not* associated with "stop" or "no admittance." The lights had to be changed to red, at no small expense to the airline. In that case, a formal human factors test was not needed to predict the problems experienced by the airline. It is common knowledge that a green light is most likely to be interpreted as "go" or "enter." Most questions about training, displays, controls, and how operators may use or abuse them are much more complex, however, and require controlled testing to be answered effectively.

Basic research findings, such as information about our sensory and cognitive capabilities and limitations, can help us identify desirable design options. However, each specific application of a technology, training program or procedure should be evaluated under the same or similar conditions as it will be used, by the same type of operator that will be using it, and while the operators are performing the same types of tasks that actual operations require.

When a human factors evaluation of a system or subsystem is warranted, it should be designed by a team that includes human factors specialists and operations specialists.[1] Operations specialists are intimately familiar with the operational environment (e.g., a specific cockpit or ATC facility). They represent the potential users and are usually operators (e.g., pilots or controllers). As long as they are operationally current (i.e., knowledgeable of current issues, procedures, and practices), they are the most appropriate source for information on user preferences and suggestions for symbology, terminology, display layout, and other factors.

Even the most experienced users should not be solely responsible for the user-machine interface, however. In fact, many years of experience can occasionally be a liability in making such decisions, since the skills and knowledge that develop with extensive experience can often compensate for design flaws that may then remain unnoticed.

For these and other reasons, it is important for operations specialists to work with human factors specialists in the planning and conduct of a human factors test. Human factors specialists are intimately familiar with the capabilities and limitations of the human system, testing methods, and appropriate data analysis techniques. They can point out potential problems that operational specialists might overlook. Working together, the two types of specialists can predict problems and head them off before they occur in actual operations. Together, human factors specialists and operations specialists are equipped to decide exactly what needs to be tested and how it should be tested.

10.3 HUMAN PERFORMANCE MEASUREMENT

"Most human factors measures have been applied at some time to air traffic controllers, but none has gained general acceptance as pertinent, valid, and complete."
(Hopkin, 1980, p. 547).

One interesting measure of controller performance examines the debriefings that occur when one controller takes over from another controller (Hopkin, 1980). Test measurement could involve such factors as the time required before the incoming controller feels prepared to take over and the amount of information that must be verbally conveyed because it is not readily apparent. As Hopkin points out, the busy controller relies on remembered, as well as displayed, information when performing essential tasks. S/he may use notations and abbreviations that serve well as self-memory aids, but are useless to an incoming controller. These measures could not be used without certain controls, since they would vary with extraneous factors, such as the experience, skill, and confidences levels of both controllers, and not just with the skill level and workload of the relieved controller.

Clearly, we will not be able to devise measures of controller performance that will be suitable for studying every question from, "How do we train people to be good controllers?" to "Which of two types of systems will be most efficient in helping to reduce the number of runway incursions?" There are, however, several measures of human performance that have wide applications in studies of ATC issues.

Measures of human performance can be subjective or objective. Subjective measures use responses that are measured in units that are defined by the individual. Because subjective measures will differ with personal criteria, it is important to make the reporting scale as standardized and structured as possible. An example of a subjective measure of workload is a controller's opinion as to how difficult a task was. What constitutes a high workload situation for one person may not be considered high workload for another person. Subjective measures are used whenever objective measures are unavailable or inappropriate. They are also used to complement objective measures. Valid, reliable subjective measure yield quantitative data that can be submitted to statistical analysis.

Objective measures of human performance use units that are clearly defined, such as number of errors, traffic counts, heart beats per minute, and blood pressure. The most commonly used objective measures of performance are response accuracy and response time. Response time measures the time required for a person to perform a specific task, or component of a task. Response accuracy measures the percentage of errors made while completing the task or the precision with which a specific task is accomplished (e.g., flying a pre-determined route, as measured by cross-track error).

It is important to point out that there is some level of subjectivity even in the most seemingly objective measure because the units and the equipment for measuring them are defined and designed by human beings. Both subjective and objective measures are useful for different purposes and often provide valuable information when used together.

The numbers of errors or omissions, when taken in isolation, are not very informative measures of controller performance. Complementary subjective measures are necessary to determine, for example, whether the errors were due to complacency and boredom, extremely high workload, task complexity, or a combination of factors. Another problem with measuring response accuracy only is that it is possible to obtain insignificant results due to either a ceiling effect or a floor effect. For example, if the response being measured is so highly practiced that errors are few (e.g., a baseline of 95 percent accuracy), any manipulated factor is not likely to have an observable effect. This is the **ceiling effect.**

Conversely, initial performance may be so poor that any manipulation will not have a measurable effect. The tests may not be sensitive enough to measure an effect beyond this very high or very low baseline. A similar problem is encountered when attempting to measure system performance in terms of aircraft accidents, near mid-air collisions, or operational errors. The baseline rate of such incidents is often too small to measure the effect of any given factor. While this is good news for the flying public, it poses measurement problems for researchers trying to measure the success of an intervention program.

Generally, if baseline performance on the measured task is extremely accurate, and it is not desirable to induce more errors by manipulating other factors (e.g., workload), then response time is generally a more sensitive measure than error rates. Differences in the response times may be observable even when the differences in response accuracy are not.

While response time appears to be a simple measure of human performance, it is actually quite complex. Response times have several components and each of these components can be affected by many different factors. (See Chapter 5, Human Information Processing, for a complete discussion of these issues.)

10.4 TEST METHODS

10.4.1 What method of testing should be used?

The testing method of choice depends on the specific problem or question under investigation and the available resources. Most importantly, the method must be appropriate to the issue. For example, one would not consider a questionnaire for measuring the time required to complete a small task, nor would one collect data on controller eye movements by asking controllers where and when they moved their eyes.

Another necessary consideration is the amount and type of testing resources available. Often, the most desirable type of test is too expensive and many compromises are necessary. The implications of these compromises need to be recognized as do their implications for the interpretation of the test results.

Field Observations

One evaluation technique that is often used is field observation. This includes any over-the-shoulder evaluations, such as sitting behind a controller and observing a specific controller activity, or sitting behind a controller team (A side and D side) and observing their interactions. One advantage to this method is that it allows investigators to make observations in the most natural setting possible. It can increase understanding of the nature of processes and problems in the work environment. Specifically, valuable insights can be gained as to where problems might occur with a specific system or procedure and why they might occur.

One task in which field observations are helpful is in trying to determine the information or cues that people use in performing a task. We, as humans, are rarely aware of all of the information that we use in performing a task. This is illustrated in a "problem" that Boeing Commercial Airplanes once had with one of their engineering simulators. After flying the simulator, one pilot reported that, "It felt right last week, but it just doesn't feel right this week." The mechanics examined everything that could possibly affect the handling qualities of the simulator. They took much of it apart and put it back together. They fine-tuned a few things, but made no substantive changes. The pilot flew the simulator again, but again reported that it still didn't "feel right." It seemed a little better, but it just wasn't right. Someone finally realized that the engine noise had inadvertently been turned off. The engine noise was turned back on and suddenly, the simulator once again "handled" like the aircraft (Fadden, personal communication).

While field observations are often useful as an initial investigative technique into a problem, their limitations often preclude broad objective conclusions. Field observation findings may be more subjective than objective, are dependent on the conditions under which the observations were made, and can actually be affected by the observation process itself.

One factor that affects the reliability of findings based on field observations is the number of observations made. For example, a conclusion based on 40 hours of voice communications is going to be more reliable (i.e., more repeatable) than one based on three hours. Furthermore, the findings based on field observations are condition-dependent. That is, **the findings must be qualified with respect to the specific conditions under which the observations were made.** For example, if observations of controller behavior were made only during periods of low to moderate workload, then the findings based on these data might not apply to situations involving high workload.

Another, and more subtle, consideration is that the very process of observation can alter what is being observed. An observer's activities, or even his or her mere presence, can affect performance. For example, depending on who the observer is (and their stated or implied mission), controllers may change their behaviors. They may, for example, become more conscientious (e.g., about phraseology). It is easy to envision how different observers (e.g., a university researcher or a supervisor conducting an over-the-shoulder evaluation) might observe slightly different behaviors exhibited by the same controller, all of which may be different from what occurs when no observer is present.

Another possibility is that the observer's presence might make a controller (particularly a new controller) nervous and induce a classic case of "checkitis." In this case, performance would be poorer than when no observer is present. Observers, or their questions, may also be distracting and this may adversely affect performance.

Questionnaires

Questionnaires are important research tools that allow investigators to collect information from many people at minimum cost. They are useful in surveying user opinion, local procedures, and individual practices and preferences, among other issues. Developing a useful questionnaire is not a simple process, however. There are experts available in questionnaire development. While anyone can call themselves an expert in questionnaire development, a successful consultant in questionnaire design is likely to have an advanced degree and relevant experience. There are also guidelines for developing and administering useful questionnaires (Kidder, 1981).

The first rule of questionnaire design is that questions should be simple and direct. The probability of confusing questions resulting in different people interpreting the questions or rating scales differently should be minimized. Confusing or ambiguous questions should be eliminated. The best way to accomplish this is to administer the questionnaire to a small group of individuals who are part of the target population (e.g., controllers) and see how they interpret the questions. It is also very helpful to ask for their feedback on the format of the questionnaire, the clarity of the questions, and other related issues.

While it is true that the best questions are simple and direct, care must be taken in the specific wording of the questions. A question with an obviously desirable answer will not yield informative results. For example, in a survey on cockpit and cabin crew coordination, Cardosi and Huntley (1988) wanted to assess crewmembers' knowledge of sterile cockpit procedures. The most direct question, "Do you know your airlines's procedure for sterile cockpit?," would probably have resulted in crewmembers answering in the affirmative, whether or not they were certain of the procedure. Instead, we asked, "What is your airline's procedure for sterile cockpit?" It was an interesting finding in itself that different crewmembers from the same airlines gave different answers. Second, **the questions need to be unbiased, both individually and as a set.** Individual questions can be biased in terms of their wording. For example, asking "How much easier is it to use trackball X than trackball Y?," presumes that trackball X is easier. Respondents are unlikely to report that X is actually more difficult. An unbiased way to present the same question is, "Compare the ease of using trackball X to using trackball Y." This question would be answered with a scale ranging from "X is more difficult" to "Y is more difficult" with a midpoint of "X and Y are the same."

Just as any individual question can be biased, a questionnaire may also be biased in its entirety. For example, if there are more questions about possible problems with a system than about its advantages, respondents may report feeling less favorably toward the system than if the questionnaire had more positive than negative questions.

Finally, **the questionnaire should be administered as soon as possible after the experience or task that is under investigation.** Because memory for detail can be very fleeting, it would not be advisable to show a controller a new display, and then a week later administer the questionnaire. The sooner after exposure the questionnaire is administered, the more useful the results are likely to be.

One exception to this rule is a questionnaire that is used to examine the effectiveness of a training program, that is, how much of the information that is presented in training is retained over a given period of time. For such a "test", a significant time interval (e.g., one month or longer) between exposure to the training and the questionnaire would be useful. A test with such a delay would be more effective than a test with no delay in predicting what information will be remembered and accessible for use when needed in actual operations.

Rating Scales

Rating scales are often very useful. Most scales offer five or seven choices. **Fewer than five choices is confining; more than seven choices makes it difficult to define the differences between consecutive numbers on the scale.**

Unless it is desirable to force questionnaire respondents to choose between two alternatives, rating scales should always have a mid-point. (This is one reason why an odd number of choices is recommended.) The scale should also have descriptive "anchors", that is, at least both ends and the middle values should have a word or phrase that identifies exactly what is meant by that number. This helps to minimize differences in people's own standards. For example, if the questions asks for a rating of the ease or difficulty of the use of a system Y as compared to system X, anchors should be given where the number "1" means "Much easier than X"; the number "3" corresponds to "No difference" and "5" means "Much more difficult than X." The results will be easier to interpret and, therefore, much more valuable, than those obtained by simply asking for a rating of ease or difficulty on a scale of one to five.

While user opinion is extremely valuable, there are many problems with making important design decisions by vote or consensus alone. We, as humans, are not very good at estimating our own response times, or predicting our own errors; nor do our initial preferences always match what will be most efficient in actual operations. Furthermore, there is also a tendency to prefer what is most familiar to us. Initial perceptions of new systems or subsystems may change with experience.

Even simple behaviors do not lend themselves to accurate judgments about our own actions. As part of an evaluation of a prototype navigation display, the Boeing flight deck integration team monitored pilots' eye movements as they used a prototype navigation display. The team also asked pilots to report where they thought they were spending most of their time looking. There was no systematic relation between where the pilots thought they were looking the most and where the data actually showed that they were looking most (Fadden, personal communication). The same kind of dissociation has been found in studies of controllers' visual scanning behavior; controllers report different scanning patterns than those detected by measuring instruments (Potter, personal communication). This simply shows that peoples perception of their simple actions are not always accurate.

Laboratory Experiments

It is difficult, if not impossible, to investigate issues by manipulating factors in actual operations. Such control is usually only available in a laboratory setting. The goal of an experiment is to manipulate the variables under investigation while keeping everything else constant. This careful manipulation of the key variables allows investigators to determine which of them has an effect.

One common type of laboratory experiment is a part-task simulation. In a part-task simulation, controllers use the component that is being tested in isolation. Part-task simulations are useful for studying simple questions, such as: "How long does it take to notice a particular change in the display?," or "Will the user immediately know what that symbols mean?"

A part-task simulation is an ideal way to conduct an in-depth test of a new display. It allows attention to be focused on the details of the display before it is tested in a more comprehensive manner. In addition to providing valuable initial results, a part-task simulation often points to specific areas that need to be tested further.

In a full-mission simulation, controllers perform all of their normal tasks (e.g., separate traffic and communicate with aircraft, coordinate with other controllers) as they would during their shift. The full-mission simulation is, of course, a very desirable type of test because it preserves the most realism, and thus, yields results that are easy to generalize to the real world. Full-mission simulation can give the same degree of control as a laboratory experiment, with the added benefits afforded by the realism.

The major drawback of full-mission simulation is that it is very expensive. The costs for computer time, simulator time, and salary for participating controllers, in addition to the other costs of research, can be prohibitive for all but the largest, and most well-funded, projects.

Another limitation of simulation studies that must be considered when interpreting the results is the **priming effect**. When controllers walk into a simulation knowing that they are going to participate in a test of Warning System X, they are expecting to see that system activated. They will see System X activated more times in one hour than they are likely to see in an entire shift. This expectation leads to a priming effect, which yields faster response times than can be expected when the activation of System X is not anticipated. For this reason, **the response times obtained in**

simulation studies can be faster than would be expected in the real-world and must often be considered as examples of best-case performance. How much faster response times will be in simulation than in actual operations is difficult to say as it depends on a variety of factors, particularly the specific task.

In addition to response times being faster, they are also more homogeneous in simulation studies than would be expected in actual operations. This reduced variability can result in a higher likelihood of obtaining a statistically significant difference between two groups or conditions in a simulation study than in actual operations. Since data obtained in actual operations are rarely obtainable, however, data from realistic simulation studies are a good alternative.

10.4.2 How can tests be assessed for validity and reliability?

The goal of any evaluation is to have reliable and valid results. **Reliability** refers to the repeatability of the results. If other investigators were to run the same test with the same equipment and same type of test participants, what are the chances that they would get the same results? To have repeatable results, all extraneous factors must be carefully controlled and held constant.

In any experiment, it is necessary to carefully manipulate the factors that will be examined in the study and control all other variables, if only by keeping them constant. For the results of experiment to be useful, they must be due to the factors that were manipulated and not to extraneous factors, chance, or anything peculiar to the testing situation or individuals tested. Careful controls help to ensure useful results.

Validity refers to whether the test measures what the test purports to measure. For tests to be valid, they must provide accurate measures of what they say they are measuring. For example, do tests use to select employees, actually predict how well the person will be able to do the job? Do the Standardized Aptitude Tests (SATs) actually measure one's ability to succeed in college? If the answer to this type of question is "no", then the test is not valid.

Experimental Controls. One way to help ensure that the results of a study are valid and reliable is to employ careful controls of critical factors of interest and of extraneous factors (e.g., fatigued participants) that may influence the results of the study. This is easier said then done because it is often very difficult to even identify all of the factors that may contribute to test results. Careful selection of test participants and testing conditions, however, in addition to sound experimental design and standardized execution of test procedures, will help to ensure valid and reliable results. A sound experimental design ensures that an adequate number of test participants ("subjects") are properly selected and tested (in an appropriate number and order of conditions) and that careful controls of the variables are included in the test.

Operationally-Defined Variables. One fundamental component of an evaluation that often gets neglected is the idea that test variables be operationally defined. This means that **the factors under investigation must be defined in ways that can be measured.** For example, a test to determine whether the use of the Traffic Alert and Collision Avoidance System (TCAS) increases ATC frequency congestion would begin with an operational definition of frequency congestion. A suitable measure, in this case, would be the number of ATC calls generated by TCAS-equipped aircraft (e.g., pilots contacting ATC to inform the controller of a maneuver or ask a question concerning a traffic alert) per unit time as compared to the number of traffic related calls generated by aircraft without TCAS under similar conditions.

Whether a test is designed to examine something simple, such as display clutter, or complex, such as situational awareness, all variables must be defined in terms of units that can be measured in the study.

Representative Subject Pool. Another necessary component of an evaluation is a representative subject pool. ("Subject" is the term used by researchers to refer to a participant in a test or study.) Most research on basic perceptual and cognitive processes is conducted using college students as subjects, since they are readily available at universities where much of the work is conducted. Because of this, the question often arises as to whether or not we may generalize the results to specific populations, such as pilots or controllers.

One rule of thumb is that, if the study purports to examine an aspect of behavior in which the target population would be expected to be different from college students in key ways, then the results will not be applicable. The differences between the target and test populations may be in terms of physical differences (e.g., age), or intellectual abilities (e.g,. specific skills or knowledge). Whether or not these differences prevent a generalization of test results depends upon the task. These differences can be quite subtle, but important.

For example, one approach to studying the similar call-sign problem might involve determining which numbers are most likely to be confused when presented aurally. A sample research question would be, "Is 225 more likely to be confused with 252 or 272?" This is a relatively simple question and the results of the test would comprise a confusability matrix. Because this is a simple auditory task, pilots would not be expected to perform much differently than college students, with the exception of the differences attributable to hearing loss due to age and exposure to noise. In this case, performance depends primarily on the ability to hear the differences between numbers and results of experiments performed with college students as subjects are likely to be applicable to pilots.

Now consider a superficially similar, but technically very different, task. If the experimental task was to look at the effect of numerical grouping on memory for air traffic control messages, subjects might listen to messages with numerical information presented sequentially (e.g., "Descend and maintain one one, zero thousand. Reduce speed to two zero knots. Contact Boston Approach one one niner point six five."), and messages with numerical information presented in grouped form (e.g., "Descend and maintain ten thousand. Reduce speed to two twenty knots. Contact Boston Approach one nineteen point sixty-five.") Since a pilot's memory for that type of information is going to be very different from a college student's memory of that information, mostly because it is meaningful to the pilot, results obtained by using college students would probably not be directly applicable to pilot populations.

One important aspect in which subjects should be representative of the target population is in terms of skill level. It is highly unlikely that an experienced controller can successfully train himself to react or think like a new trainee. A below-average controller (or an average controller on a bad day) is likely to experience more difficulties with a new system than a skilled controller. It is very difficult for a highly experienced operator to predict how people without prior knowledge or specific experiences will perform a certain task or what mistakes they are likely to make. Exceptional skill can enable an operator to compensate for design flaws—design flaws which, because of the skill, may go unnoticed.

Controlling Bias. While it is important that the people used as subjects are as similar as possible to the people to whom you want to generalize the results, it is also important that the subjects' and experimenters' biases don't affect the results of the test. If participants have their own ideas as to how the results should come out, it is possible for them to influence the results, either intentionally or unintentionally. It is not unusual for subjects to be able to discern the "desirable" test outcome and respond accordingly. To prevent this, investigators must take steps to control subject bias. For example, studies designed to test the effectiveness of a new drug often employ a control group that receives a placebo (sugar pill). None of the subjects knows whether he or she is in the group receiving the new drug or in the group given the placebo. Drug studies typically go a step further to control bias and conduct "double-blind" studies. This means that even the experimenters who deal with the subjects do not know who is receiving the placebo and who is receiving the drug.

In aviation applications, it is usually impossible to conduct a test (e.g., of new equipment) without the participants knowing the purpose of the test. Furthermore, this is often undesirable, since subjects' opinions (e.g., of a new display) can be a vital component of the data.

One solution to the problem of controlling or balancing the effects of biases and expectations is the use of a **control group**. This group of subjects is tested under the same conditions (and presumably would have the same expectations) as the experimental group, but is not exposed to the tested variable.

For example, consider a test designed to examine the effectiveness of a program to teach effective communications techniques to controllers. The effectiveness will be defined as a significant reduction in the numbers of readback and hearback errors experienced by individual controllers per hour. If the training program is to be compared to an existing program, then the performance of controllers who were trained in each program could be compared. Controllers trained in the new program would be the experimental group, and controllers trained in the existing program would constitute the control group.

If the training program was a prototype and there was no such comparison to be made, then the performance of one group of controllers could be monitored before and after the exposure to the training program. Alternatively, the performance of controllers trained with the new program (the experimental group) could be compared to that of controllers who did not receive this training (the control group). In this case, however, it would be important to control for the other conditions that can contribute to hearback and readback errors (e.g., knowing that your errors are being recorded, workload levels, traffic mix, fatigue, etc.).

For the comparison between the two groups to be meaningful, all measures of errors would need to be taken either with or without the knowledge of the controllers; a comparison between a trained group that knew they were being monitored and a control group that did not know they were being monitored would be meaningless.

Representative Test Conditions. It is usually desirable for the test conditions to be as representative as possible to "real world" conditions. While the engineer looks at a system and asks, "Does it perform its intended function?", the human factors specialist wants to know if the controllers or other operators are able to use the system effectively under the conditions under which it will be used. Because of this, the key conditions included in the test must be as representative as possible to actual operating conditions so that the results of the test can be generalized to actual operations. Important conditions may include, but are not limited to,: varied workload levels, traffic loads and complexity, weather conditions, ambient illumination levels (i.e., lighting conditions), and ambient noise conditions. For example, if a display is designed to be used in an ATC tower, then it is important to ensure that it is easily read in a wide variety of lighting conditions (e.g., from direct sunlight to dim lighting).

It is often important to include the "worst-case" scenario in addition to representative conditions in a test. Most human factors evaluations must include a worst case test condition, since it is the worst case (e.g., combination of failures) that often results in a dangerous outcome. For example, if it is important that a time-critical warning system be usable in all conditions, then the operator response time that is assumed by the software's algorithm needs to take this into account. In this case, in addition to measuring how long will it take the average person under average conditions to respond to the system, the longest possible response time, or response time at the 90th, 95th, or 99th percentile, should also be measured. Such worst case response times should be obtained under worst case conditions, such as very high workload.

Counter-balancing. One control that is not necessary in the engineering world but that can be critical in the human factors world is counter-balancing. When testing physical characteristics, such as height and weight of a person, or the range of two different radar systems, it doesn't matter which one is tested first; the test of the first system or characteristic will not affect the outcome of the test of the second. When testing human performance, however, order effects are common due to such factors as learning and fatigue.

There are two possibilities of how human performance can change during the course of the test–it can get better or worse. Performance may improve because exposure to the first system gives subjects information that helps them in using the second system. This is called **positive transfer**. For example, in a test of two data input devices, it would be reasonable to have controllers use each of them and measure the time required to perform specific tasks (i.e., response time) with each system. The number of errors made in the data input process (i.e., response accuracy) would also be measured. Performance with System A could be compared to performance with System B to determine which of the two systems is preferable.

If the procedures for two systems are similar (e.g., in terms of keypad layout and the order of the information input) but new to subject controllers, then the practice acquired during test of System A might improve their performance with System B over what it would have been without the experience gained during the first test.

If, however, the two systems are physically similar, but require different procedures to operate, then the experience acquired with the use (test) of System A would probably impair performance with System B. Performance with System B would have been better with no previous experience with System A. This phenomenon is referred to as **negative transfer**. Negative transfer is not limited to testing situations, it also occurs in learning new tasks, or performing familiar tasks with new equipment. (See discussion of negative transfer on p. 181.)

One way to avoid the possibility of positive or negative transfer influencing test results is to balance the order of conditions. For example, in a comparison of two radar displays, a test could be conducted in which each controller uses both displays, with half using Display A first and half using Display B first. If there are more than two conditions or scenarios (e.g., different levels of workload), then the order should be randomized for each set of subject controllers. For example, if the operational testing and evaluation includes three weeks of testing with a different group of controllers participating as subjects each week, then the order of the scenarios (e.g., representing low, medium, and high workload) would be randomized for each group. An alternative to counter-balancing is to test one display with half of the controllers and the other display with the controllers. In this case, it is particularly important to ensure that there are no important differences in the two controller groups (e.g., in terms of skill level).

There is another reason why performance may deteriorate over the course of a test. If the test is extremely long or the task is very tedious, performance may suffer due to a fatigue effect. When fatigue may be a factor in any test, careful controls, such as the use of an appropriate control group or balancing or randomizing the order of conditions, must be considered.

A study, intended to investigate the effects of fatigue on flight crews and flight crew errors, illustrates this point. In this investigation of the effects of fatigue on flight crew errors, two groups of active line pilots flew a scenario similar to a line flight in a full-mission simulation (Foushee, Lauber, Baetge, and Acomb, 1986). Ten flightcrews flew the scenario within two to three hours after completing a three-day, high-density, short-haul duty cycle. The other ten flightcrews flew the test scenario after a minimum of three days off. The results showed that while the "post-duty" crews were more fatigued than the "pre-duty" crews, their performance was significantly better than that of the pre-duty crews. The better performance was not attributable to fatigue, but to a personal familiarity that developed over their duty cycle. The crews who had flown together on the duty cycle prior to the simulation got to know each other and knew what to expect from each other. This is often considered to be what led to the birth of **cockpit (or crew) resource management.**

The first part of this study did not have a control group of pilots who flew together for the same amount of time right before the simulation, but who were not fatigued. In the second part of this study, a subsequent analysis of the data showed that the superior performance was, indeed, due to familiarity with the other crewmembers and not due to fatigue.

10.5 ANALYSIS OF TEST RESULTS

Once a human factors test has been conducted, the next step is to analyze results and present them in the simplest and most straightforward manner. The goals of data analysis are to describe the results and, where applicable, to determine whether there are important differences between groups or conditions of interest.

10.5.1 What are descriptive statistics?

The methods used to organize, summarize, and communicate the meaning of a large set of numbers (e.g., response times or error rates) with the fewest possible numbers are **descriptive statistics.** Following are brief introductions to various types of descriptive statistics likely to be used for testing and evaluation.

Measures of Central Tendency: Measures of central tendency seek to describe a set of data (e.g., a set of reaction times) with a single value. The most commonly cited measures of central tendency are the arithmetic mean, the median, and the mode.

The mean: The mean is computed as the sum of all the scores (e.g., response times or error rates) divided by the number of scores. For example, suppose we wanted to assess how much time should be allotted for controllers to notice a subtle change in a display (e.g., a change in a visual indicator from a normal status to a "warning" status).

One method of determining the time required, would be to conduct a simulation study and measure the time between the change in the display and the controller's response to the change (or some other indication that they noticed the change). In this example, assume that each controller saw the change in the display in eight different scenarios, so that we have eight different response times or "scores" (in seconds) for each controller. One way of describing the time required for an individual controller to notice the changes in the display would be to calculate the average (or mean) response time for that controller. For the sake of this example, assume that we have the following eight scores (in seconds):

5.5, 6.2, 6.8, 8.5, 10.2, 10.4, and 14.2;

the mean would be calculated as: (5.5 + 6.2 + 6.8 + 8.5 + 10.2 + 10.4 + 14.2)/8 or 8.6 seconds.

The mean is considered to be the fulcrum of a data set because the deviations in scores above it balances the deviation in scores below it. The sum of the deviations about the mean is always zero. Because of this, the mean is very sensitive to outlying scores, that is, scores that are very different from the rest. A very high or very low score will tend to pull the mean in the direction of that score. While the mean is more frequently cited than the median or the mode, it is not always appropriate to cite it alone for this reason.

The median: The median is the score at which 50 percent of the scores fall above it and 50 percent of the scores fall below it. With an odd number of scores, the median is the score in the middle when the scores are arranged from lowest to highest. With an even number of scores, the median is the average of the two middle scores. In the example of the set of response times cited above, the median would be the average of 6.8 and 8.5 or 7.6 seconds.

The median is frequently cited in addition to the mean when there are some scores that are very different from the rest. While these discrepant scores can have a large effect on the mean, they have very little effect on the median. When there are a few scores that are very different from the rest, then the median (as well as the mean) should be cited in the data analysis.

- **The mode: The mode is the most frequently occurring score.** In our example data set, the mode is 6.8 seconds, since it is the only score that occurs more than once. It is possible, especially with very small data sets, to have no mode. In very large data sets, it is possible to have multiple modes. While the mode is easy to identify, it is also very unstable (in the sense that it can change dramatically with the addition or deletion of a single score). For that reason, it is not as useful as the mean or median.

- **Measures of Variability**: A measure of central tendency, when presented in isolation, cannot fully describe the test results. In addition to the mean or median, we also need to know how close or disparate the scores were. In other words, how homogeneous were the scores as a group? To answer this question, we need to compute a measure of variability, also known as a **measure of dispersion**. The most commonly used measure of dispersion is the standard deviation. The standard deviation takes into account the number of scores and how close the scores are to the mean.

The standard deviation (abbreviated as "s" or "s.d.") is the square root of the variance. The variance (s^2) equals the squared deviations of each score from the mean divided by the total number of scores. One equation for computing the variance is as shown in Figure 10-1:

Figure 10-1. Equation for Computation of Standard Deviation

$$s = \sqrt{\frac{\Sigma(X-\bar{X})^2}{n-1}}$$

Where:

Σ is the summation sign
X represents each score
$\bar{X}$ equals the mean of the distribution, and
n equals the number of scores in the distribution

To compute the standard deviation in this way, we subtract each score from the mean, square each difference, add the squares of the differences, divide this sum by the number of scores (or the number of scores minus one), and take the square root of the result. Relatively small standard deviation values are indicative of a homogeneous set of scores. If all of the scores are the same, for example, the standard deviation equals zero. In our sample set of data used to compute the mean, the standard deviation equals 2.9 seconds.

Another use of the standard deviation is that it helps us to determine what scores, if any, we are justified in discarding from the data set. Studies in visual perception, for example, often use stimuli that are presented for very brief exposure durations (e.g., less than one-half of a second). In this case, a sneeze, lapse in attention, or other chance occurrence could produce an extraordinarily long response time. This data point would not be representative of the person's performance, nor would it be useful to the experimenter. What objective criterion could be used to decide whether this data point should be included in the analysis?

In the behavioral sciences, it is considered acceptable to discard any score that is at least three standard deviations above or below the mean. In our sample set of data, three times the standard deviation is 8.7. Since the mean is 8.6 seconds, scores of 17.3 (8.7 + 8.6) and above could legitimately be excluded from the data analysis, since they would be presumed to be due to conditions that were not related to the factors being tested (e.g., an equipment malfunction or an unusually long lapse in attention). (In this case, it is impossible to have a score three standard deviations below the mean, because it would indicate a negative response time.)

Percentiles: Percentiles provide a frame of reference for interpreting data. For example, a score of 500 (out of 1000) on an aptitude test doesn't sound impressive, if you think of it as 50% correct. However, the same score of 500 would be very impressive if it was at the 90th percentile. A score at the 90th percentile means that 90% of the rest of the group that took the test (or a comparison group, such as a group that took the test last year) scored below 500. In operations research, we are often concerned with "worst-case" scenarios. For example, software programs designed to give advice to pilots or controllers contain a time parameter for human performance. The traffic alert and collision avoidance system for pilots (TCAS) assumes that a certain amount of time (and no more than that time) will elapse between the time the advisory is given to the pilot and the time the pilot maneuvers the aircraft. Similarly, a conflict resolution algorithm for controllers would need to incorporate the time required for the controller to notice and understand the advisory and relay the required maneuver to the pilot. How long this will take will depend upon a variety of factors. For example, the time required for successful transmission of this message can be expected to vary as a function of frequency congestion, pilot response time, and other factors. Calculation of the mean and standard deviation of the time required would be useful in describing the data. However, the conflict resolution algorithm needs to be able to accommodate longer than average message transmission times. Since programs need to be able to handle less than optimum conditions, operations researchers often consider the 90th, 95th and 99th percentile. Ideally, every program would like to design to the 99th percentile to be able to handle the "worst-case" scenario. However, since there are always trade-offs (such as the number of valid resolutions that can be calculated), the 99th percentile is not always suitable. While percentiles provide a frame of reference for interpreting the data, how the percentiles should be used will depend upon the problem.

- **Correlation: Correlation is a commonly used descriptive statistic that describes the relation between two variables.** A correlation coefficient is reported as "r = x", where "x" equals some number between negative one and one. When two variables are unrelated (e.g., the number of rainy days per month in Kansas and the cost of airline fares), the correlation coefficient is near zero. A high positive "r" indicates that high values in one variable are associated with high values in the other variable. A high negative "r" indicates that high values in one variable are associated with low values in the other variable. A correlation of .7 or greater (or -.7 or less) is usually regarded as indicative of a strong relation between the two factors.

An important note about correlation is that even a very high correlation (e.g., r =.90) does not imply causality or a cause-effect relationship. A correlation coefficient merely indicates the degree to which two factors varied together, perhaps, as a result of a third variable that remains to be identified.

Another way in which the correlation coefficient is useful is that, when squared, it indicates the percentage of the variance that is accounted for by the manipulated factors. For example, with a correlation coefficient of .7, the factors that were examined in the analysis account for only 49 percent of the variance (i.e., the variability in the data). The other 51 percent is due to chance or things that were not controlled.

10.5.2 What are inferential statistics?

The statistics discussed above describe the test results and are, therefore, referred to as descriptive statistics. **Inferential statistics** are used to determine whether two or more samples of data are significantly different: for example, if performance on System A is significantly better or worse than performance on System B. (Just because the results are different in absolute terms does not mean that they are significantly different in statistical terms.)

The most commonly cited inferential statistics are the t-test and analysis of variance. These are *parametric* statistics and have a clearly defined underlying set of assumptions about the quality and nature of the data. If these assumptions are seriously violated, or the analysis is inappropriate for the experimental design, then *nonparametric* statistics should be used; otherwise, the conclusions based on the analysis may be invalid. (For a complete discussion of nonparametric statistics, see Siegel, 1956).

- **Student's *t* Ratio: Student's *t* ratio (commonly referred to as a t-test) compares two different groups of scores and determines the likelihood that the differences found between them are due to chance.** For example, *t*-tests would be appropriate when comparing the results of two groups of scores, whether it be the performance of the same group of controllers with System A and System B, or the performance of two groups of controllers: one using System A and the other using System B. When both sets of scores are taken from the same group of people, Student's *t* ratio for correlated samples is appropriate. When the scores of two different groups of people are examined, Student's *t* ratio for independent samples is appropriate. Both types of *t*-tests look at the differences between the two groups of scores with reference to the variability found within the groups. They provide an indication as to whether or not the difference between the two groups of scores is statistically significant. The results of a *t*-test are typically reported in the format shown in Figure 10-2.

Figure 10-2. Equation for Reporting of *t*-Test Results

$$t(df) = x, \ (p < p_o)$$

Where:

df = the degrees of freedom
x = the computed t-value
p_o = the probability value

For example, $t(20) = 3.29, (p < .01)$

Degrees of freedom (df) refers to the number of values that are free to vary, once we have placed certain restrictions on the data. In the case of a t-test for correlated samples, the number of degrees of freedom equals the number of subjects minus one. For independent samples, df equals the number of subjects in one group added to the number of subjects in the other group minus two. In both cases, as the number of subjects increases (and, hence, the number of df increases), a lower t-value is required to achieve significance.

The formulas for computing a t-ratio (and all of the statistics discussed in this chapter) can be found in *Experimental Statistics* (Natrella, 1966) and in most statistics textbooks.

- **Statistical Significance: The p value relates to the probability that this specific result was achieved by chance.** This is true not only for the t-values, but for all other statistics as well. A "$p < .01$" indicates that the probability that this result would be achieved by chance, and not due to the manipulated factors, is less than one in 100. When the results are significant at the .05 level, (i.e., $p \leq .05$), the chances of the results occurring by chance are 5 in 100, or less.

  Very often, this statistic is cited at the end of a statement of the results (e.g., "The number of errors was significantly higher in the high workload condition than in the low workload condition ($t(15) = 2.25$, $p < .05$.).") It can also be used to show that there were no statistically significant differences between two conditions (e.g., "The number of errors in the high workload conditions was comparable to the number of errors in the moderate low condition ($t(15) = 0.92$, $p >$. 10).") It cannot, however, be used to prove that there are no differences between the two groups, or that the two groups are the same.

For comparisons among more than two groups or more than two conditions in the same test, performing *t*-tests between all of the possible pairs would not be the best approach. A more appropriate test is **Analysis of Variance (ANOVA)**. Analysis of Variance is similar to a *t*-test in that it examines the differences between groups with respect to the differences within groups. In fact, when there are only two groups, an analysis of variance yields the same probability value as the *t*-ratio.

- **ANOVA:** ANOVA permits users to divide all of the potential information contained in the data into distinct, non-overlapping components. Each of these portions reflects a certain part of the experiment, such as the effect of an individual variable (i.e., a main effect), the interaction of any of the variables, or the differences due solely to individual subjects. Each main effect and interaction is reported separately, as shown in Figure 10-3.

Figure 10-3. Equation for Reporting Analysis of Variance Results

$$F(df,df) = x, (p < p_o)$$

Where:

df = the degrees of freedom
x = the computed F statistic
p_o = the probability value

For example, $F(2,24) = 7.78, (p < .01)$

For an ANOVA, the two reported degrees of freedom depend upon the number of subjects and the number of conditions or levels of effects.

As a hypothetical example, consider a simulation study of the operational effects of transmitting pilot-to-controller and controller-to-pilot communications via satellites. (For an actual study that is very similar to the hypothetical one described here, see Nadler, et al., 1992.) This method of transmission would impose a delay of approximately one-half second between the time the controller keys the microphone and the time the pilot was able to hear the beginning of the transmission. Pilot transmissions to controllers would be similarly affected.

One effect that satellite transmission might be expected to have on operations is to increase the number of blocked transmissions ("step-ons"), since the delay makes it possible for both controllers and pilots to key their microphones without realizing that there is an incoming transmission. (The number of pilot-pilot step-ons would not change, as the pilots would still be able to hear the beginning of the other pilots' transmissions without a delay.) Without this delay induced by satellites, blocked transmissions are due solely to two or more people (controller and pilot or two pilots) attempting to transmit at the same time and to stuck mikes. Since the probability of two individuals trying to transmit simultaneously is logically a function of frequency congestion, the number of transmissions on the frequency would be an important experimental variable.

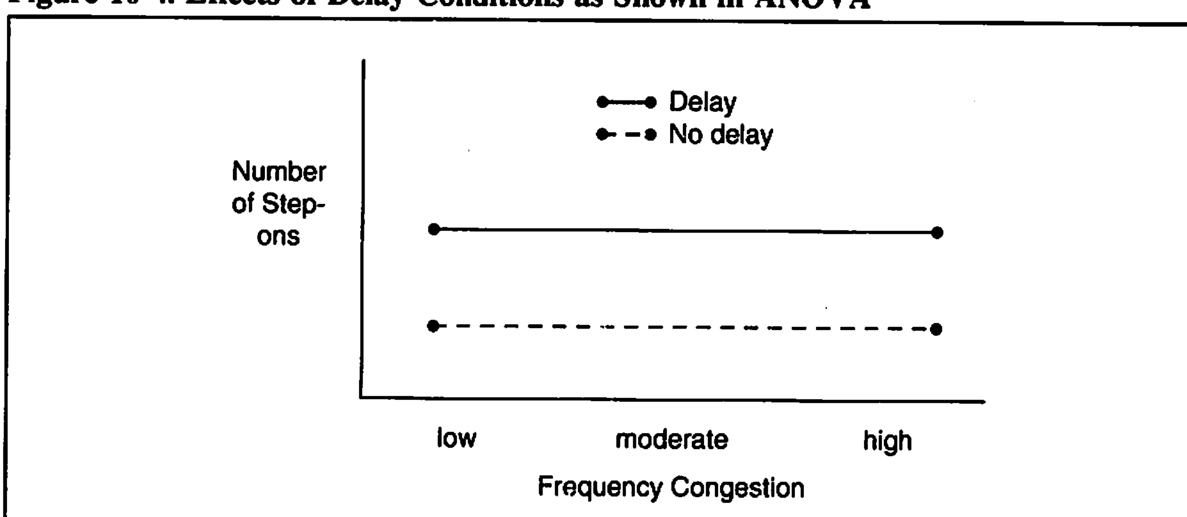
In this simulation study, two independent variables–the number of aircraft on the frequency and whether or not there is a communication delay–would be manipulated. Their effect on the number of step-ons (the dependent variable) would be measured. In this example, we have two levels of delay: 500 milliseconds. to simulate the satellite condition, and no delay to simulate the present system.

We are careful to ensure that the number of aircraft on the frequency generates different levels of frequency congestion. We categorize these levels of frequency congestion into "low," "moderate," and "high," based on data obtained from actual operations. Since we have two levels of delay and three levels of frequency congestion, this is referred to as a two by three experimental design.

Furthermore, we have a completely balanced design. This means that we have an equal number of hours of voice recordings in each combination of delay-frequency congestion conditions. We are statistically confident that we have an adequate number of different controllers and number of hours of data. We are also careful to keep all other conditions constant (which is always easier said than done).

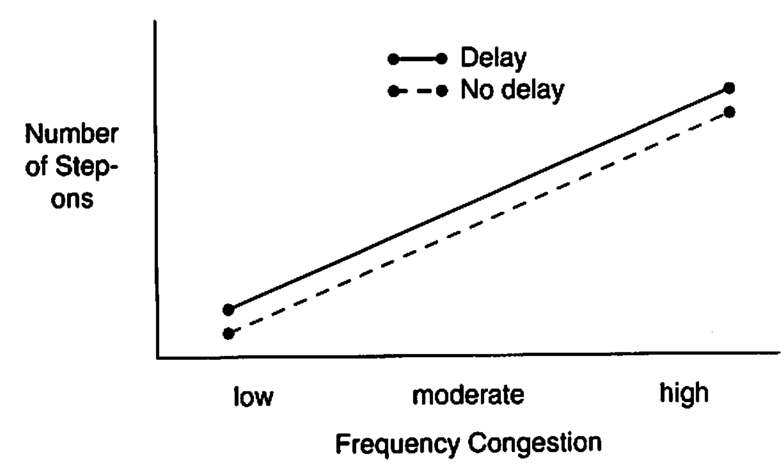
The three sources of variation in our ANOVA are the effects of delay, frequency congestion, and subjects (i.e., differences in the number of step-ons associated with different controllers). The results may show that the only significant effect is that of delay, meaning that the number of step-ons was significantly different for the two delay conditions. Graphically, this possibility might look like Figure 10-4:

Figure 10-4. Effects of Delay Conditions as Shown in ANOVA



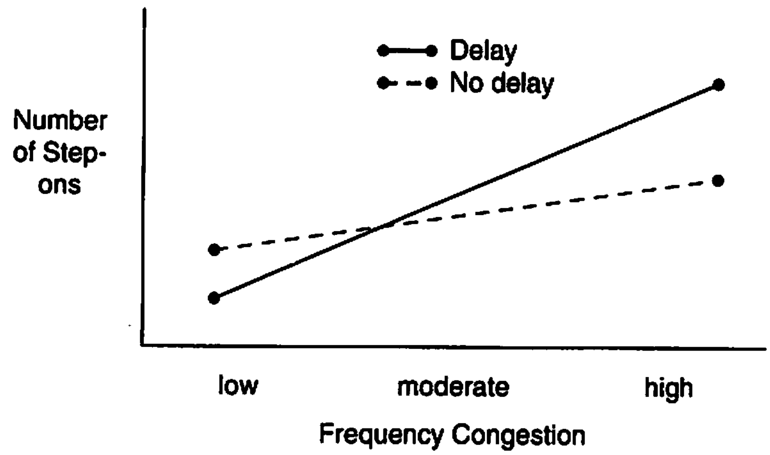
Another possible result is that the only significant effect was due to frequency congestion. This could mean that the number of step-ons increased with frequency congestion regardless of the delay condition. Graphically, this possibility might look like Figure 10-5.

Figure 10-5. Effects of Frequency Congestion as Shown in ANOVA



In addition to one, both, or neither of the effects of delay and frequency congestion being significant, a significant interaction may occur. A significant interaction would occur if, for example, there was no difference between the delay conditions at the lowest level of frequency congestion, but there was a significant difference at the highest level of frequency congestion. Graphically, this possibility might look like Figure 10-6.

Figure 10-6. Combined Effects of Delay Conditions and Frequency Congestion as Shown in ANOVA



These are only examples of the kind of results that may produce significant main effects or a significant interaction. There are many other possibilities. Of course, only a statistical analysis can determine whether differences portrayed on a graph are significant. Interpretation of test results is usually not simple, particularly with complex experimental designs. For this reason, human factors specialists with expertise in experimental design, but preferably statisticians, should be involved in the design of the research and the analysis of the results.

- **Regression Analysis**: A special case of analysis of variance that is often used is regression analysis. Regression analysis takes the data and fits it to a mathematical function. The function may be a straight line, a parabola, or any other function. The analysis provides an indication of how well the data fits that particular function.

One of the advantages of regression analysis is that it is very forgiving of empty cells in an experimental design (i.e., conditions in the design that do not have as many data points as the other conditions). For example, if we wanted to test how many mistakes controllers were likely to make with a certain data input system, but were most interested in the number of errors to be expected under conditions of high workload, then we might run a test with the majority of responses being in high workload conditions. Perhaps some controllers would only be tested in the high workload condition. Because of this asymmetry of data points in the high and moderate workload conditions, ANOVA would not be the most appropriate analysis; regression analysis, however, would still be appropriate.

Regression analysis also has some predictive value that analysis of variance does not. Regression analysis is often used to project from the data obtained in an experiment to situations that were not included in the test. In our hypothetical example, regression analysis would be appropriate if communication delays of 0 milliseconds, 250 milliseconds, and 500 milliseconds were tested and we wanted an estimate of the number of step-ons that could be expected at delays of 300 or 600 milliseconds.

When using regression analysis in this way, it is important to remember three points. First, the projection can only be as good as the fit of the data to the mathematical function. Second, all other things being equal, an estimate between two data points inspires more confidence than a projection beyond (above or below) the values included in the test. Third, **confidence in the projection decreases as the distance between the hypothetical or projected point and the value that was included in the test increases.**

10.5.3 What are the differences between statistically and operationally significant results?

A final note about data analysis concerns the differences between statistically significant and operationally significant results. Most statisticians only seriously consider results that are statistically significant at the .05 level or better. This enables investigators to be reasonably certain that their findings were not due to chance and thus, are likely to be repeatable. A statistically significant difference may, however, be very small as long as it is consistent. This may or may not be operationally useful. This difference between statistical significance and operational significance is often overlooked. A difference in response times of half of a second may be statistically significant, but may not be operationally important, depending upon the task.

On the other hand, when the experimental focus is actual operations, results that are not statistically significant at the .05 level may still be important. For example, if the focus of the experiment is serious operator errors that could significantly affect flight safety, then we may choose to conservatively consider results that are statistically significant only at the .1 level. The standard criteria for acceptance of statistical significance at the .05 level should not be used to ignore potentially meaningful findings. It may also be the case that statistically significant results would be attainable with a more powerful test or change in research design (e.g., by utilizing better experimental controls or by increasing the number of subjects). The decision as to what level of significance is to be used should depend on the nature of the question that the test is designed to answer.

10.5.4 An Example

Human factors specialists, controllers and other specialists (e.g., training specialists, statisticians, etc., as needed) must work together to properly evaluate a new system design or an important modification to an existing system. Only a team of these specialists can ensure that the necessary questions are properly addressed with a well-designed test and data analysis.

As an example, suppose we wanted to design and evaluate a modification to the controller's display that would inform the controller when a pilot is getting a resolution advisory (RA) from the Traffic Alert and Collision Avoidance System (TCAS). TCAS II RAs may direct a pilot to climb, descend, maintain an existing vertical speed, or avoid a certain vertical speed. (TCAS IV will offer resolutions in both the horizontal and vertical planes.)

I. Display Design

The following questions should be asked during the design phase (or later if they were not already addressed):

What should be displayed?

- Should every RA be displayed or should it only be displayed if the aircraft maneuvers?
- Should the displayed maneuvers include preventive RAs (e.g., "Limit vertical speed"), or should they be limited to corrective RAs (e.g., "Climb").
- Should the direction of the RA (i.e., climb or descent) be displayed?
- Should only the initial RA be displayed or should "increases" (i.e., to increase vertical speed) and "reversals" (e.g., a change from a "climb" RA to a "descend" RA) also be displayed?

How should it be displayed?

- What should the alert look like?
- Should the indication be attached to the data block, be displayed on the tabular list, or be located elsewhere?
- Should the maneuver be displayed in graphic or text format?
- Should/could "TCAS↓" be displayed on the data block?

  - Should this text blink?

  - Should the entire data block blink?

- On displays where use of color is an option, what color should the alert be?

- How should changes in the RA be displayed to the controller?

All of these questions have technical considerations that are inseparable from human factors considerations. It may be the case that the technology available is not capable of supporting the preferences of controllers and human factors specialists, both in terms of what is displayed or how it is displayed. For example, the preference may be to only display the RA if the aircraft begins to maneuver. However, if the source of the data that determines when the aircraft begins to maneuver is limited to radar returns (rather than from the flight data computer via data link), then the information may arrive too late to be useful.

The answers to these questions depend, in part, on operational decisions that need to be made about the importance, urgency and usefulness of such an alert. For example, if the alert is determined to be something that requires the controller's immediate attention, then it should be displayed differently than information that can be reviewed at the controller's discretion.

How are controllers expected to use the information?

Will controllers use the information to maneuver or advise other aircraft? If the controllers intend to use the information for immediate strategic planning and aircraft separation, then the information needs to be current and available to the controller as long as it is relevant. If this is not possible or not advisable, then the controllers need to be aware of the limitations of the information available to them.

Do controllers refer to the display only after they notice that an aircraft has deviated from an assigned altitude? Do controllers use the information to "wait and see" what the aircraft will do? These questions will need to be addressed by operations personnel so that the answers can help to determine how the information should be displayed.

How efficiently should the alert attract attention?

There is a trade-off between the efficiency with which an alert will attract your attention and the potential for the alert to become distracting and annoying. An alert designed to automatically attract your attention in every instance can be distracting when you are trying to concentrate on something else. All such alerts have the potential for being a nuisance if they go off for no apparent reason or if they provide no additional information (e.g., the pilot already informed you that they were climbing in response to an RA). Considerations for determining how efficient the alert should attract attention include: how often the alert will be seen, under what circumstances it will most likely be presented (e.g., high or low workload), how often will the alert give the controller useful information (versus how often the alert will be a nuisance), and whether an immediate response from the controller is deemed necessary.

II. Testing and Evaluation

The operational and human factors testing and evaluation should be designed to answer the following questions:

Is the display in a format that is easily read and understood?

- What errors are controllers likely to make in reading and interpreting the display and how probable are these errors?

Does the display provide useful information or is it a nuisance?

- What percentage of the time, and under what circumstances, is the alert useful?

- When the alert is useful, how do the controllers use the information? (Was the message used as intended? Were other uses identified? If so, should changes in the display be considered to accommodate these different uses?)

Does the display efficiently attract the right amount of attention?

- How much time is required for a controller to notice the display? If too much time is required, then the display will need to be changed so that it is more easily noticed. If it attracts attention quicker than is deemed necessary, then the display may be made less salient, so that it is not as distracting.

The intent of any display should be to give the controller only the information that is needed, when it is needed, and in the format that it is needed. Evaluations, from the design phase through operational testing, must help to ensure that the system performs as intended by identifying and avoiding potential design flaws and other human factors "traps".

In the case of this example, these questions would best be addressed with simulation studies (such as those conducted at the simulation facility at the FAA Technical Center in Atlantic City New Jersey). Controllers would become familiar with the new display and then begin the test by performing their normal duties. Scenarios of relatively high workload would be presented to the controllers to ensure that the they are not able to devote more attention to the display than they could in actual operations.

Although more than one testing procedure would be appropriate in this case, one test procedure that could be used is to ask the controllers to read the TCAS alert into the microphone when it appears. This would record what the controller thought the alert said and the time that they noticed the alert. To assess whether the display efficiently attracted attention, the time between the onset of the alert and the time the controller noticed it would be measured.

The purpose of this measure of controller response time is not to determine how quickly controllers can respond, but rather, its purpose is to determine how much time should be reasonably allocated for a controller to respond to the alert. This is often important information for system designers and software developers. It can help to determine which information should be displayed and how it should be displayed.

Errors in reading the display would also be recorded. Again, the purpose of this measure is solely to test the adequacy of the display, not to test controllers. Information on errors made in reading the display would be gathered to establish a likely error rate and to identify patterns of errors that could point to necessary changes in the display. Finally, a questionnaire could be provided to the controllers after they have participated in the test. The questionnaire and interviews with the controllers could be used to assess whether the controllers thought that the information provided by the display was useful and how the controllers used the information. Their opinions on the display format could also be valuable. Any significant design changes that resulted from the initial test would then need to be evaluated in a new test.

This iterative approach to display design and evaluation takes lessons learned early in the evaluation process (or from evaluation of prototype displays) and incorporates them into the development of the final display design. Any remedial design changes need to be tested to ensure that they do not induce other problems. Only through thorough testing can confidence in all aspects of the display be assured.

10.6 REFERENCES AND SUGGESTIONS FOR FURTHER READING

Berson, B. L., Po-Chedley, D. A., Boucek, G. P., Hanson, D. C., Leffler, M. F., and Wasson, R. L. (1981). Aircraft alerting systems standardization study. In *Volume II: Aircraft alerting system design guidelines* (Report No. DOT/FAA/RD-81/II). Washington, DC: Department of Transportation/Federal Aviation Administration.

Boff, K. R., and Lincoln, J. E. (Eds.) (1988). Engineering data compendium: Human perception and performance. In *Volume III,* (p. 1862). Wright-Patterson Air Force Base, OH: Harry G. Armstrong Aerospace Medical Research Laboratory.

Boucek, G. P., Erickson, J. B., Berson, B. L., Hanson, D. C., Leffler, M. F., and Po-Chedley, D. A. (1980). *Aircraft alerting systems standardization study* (Phase I final report, Report No. FAA/RD-80-68). Washington, DC: Department of Transportation/Federal Aviation Administration.

Boucek, G. P., Pfaff, T. A., White, R. W., and Smith, W. D. (1985). *Traffic alert and collision avoidance system - operational simulation* (Report No. DOT/FAA/PM-85/10). Washington, DC: Department of Transportation/Federal Aviation Administration.

Boucek, G. P., Po-Chedley, D. A., Berson, B. L., Hanson, D. C., Leffler, M. F., and White, R. W. (1981). Aircraft alerting systems standardization study. In *Volume I: Candidate system validation and time-critical display evaluation* (Report No. FAA-RD-81-381). Washington, DC: Department of Transportation/Federal Aviation Administration.

Boucek, G. P., White, R. W., Smith, W. D., and Kraus, J. M. (1982). *Traffic alert and collision avoidance system - Developmental simulation* (Report No. DOT/FAA/RD-82/49). Washington, DC: Department of Transportation/Federal Aviation Administration.

Cardosi, K., and Boole, P. (1991). *Analysis of pilot response time to time-critical air traffic control calls* (Report No. DOT/FAA/RD-91/20). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Cardosi, K., Warner, M., Boole, P., Mengert, P., and DiSario, R. (1992). *Controller response to conflict resolution advisory prototype* (Report No. DOT/FAA/NA-92/1). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Cardosi, K., Burki-Cohen, J., Boole, P., Mengert, P., and DiSario, R. (1992). *Controller response to conflict resolution advisory* (Report No. DOT/FAA/NA-92/2). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Cardosi, K., and Huntley, M. S. (1988). *Cockpit and cabin crew coordination* (Report No. DOT/FAA/FS-88/1). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Foushee, C., Lauber, J., Baetge, M., and Acomb, D. (1986). *Crew factors in flight operations III; The operational significance of exposure to short-haul air transport operations* (NASA Technical Memorandum 88322). Moffett Field, CA: National Aviation and Space Administration.

Hopkin, D., The measurement of the air traffic controller. (1980). *Human Factors, 22*, 547-560.

Kidder, L. H. (1981). *Selltiz, Wrightsman, and Cook's research methods in social relations* (4th ed.) (p. 483). New York: Holt, Rinehart and Winston.

Nadler, E.D., DiSario, R., Mengert, P., and Sussman, E.D. (1990). *A simulation study of the effects of communication delay on air traffic control* (Report No. DOT/FAA/CT-90/6). Washington, DC: Department of Transportation/Federal Aviation Administration/Volpe National Transportation Systems Center.

Natrella, M. G. (1966). *Experimental statistics* (National Bureau of Standards Handbook 91). Washington, DC: National Bureau of Standards.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

10.7 ENDNOTES

1. Since there is no licensing procedure for human factors specialists, they can only be judged by their credentials and previous experience. A successful (i.e., helpful) human factors specialist is likely to have an advanced degree (in an area of experimental or engineering psychology, human factors, or a related field) and experience in a specific area (e.g., ATC).

APPENDIX 10A. CHECKLIST ITEMS

Test Methods

1. The proposed method of testing is appropriate (10.4.1).

2. If field observations are used, they were obtained under conditions that are representative of the full scope of actual operations (10.4.1).

3. If questionnaires are used:

    The questions are simple and direct (10.4.1).

    The questionnaire is unbiased (10.4.1).

    The questionnaire will be administered as soon as possible after the task (10.4.1).

    The questionnaire will be given to a representative set of users (10.4.2).

4. The test variables are operationally defined in the test plan (10.4.2).

5. The test participants (subjects) are representative of the user population (10.4.2).

6. The test design controls for subject bias (10.4.2).

7. The conditions (scenarios) included in the test plan are representative of actual operations (10.4.2).

8. If there is more than one condition, the order of the conditions is counterbalanced or randomized (10.4.2).

Analysis of Test Results

1. The proposed methods of data analysis are appropriate for the test design (10.5).

2. The measure of central tendency (i.e., mean, median, or mode) chosen is the most appropriate for the test results (10.5.1).

3. A measure of variability is included in the analysis (10.5.1).

4. If a measure of correlation is used or proposed, it is appropriate for the data (10.5.1).

5. If a t-test is used or proposed, it is appropriate (for a comparison between two groups) (10.5.2).

6. If an Analysis of Variance (ANOVA) is used or proposed, it is appropriate (for a comparison between more than two groups) (10.5.2).

7. In any report of test results, the level of statistical significance is cited appropriately (10.5.2).

8. If a regression analysis is used, all of points that are predictions and projections are close to the actual data (10.5.2).

9. In the interpretation of test results, the distinction between statistically and operationally significant results is clear (10.5.3).

APPENDIX A-HUMAN FACTORS CONSIDERATIONS FOR COLOR DISPLAYS FOR ATC

By Daniel J. Hannon, Ph.D.

1. Introduction

The visual display of information for the purpose of air traffic control (ATC) requires the use of multifunction electronic display equipment. Recently, the Federal Aviation Administration (FAA) has considered the use of raster-scan, shadow mask, cathode ray tubes (CRTs) for the display of color coded air traffic information. Although the use of color-capable displays has been considered by the FAA since the mid-1970s, controllers currently use monochromatic displays for most of their duties. Similarities among the different ATC settings (e.g., TRACON and Tower) has prompted the request for uniform display software and coding conventions to be used in all ATC applications on these display screens. However, differences between the settings in terms of operator location, display screen location, viewing angles of the observers to the display screens, ambient lighting, and the presence or absence of sunlight glare make it difficult to apply one standard display configuration to all ATC locations. The need to consider commercial-off-the-shelf (COTS) equipment further limits the possibilities for creating a single ATC display system. Insuring the adequate display of information for use by controllers in all ATC settings necessitates the consideration of human factors guidelines in regard to the physical characteristics of the display, the display output, and the display environment, as well as to the layout of the information on the displays.

This document provides an evaluation and discussion of the environmental factors and display characteristics that influence the perception of color from displays by human observers. Although the information is organized into specific sections, the sections are not completely self contained chapters. To some extent, the information is presented in a sequential order, with succeeding sections building on previous sections. Section 2 briefly traces the history of ideas regarding the use of color on ATC displays. Section 3 focuses on current guidelines for the use of color on displays of air-traffic information. Section 4 provides an evaluation of the impact of high ambient lighting, particularly sunlight, on displays of color coded information. In section 5, techniques for improving the sunlight readability of displays are presented. Section 6 considers the impact of the viewing environment on the perception of color coded information. Section 7 briefly discusses the visual performance characteristics of current flat panel devices. In section 8, the use of window shades for reducing the effect of sunlight on color displays in the ATC tower is considered.

The goal of this report is primarily to identify issues that should be considered in the selection and implementation of color capable electronic displays in ATC settings. Specific recommendations have deliberately been avoided, to the extent possible, in order to foster further consideration of the issues.

2. Background Information on the Use of Color Coded Information on Air Traffic Displays

The FAA has considered the use of color-coding on displays of air traffic information in the different ATC environments (e.g., TRACON and Tower) for many years. Improvements in display technology since the mid-1970s have resulted in continued interest in the incorporation of color displays in ATC settings. Despite the interest in the use of color for ATC displays, and the effort that has been devoted to developing color capable ATC display systems, there is a striking lack of empirical data to support the decisions that must be made in order to achieve safe and effective color displays. Although human factors guidelines have been generated for color ATC displays, these guidelines have been distilled primarily from limited studies of human color perception and from expert opinion (Campbell, White, and Hamilton, 1990). In this section, relevant literature on the use of color in ATC is summarized in a chronological order in an attempt to separate empirically derived information from information that is primarily based on conjecture. This analysis will provide a foundation for the interpretation of information contained in subsequent sections as well as point to areas that require further investigation.

Early work on the introduction of color to the process of air traffic control concluded that color coding may be beneficial to controllers (Connolly, Spanier, and Champion, 1975). The color-capable, CRTs that were considered by these authors employed a red-green penetration phosphor. With this type of phosphor, electron beams in the CRT penetrated the phosphor on the screen to different depths. The depth of penetration created differential excitation of the red and green layers resulting in emitted light that appeared red, orange, yellow or green. Even with this limited color gamut, researchers concluded that the color coding of air traffic information was helpful in retrieving information from overlapping data blocks. The researchers also found, however, that performance on the task of maintaining aircraft separation was not enhanced by the application of color. These authors also noted that the unreliability of color displays was a major obstacle to the incorporation of color displays in ATC settings.

Kinney and Culhane (1978) conducted a survey of the literature also on the topic the application of color to ATC displays. These authors attempted to identify: 1) whether color helps or hinders performance, 2) how color should be used, 3) which colors should be used, and 4) what additional information was needed before color coding could be introduced onto ATC displays. Unfortunately, their main conclusion was that very little was known about how air traffic controllers performed with color displays. Due to the paucity of information, these authors were unable to answer the first three of their questions. However, they were able to glean some insights from literature on human color perception and they used this information to develop some recommendations for the application of color to ATC displays.

In a later study, Kinney and DeCicco (1982) and Aschenbach, Kopala, and Douglass (1982) jointly attempted to evaluate the use of color on ATC displays in an operational setting. Color capable displays were introduced into an actual ATC setting and survey data was collected on controllers' evaluations of the new displays. The displays could be configured for either all green information, or in a four color mode including green, yellow, orange and red symbology. Controllers rated the color displays as no better, and sometimes worse than existing monochromatic displays. The subjective nature of the data collected limited the conclusions to qualitative statements about the displays employed in the study. Numerous technical difficulties also were noted. These authors were reluctant to draw conclusions beyond their specific study.

Concurrent interest in the application of color displays to ATC environments by European agencies led to further developments in the 1980s. Narborough-Hall (1985) advocated the use of color in ATC displays and provided several guidelines for color usage as well as a thoughtful analysis of potential advantages and disadvantages to color coding ATC display information. In addition to the studies noted, this author also summarized findings from efforts by European agencies to evaluate color displays in ATC settings. Again, little in the way of empirical findings from direct studies of color in ATC displays were offered.

A somewhat more recent study on color use in ATC was conducted for the Canadian government by Campbell, White, and Hamilton (1990). These authors also reviewed relevant color perception literature and related human factors guidelines in order to develop recommendations for the application of color to ATC displays. They also noted the lack of solid empirical evidence to support the use of color on ATC displays. Using a somewhat unique approach, they reported the contents of conversations they had with several people in the aviation research community, including people in the US, Canada, and Europe. Although there were a few statements made about pending research, a number of the respondents clearly indicated that they were pursuing development of color displays for ATC, or they were pursuing the creation of guidelines for the application of color to ATC displays, based on existing knowledge of human color perception, and common sense. This included existing computer-human interaction (CHI) guidelines used by the FAA.

In summary, there are many more guidelines and recommendations for the use of color in ATC displays than there are empirical reports to support them. The sequence of studies reported here indicates that while there has been much speculation about the use of color, there is a paucity of behavioral data that is needed for the careful application of color to ATC displays. The prevailing approach to the incorporation of color into ATC displays is to rely on extrapolations from what is known about color perception and the task of the air traffic controller, as well as to consult existing human factors guidelines (e.g., military standards) on the use of color. The guidelines that do exist have been derived from a thoughtful consideration of the situation. However, specific applications of whether specific color coded information on ATC displays will be a benefit or a burden to air traffic controllers remains unknown.

The remainder of this document attempts to provide some guidance for the selection of color display equipment for use in ATC environments from the perspective of the interaction of the human observer with the display device. Due to the general lack of direct empirical data from which to derive conclusions, the information that is provided is also based to some extent on conjecture. The inclusion of specific numeric data and other information should not be construed as an indication that sound empirical data are no longer needed. In several instances, issues are raised for which there are no simple answers. These are clear indications that some direct study is needed.

3. Current Guidelines for the use of Color on ATC Displays

In this section, the current guidelines for using color on ATC displays are considered from the perspective of the display characteristics that will be necessary to depict the colors and the impact on the human operator. The first consideration is the source of the guidelines and the foundation for the information contained in them. Second, the requirements for color on ATC displays will be briefly summarized. Next, the specific colors called for in the guidelines will be considered with respect to their physical properties on the screen and how they are likely to be perceived by the observer. Two different color specification schemes will be considered. Issues related to the choice between different color schemes will also be discussed. Finally, methods for creating different shades of a color are presented.

3.1 Sources of Information on Color Specification

There are two documents that the FAA has used to provide guidance on the selection of colors for ATC displays - the specifications for the Advanced Automation System (AAS) and CDRL EN09. These documents will be referred to here as AAS and ENO9. The AAS document provides color specification for TRACON and Tower displays under the AAS system. The most recent document available was updated in 1990. In this document, color displays were intended for use in TRACON, and monochromatic displays were prescribed for the tower. The ENO9 document establishes CHI guidelines for a system that has superseded the AAS project. The most recent version of this document was created in 1993. Color specification is also limited to the TRACON environment. The information in both of these documents is often referenced by the FAA and industry display engineers with regard to color specification and ATC displays.

The color specifications provided in AAS appear to have been tailored to the specific CRT that was being procured for the system. Color selection utilized each of the red, green and blue phosphors in isolation, their binary combinations for producing yellow, cyan, and magenta, and the combination of all three to produce white. Although this is a functionally efficient scheme because it maximizes the color differences, it is not clear that the choice of colors is optimal for the task of ATC. The ENO9 color scheme is also based on the same CRT as the AAS system. However, the specification for each color required specific levels of each of the three phosphors. The basis for this selection of colors is not clear in the ENO9 document. Unfortunately, the original authors of this document are no longer employed with the organization that produced it, so the purpose of this color selection scheme remains uncertain. Once again, it is not clear if this color scheme has been optimized for the task of ATC. Since it cannot be established, based on preliminary information, that either the AAS or the ENO9 color scheme is superior to the other, both schemes are included in the analyses below.

3.2 Color Requirements for ATC

The AAS and ENO9 color schemes are based on a core set of color requirements for ATC displays. Seven different, simultaneously presented colors are required: red, yellow, green, cyan, blue, magenta, and white. Maximum and minimum luminance levels are also specified. The color of the information presented to the controller is based on the following code:

Green	- general color used for most information on the display.
White	- used for new information or information that has changed.
Yellow	- used to indicate a warning or a situation that requires attention.
Red	- used for emergency situations that require immediate attention.
Blue	- used for graphic weather information and readout fields.
Cyan & Magenta	- used for operational information such as aircraft direction.

This particular color code serves two basic functions. First, color assignments to different types of information, such as blue for weather, and green for aircraft, allow for the perceptual grouping of similar information that is spread out over the display screen. Second, the white, yellow and red color assignments provide information on the status or condition of the displayed item.

The two different functions of the color code require different levels of color specification. A requirement for color assignments when the color code is used for grouping similar information is that the colors look different to the observer. Color discrimination is more important than absolute color identification. The specific choice of blue, in this case, is not as critical as the requirement that it look different from the green that is used. Color codes that provide an alerting signal, or status information, however, require that the specific color be identified. For example, information displayed in yellow must look distinctly yellow if it is to be interpreted quickly as a warning to the controller. A yellow target that is pale may be mistaken as white and not command the attention of the observer. In general, **in instances**

where absolute color identification is required, the number of different colors on a display should be minimized (Merrifield & Silverstein, 1986).

3.3 The ENO9 and AAS Color Schemes

In this section, the color schemes detailed in ENO9 and AAS are considered from the perspective of display requirements and the likely appearance of each color to a human observer. Information is first provided on the color capabilities of the SONY DDM-2802C color CRT that was used in the creation of the color sets. This will allow more readily for the interpretation and comparison of the two different color schemes. The specifics of the ENO9 scheme will be presented next, followed by the AAS scheme.

3.3.1 SONY DDM-2802C CRT

The SONY monitor is a 20" by 20", shadow mask, CRT, with 2048 addressable pixels in both horizontal and vertical dimensions. The pixel pitch (i.e., distance between RGB groups) is listed as .31 mm which results in approximately 1600 resolvable pixels in both dimensions. The chromaticity coordinates of the three phosphors are listed in Table 1 and plotted in Figure 1.

Table 1
Chromaticity Coordinates for SONY CRT

Color	x	y
red	0.625 ± 0.03	0.340 ± 0.02
green	0.280 ± 0.02	0.595 ± 0.03
blue	0.155 ± 0.03	0.070 ± 0.015
white	0.271 ± 0.03	0.286 ± 0.03

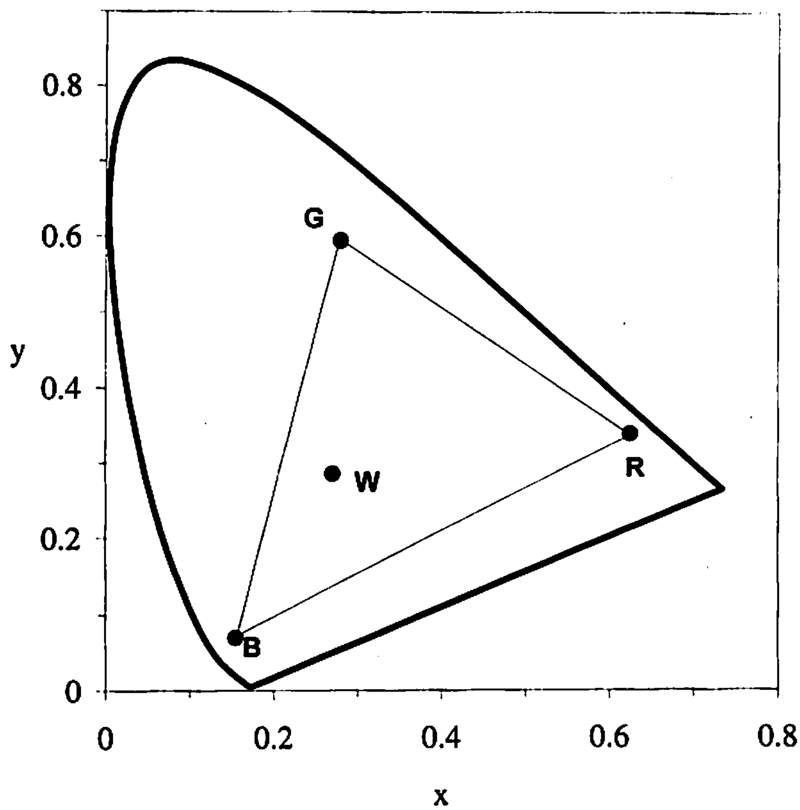


Figure 1. Color gamut for SONY DDM-2802C monitor on the 1931 CIE chromaticity diagram. Filled symbols represent chromaticity coordinates of the red (R), green (G) and blue (B) phosphors. White (W) is also plotted.

The plot in Figure 1 depicts the CIE chromaticity coordinates of the SONY phosphors on the 1931 CIE chromaticity diagram. The CIE chromaticity measurement system converts spectral measurements, through a system of linear equations, into a unitless coordinate system. The x and y coordinates represent a specific color in an abstract color space. The closed figure that is plotted with the heavy line represents the locus of all visible colors. Monochromatic lights plot along the outer edge. White or achromatic appearing lights tend to plot near the center of the closed figure. Broadband colors plot in-between the center of the figure and the outer edge. The relative distance of a point between the center and the outer edge is an indication of the purity or saturation of the color.

Most CRT monitors rely on the admixture at the eye of three primary colors to produce the gamut of colors seen on the screen. Each phosphor is a relatively broadband light source and, as such, can be plotted on the chromaticity diagram. The red, green and blue phosphors (indicated with R, G, & B) form a triangle in this color space. All of the colors that can be produced by the SONY CRT lie within the boundaries of the triangle plotted in Figure 1. The point labeled W represents the white point indicated in the SONY specifications. The location of this point nearer to the blue (B) value indicates that the appearance of this white may be somewhat bluish.

3.3.2 The ENO9 Color Set

Specific CIE chromaticity coordinates for colors were not provided in the ENO9 document for the seven chosen colors. However, RGB coordinates were specified. The difference between the RGB specification system and the CIE system is that the CIE system is independent of the type of display device chosen. The RGB coordinates are device-dependent in that specific values for the red, green and blue electron beams in the CRT are dictated. Two different CRTs can be set up to use the same RGB levels and produce very different colors. However, if CIE coordinates are matched between colors displayed on different CRTs, then the colors will look the same. The only way to work with the ENO9 color specification scheme was to evaluate the colors on the SONY CRT. In order to allow for comparison between the ENO9 color scheme and the AAS color scheme (as well as others), the ENO9 colors were measured on a CRT with a colorimeter that provided CIE color coordinates. The TRACON laboratory facility at the MITRE Corporation in Bedford, MA, provided the setting for measuring the colors. The data reported below on the ENO9 color set are based on color measurements taken from a single CRT at that facility. Although care was taken in data collection, it is always the case that no single measurement of a displayed color can be regarded as infallible. Independent verification of the CIE coordinates is recommended before drawing conclusions with regard to the ENO9 color scheme.

Table 2 provides the x and y chromaticity coordinates for the ENO9 color scheme as well as the luminance of each of the displayed colors. The chromaticity coordinates are plotted in Figure 2. The seven filled circles labeled green, red, white, blue, cyan, magenta, and yellow are the measured chromaticity coordinates based on the RGB values listed in the ENO9 document. With the exception of Green, all of the colors are located relatively close to the plot for White. This indicates that the appearance of all of them will be somewhat subdued. Both Blue and Cyan plot near the line connecting the G and B phosphors of the SONY monitor, indicating that they are comprised primarily of these two phosphors. Magenta, Red and Yellow are all plotted within the boundary of the triangle indicating that all three phosphors influence the appearance of these colors. Green coding uses only the G phosphor.

Table 2
Chromaticity Coordinates and Luminances for the ENO9 Color Set

Color	x	y	luminance (fL)
blue	0.195	0.204	19.5
cyan	0.206	0.271	31
green	0.279	0.595	11.9
magenta	0.277	0.266	22.4
red	0.392	0.299	11.8
white	0.272	0.311	34.5
yellow	0.356	0.470	29.5

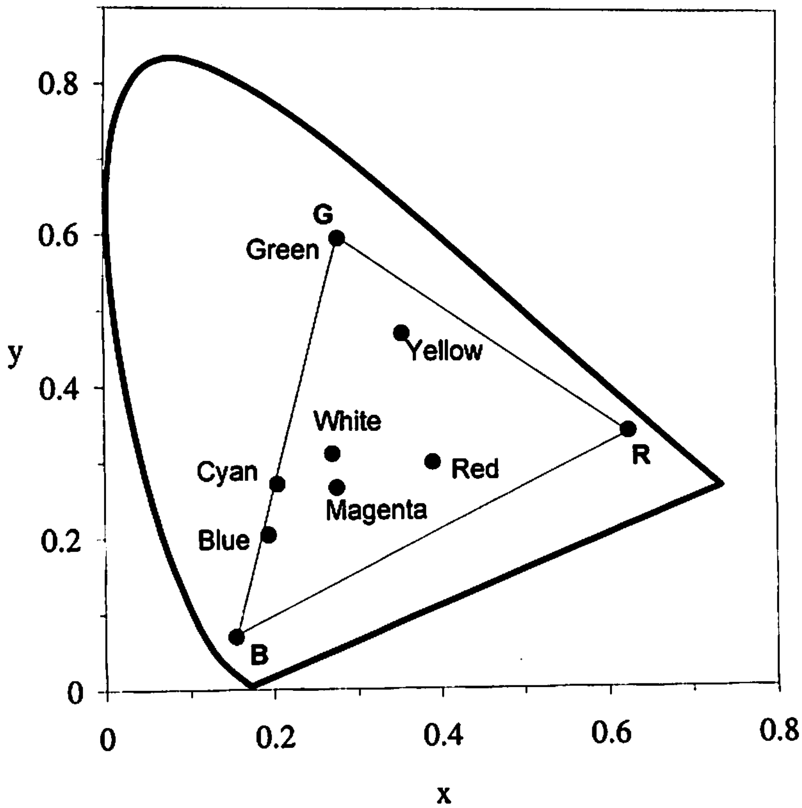


Figure 2. Chromaticity Coordinates of ENO9 Specified Colors as Measured on a SONY Monitor

3.3.3 The AAS Color Set

The AAS color specification did use the CIE system. The chromaticity coordinates are repeated here in Table 3. Figure 3 plots these values on the chromaticity diagram. Filled triangles indicate the colors in the color scheme. The coordinates of the SONY phosphors are indicated with filled circles, and the color gamut triangle is shown as well. The Green, Red and Blue colors utilize the G, R, and B phosphors directly and therefore are collocated at these points. The White point is near to the center of the triangle indicating a neutral appearance. Magenta, Cyan and Yellow plot on the triangle representing the color gamut of the SONY monitor. These colors also plot at the midpoint of each segment of the triangle. These two facts indicate that the colors are generated with only two of the screen phosphors, and that a somewhat equal balance between the two was used.

Table 3
Chromaticity Coordinates and Luminances for the AAS Color Set

Color	x	y	luminance (fL)
blue	0.155	0.070	0.50
cyan	0.217	0.332	3.13
green	0.280	0.595	2.80
magenta	0.390	0.205	0.905
red	0.625	0.340	0.75
white	0.350	0.350	2.40
yellow	0.453	0.467	2.05

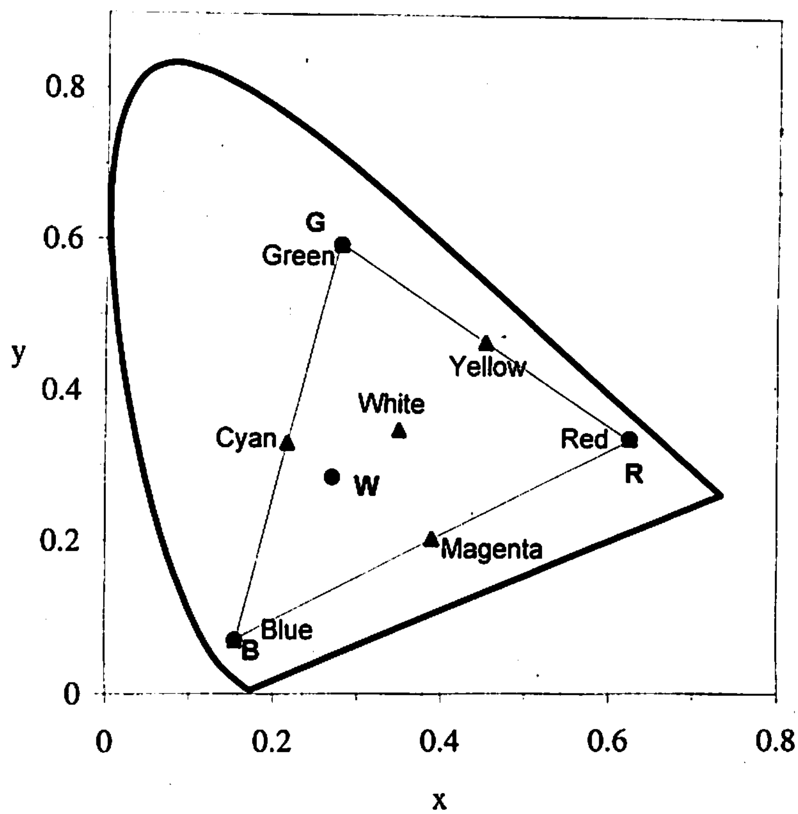


Figure 3. Chromaticity Coordinates of AAS Color Set

3.4 A Comparison of the ENO9 and AAS Color Schemes

A comparison between the ENO9 and AAS color schemes is possible at this point since both sets of colors are now plotted in a common coordinate system. Both color sets were derived for use in the TRACON environment. However, a comparison between Figures 2 and 3 indicates that the colors chosen were quite different. Overall, the ENO9 colors are less saturated or vivid in comparison to the AAS colors. The AAS system utilizes single and binary combinations of the three screen phosphors to create the most saturated colors possible on the monitor. The ENO9 system creates colors using greater mixtures of all three phosphors resulting in a paler appearance. However, the luminance values presented in Tables 2 and 3 indicate that the ENO9 colors will have a brighter appearance than the AAS colors.

How saturated should display colors be?

Narborough-Hall (1985) suggested that desaturated colors would be best for ATC environments and noted potential problems with using a more saturated color set. Saturated colors can create strong chromatic afterimages, and saturated blue and red targets can create difficulties for accommodation and induce chromostereopsis, a condition in which blue and red targets appear at different depths. (For an illustration of this phenomenon, see Figure 3-11 in Chapter 3.) Alternatively, Aschenbach, Kopala, and Douglass (1982) noted that controllers complained about the lack of distinctness of the colors used in their study. The alerting feature of a color is potentially diminished if it is desaturated. Further, desaturated colors tend to have a greater level of similarity in appearance than saturated colors. Small differences in color appearance are easily overlooked under time-critical situations. This implies that, under some conditions, maximal saturation should be employed. For example, if a display is to be used in an ATC tower during the day, maximal saturation would be critical, so that the colors are distinct. At night, however, when the ambient illumination is much less, the same display (at the same settings) would be difficult to view and could induce chromostereopsis. Colors to be used on an ATC display must be carefully chosen according to the purpose of the display and the viewing environment.

Color by color comparisons between the two schemes also reveal important differences between ENO9 and AAS:

- The choice of Red was very different between the two schemes. The ENO9 Red was quite desaturated in comparison to the AAS Red. However, both appeared fairly distinct from the other colors within each respective scheme.

- The choice of Yellow in ENO9 was less saturated than the AAS Yellow, and it was closer in appearance to Green than the AAS Yellow.

- Cyan and Blue within the ENO9 set were closely spaced indicating that they had a similar appearance. Cyan was also close to white indicating that although it plots near the triangle border, Cyan was also a desaturated color. This condition results from the particular choice of white in the ENO9 system.

- Within the AAS system, Cyan and Blue were as different in appearance as possible. The AAS Cyan was relatively more saturated than its ENO9 counterpart. The AAS Blue was very saturated compared to the ENO9 blue. Highly saturated blue colors are not recommended for visual displays. Aside from the reasons stated above, decreasing sensitivity to short-wave light with increasing age and the inherently low luminance of blue phosphors limit the usefulness of saturated blue targets. Suggested alternatives to the use of saturated blue targets is to use cyan or magenta, or to desaturate the blue appearing light.

- The ENO9 Magenta was quite desaturated in comparison to the AAS Magenta. Under casual observation conditions, this author noted that the ENO9 Magenta was not distinguishable from white.

- As noted above, the white points chosen in the two sets were also different in appearance. The AAS White was much more neutral than the ENO9 White, which had a distinctly blue tint. The functional reason for this difference was difficult to determine. At times a white point in a color set is shifted in a given direction from neutral to compensate for chromatic aspects of ambient lighting. More information is required to understand this difference.

There are no published reports to indicate that any studies have been conducted to examine the effects of the ENO9 or the AAS color set on controller performance. Based on the chromaticity coordinates alone, there is no foundation for recommending one set over another. There are potential problems with both schemes. Any color scheme that is chosen should be tested before implementation to determine whether it is suitable for the task and the environment.

3.5 Creating Shades of Color

In addition to the seven colors specified in the two color schemes, additional shades of some of the colors may be needed. This is particularly appropriate for the use of blue, which may use shades of color to indicate different intensities. There are at least three ways to vary the shade of a color for coding purposes: 1) modify hue, 2) modify saturation, 3) modify intensity. Each is discussed below.

3.5.1 Shades of Color: Hue

Small changes in the hue of a color may be an effective way to create different color shades. The hue dimension of a color is what is most commonly referred to as the color name (i.e., red, blue, green, etc.). On a CRT, small amounts of a second color can be added to an existing color to create different shades. For example, incremental amounts of green added to blue can result in a series of blue-green colors that may be effective for shading. Although the colors are primarily blue, the amount of green could indicate the intensity of the situation.

3.5.2 Shades of Color: Saturation

Modifications in the saturation of a color can also be used to vary the shade. Saturation, or color purity refers to the dimension of a color that defines it's appearance relative to a neutral color (e.g., gray). Every color can be described as having a chromatic (e.g., blue) and an achromatic (i.e., white or black) component. The relative balance between the chromatic and achromatic components determines the saturation. Desaturated colors have a pale appearance and are dominated by the achromatic component. As the balance is shifted toward the chromatic component, the color becomes more saturated in appearance. The saturation dimension, that is the relative balance between chromatic and achromatic components, can be varied in a manner similar to the hue variation noted above, in order to achieve different shades of a color. To create shades of blue, small amounts of both the green and red phosphor must be added to the contribution of the blue phosphor to achieve a series of desaturated blue shades. As with the example of hue variation, the level of saturation could be used to indicate the intensity of the information being displayed.

3.5.3 Shades of Color: Brightness

The third method for creating shades of a color is to modify the brightness of the color. Shades of blue, for example, can be created by decreasing the output of the blue phosphor. The brightness of the color could also serve as a code for the intensity of the information.

3.5.4 Shades of Color: Discussion

The magnitude of the difference between shades of a color must be chosen carefully, regardless of the dimension that is modified. Small color differences may not be noticed by the observer on a consistent basis. Large differences in hue, saturation or brightness between shades of a color may result in confusion. Continuing with the example of adding green input to create shades of blue, the addition of too much green will shift the appearance of the color from primarily blue to primarily green. Should the situation arise that a primarily green appearing shade of blue is displayed in isolation, the color will not be easily interpreted. There is a limited range for each dimension over which useful shades can be defined, and a finite number of shades within that range. Only empirical testing will determine what will work.

Modifications to the hue or saturation of a color to create different shades may also result in differences in brightness between two shades of a color. Increasing the contribution of one or two phosphors without decreasing the level of the main phosphor will result in a higher luminance for the new shade. For example, the addition of light from the green phosphor to light already emitted by the blue phosphor will result in both a hue shift, and an increase in luminance. Only a proportional reduction in the light emitted by the blue phosphor will maintain the luminance level between the two shades. If brightness uniformity is essential, then shade variations must be carefully selected and measured, and the display equipment must be calibrated and monitored for fluctuations in color appearance.

Alterations in the saturation or luminance of a color to create different shades may also create unintended hue differences between two shades. As a color is desaturated, there is an increasing chance of a shift in hue depending on its wavelength. This phenomenon is known as the Abney hue shift.

In the luminance dimension, another type of hue shift can occur. This phenomenon is called the Bezold-Brucke effect. At low luminance levels, red and green color appearances are dominant. At higher levels, yellow and blue appearances are more dominant. Therefore, an orange appearing area at low luminance may become distinctly more yellow in appearance as the luminance is increased.

The preceding discussion indicates that the creation of shades of color must be performed carefully, with deliberate attention paid to the dimensions being altered. An alternative to the use of shades of color that may be less problematic for the designer is to use variations in fill-patterns. Different fill-patterns can be created by altering the number and arrangement of pixels in an area that are active at any time. The number of active pixels in a pattern can be used as a code for the intensity of the information. This method has the advantage that hue, saturation and brightness can be held constant over the different fill-patterns. The main disadvantage is that fill-patterns require the use of relatively large areas on the display screen.

4. Color Specification for the ATC Tower

The ATC tower imposes a unique set of restrictions on the selection of colors for CRT displays. The wide range of ambient illuminations in the tower requires careful color selection. Neither the ENO9 nor the AAS color schemes were designed for use on displays used in ATC towers. If color displays are to be included in the tower environment, however, it is preferable to use the same, or at least similar color schemes to those that will be used in other ATC settings. In this section, information is provided on the specification of colors for display on CRTs in the ATC Tower. The unique ambient lighting conditions of an ATC tower are discussed first, along with the implications for the display of colors on a CRT. Specific attention is focused on the problem of sunlight readability. Second, estimates of the performance of the ENO9 and AAS color schemes in the tower environment are provided. The relative merits of each scheme are considered. Third, an additional color scheme that has been designed for high ambient conditions is presented and compared to the ENO9 and AAS color schemes. Finally, suggestions for the selection of additional colors to the set of seven are provided that take into account the difficulties of the tower environment.

4.1 Ambient Lighting in the ATC Tower

The Tower environment is the most demanding on display screens in terms of the ambient lighting conditions that must be accommodated by the display equipment. Daytime light levels have been measured in the vicinity of 6500 fc. At night, however, the tower remains dark so that the controllers can remain adapted to the outside environment. From the perspective of the display of colors on a CRT, the high ambient daytime light levels create the most difficulty. CRT devices rely on the emission of light for the generation of an image. In order for a CRT image to be visible, the emitted light must, at a minimum, be relatively close to the average light level in the environment. High ambient conditions create the need for a high level of light output from the CRT. The AAS specifications for tower displays require a minimum luminance of 75 fL. The brightness of these displays can be turned down to accommodate lower ambient conditions.

4.1.1 Luminous Contrast and Sunlight Readability

An additional consideration to the luminous output of ATC tower displays is the minimum contrast ratio of displayed information. Contrast is a measure of the difference in luminance between displayed information and its background. The lower the contrast, the less visible is the displayed information. Contrast ratios are calculated by dividing the luminance of the displayed information by the luminance of the background. The following equation can be used to determine the contrast ratio:

$$\textbf{contrast ratio} = L_i / L_b$$

where L_i is luminance of the displayed information and L_b is the luminance of the background. The minimum contrast ratio allowed on ATC displays by the AAS document was 9.0. However, under 2,000 fc of incident illumination, the minimum contrast ratio allowed was 2.0. (Note, however, that under direct sunlight, the incident light in an ATC tower can exceed 6,500 fc.)

On most ATC displays, the information is displayed against a black, or non-emitting background. This typically results in a relatively high contrast ratio. Ambient light reflected off of the CRT screen, however, adds a veiling luminance to the entire display. This condition tends to reduce the contrast of the displayed information as evident in this equation:

$$\textbf{contrast ratio} = (L_i + L_r) / (L_b + L_r)$$

where L_r is the luminance of the ambient light reflected from the display surface. High amounts of reflected light from the display screen, as is the case when sunlight is incident upon the CRT, severely reduce the contrast of the displayed information. When the contrast between the displayed information and the background falls below a threshold level, the displayed information is no longer visible.

4.1.2 Chromatic Contrast and Sunlight Readability

Sunlight incident on a display screen not only affects the luminous contrast, but the chromatic contrast as well. Chromatic contrast can be thought of as the apparent difference between two colors. Quantitatively, this difference is represented by the distance between two points on the CIE chromaticity diagram (e.g., see Figure 1). When sunlight is reflected from a CRT, it mixes with the light emitted from the display. The result is that all of the emitted colors become desaturated, or washed-out. Since the sunlight blends with every color on the screen, all emitted colors are modified in the direction of the reflected light. The overall effect is a reduction in chromatic contrast. The degree of color appearance change of an emitted color toward the reflected sunlight is in direct proportion to the luminance of the two lights. The stronger the reflected light, the more chromatic contrast is lost.

Loss of chromatic contrast with sunlight is illustrated in Figure 4. The filled symbols represent the chromaticity coordinates of the R, G, & B phosphors from a hypothetical CRT with no reflected ambient light from the screen. The open symbols represent the R, G, & B coordinates when sunlight is reflected from the screen. The central symbol is a plot of the chromaticity coordinates of the reflected light. Notice that all three of the phosphors shifted along vectors toward the chromaticity of the reflected light. Also, note that the resulting color gamut, indicated by the dotted triangle, is now considerably smaller than the original color gamut measured in the absence of reflected light. **The ultimate effect of light (particularly sunlight) reflecting off of a display screen is a reduction in the color gamut, or range of possible colors that can be viewed by the observer.** Additionally, all displayed colors shift in appearance toward the color of the reflected light.

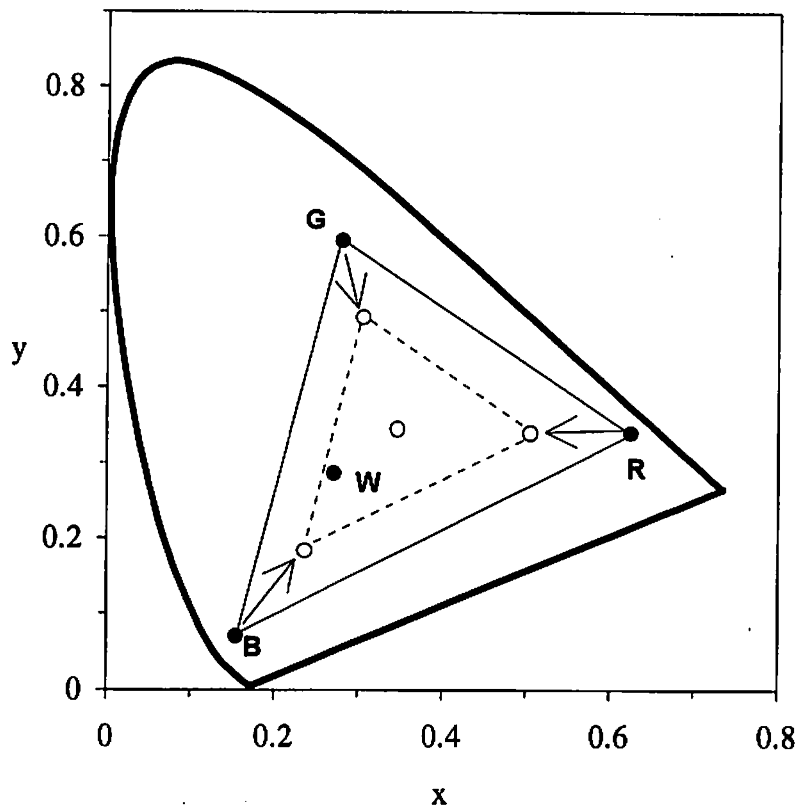


Figure 4. Illustration of restriction of color gamut due to sunlight reflected from display screen. Filled symbols represent chromaticity coordinates of the phosphors without sunlight. Open symbols represent coordinates with added sunlight.

4.2 Estimated Performance of ENO9 and AAS Color Schemes on ATC Tower Display with Reflected Sunlight

The CIE color specification system allows for the quantitative prediction of aspects of the appearance of colors on CRT displays under direct sunlight. In this section, the chromaticity coordinates of the ENO9 and AAS color schemes are predicted under hypothetical sunlight reflection conditions. Both chromatic and luminous contrast reductions are considered and the merits of both schemes are discussed.

4.2.1 Assumptions Underlying Predictions

There were a few assumptions that had to be made before calculating the new chromaticity coordinates. First, an illuminance level and chromaticity coordinates for the incident sunlight had to be determined. To obtain those values, actual illuminance and chromaticity measurements were taken at Logan International Airport, in Boston, MA, at approximately 11:30 AM on 3/10/95. The values are listed in Table 4.

Table 4
Chromaticity Coordinates and Illuminance Level of Sunlight

	x	y	illuminance (fc)
sunlight	0.333	0.360	6325

A second assumption was in regard to the amount of sunlight reflected from the display screen. Due to the refractive index between glass and air, a maximum reflection of about 4% is expected from untreated glass. The electronic display industry, however, has contributed a great amount of energy into reducing reflection with the application of different materials to the display surface. For this reason, two different values, 3% & 1%, were used in the calculations presented below. These values were chosen based on the author's recent experiences with the design of sunlight readable avionics displays. Although lower reflectances can be achieved, a 1% reflectance is readily available in COTS equipment. The luminance of the reflected light can be calculated as a percentage of the incident illuminance in fc and expressed as fL.

A third assumption was that the luminous output of each of the phosphors on the display screen would summate to 75 fL for white appearing light. The relative levels of each phosphor were determined from direct measurement, in the case of ENO9, and from the specifications in the AAS document for the AAS color scheme.

4.2.2 Predicted Sunlight Performance of the ENO9 Color Set

The chromaticity coordinates and contrast ratios for the ENO9 color scheme, with 3% sunlight reflection, are presented in Table 5. Inspection of the contrast ratios indicates a dramatic reduction in luminous contrast with the reflected sunlight. Figure 5 plots the chromaticity coordinates for both conditions. Filled symbols indicate the coordinates without reflected sunlight, and open symbols indicate the coordinates with reflected sunlight. It is clear from the figure that all of the colors shift in appearance toward the color of the sunlight. Further, the distances between all of the points decreased dramatically resulting in a loss in chromatic contrast. Figure 6 is a replot of Figure 5 on a smaller scale. In this figure, the shift toward the color of the reflected light is more clearly evident. The lines pointing from the filled symbols to the open symbols indicate the direction of the color shift under sunlight reflection.

Table 5
Chromaticity Coordinates and Contrast Ratios for ENO9 Color Set Under 3% Sunlight

Color	x	y	Contrast Ratio w/ 3% Sunlight
blue	0.296	0.318	1.21
cyan	0.294	0.333	1.34
green	0.329	0.377	1.13
magenta	0.319	0.337	1.24
red	0.341	0.352	1.13
white	0.315	0.345	1.37
yellow	0.338	0.382	1.32

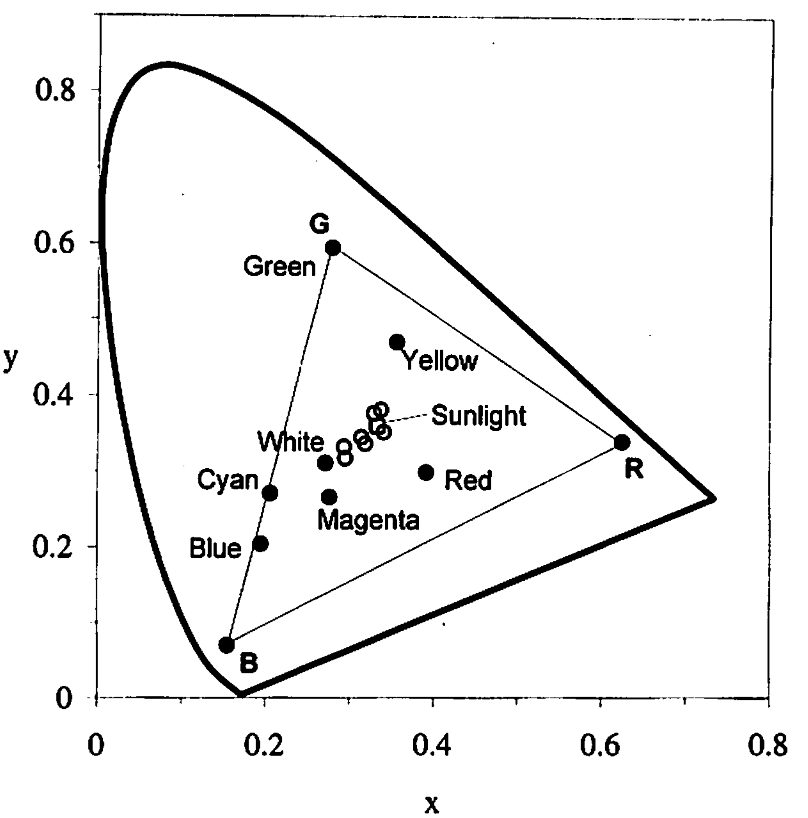


Figure 5. Estimated Chromaticity Coordinates of ENO9 Specified Colors Under 6300(fc) Illuminance and 3% Screen Reflectance

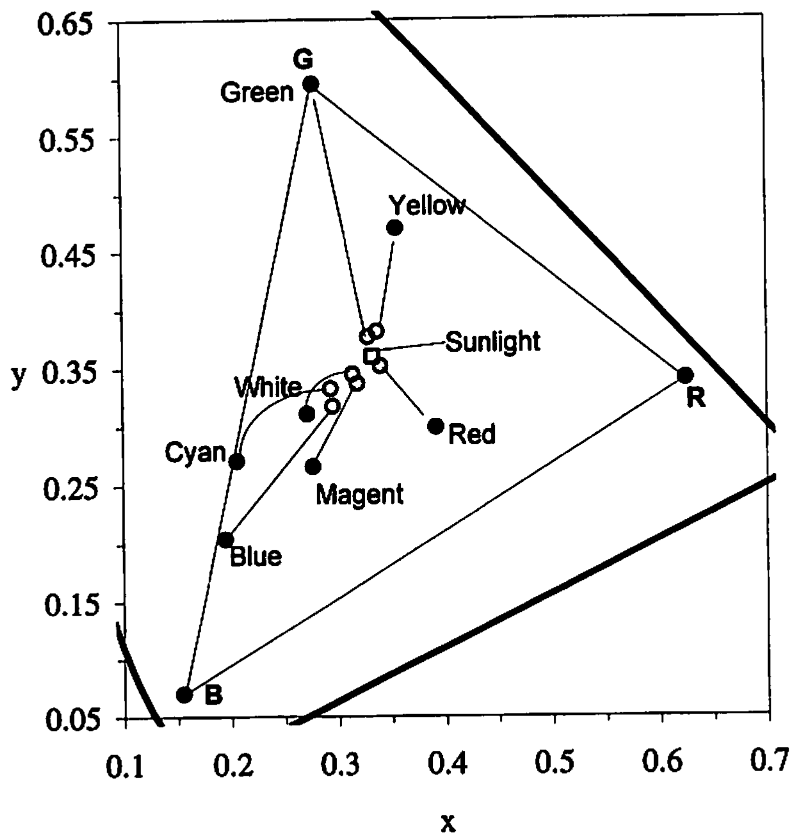


Figure 6. Estimated Chromaticity Coordinates of ENO9 Specified Colors Under 6300(fc) Illuminance and 3% Screen Reflectance-Magnified View

In contrast to the case of 3% sunlight reflection, Table 6 presents chromaticity coordinates and contrast ratios for the ENO9 color scheme under conditions of 1% reflection. A comparison of the contrast ratios between Tables 5 and 6 indicates that luminous contrast improved when less light was reflected from the CRT screen. Figure 7 indicates that chromatic contrast was reduced under these conditions as well. Although all of the colors shifted in the direction of the reflected sunlight, the magnitude of the shift was not as great as is seen in Figure 6.

Table 6
Chromaticity Coordinates and Contrast Ratios for ENO9 Color Set Under 1% Sunlight

Color	x	y	Contrast Ratio w/ 1% Sunlight
blue	0.260	0.278	1.63
cyan	0.260	0.309	2.01
green	0.323	0.405	1.39
magenta	0.305	0.313	1.73
red	0.352	0.341	1.38
white	0.299	0.332	2.12
yellow	0.343	0.407	1.96

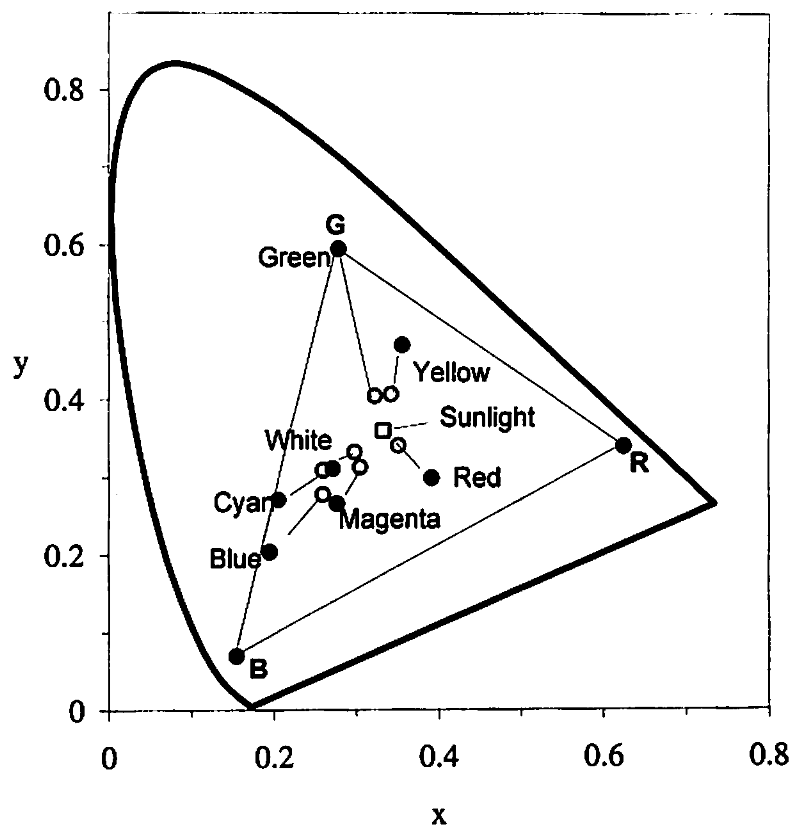


Figure 7. Estimated Chromaticity Coordinates of ENO9 Specified Colors Under 6300(fc) Illuminance and 1% Screen Reflectance

It is clear from the preceding analysis, that a 3% reflection from the CRT under the conditions of direct sunlight found in an ATC tower will lead to an undesirable result. However, although there are large improvements with only 1% display screen reflectance, the ENO9 colors still exhibited a reduction in chromatic contrast.

4.2.3 Predicted Sunlight Performance of AAS Color Set

The extreme loss of chromatic contrast that results from the reflection of 3% of the incident sunlight renders color displays unusable. This point is illustrated in Figure 5 (in section 4.2.2) by the tight clustering of data points. Because of this, the 3% reflectance case for the AAS color set is not considered here. Table 7 presents the chromaticity coordinates and contrast ratios for the AAS color set under conditions of 1% reflectance. (To compare the EN09 and AAS color sets, compare Figures 6 and 7.) Chromaticity coordinates are plotted in Figure 8. The contrast ratios in Table 7 indicate that the AAS color set endured a larger reduction in luminous contrast with the addition of 1% of the sunlight. Figure 8 reveals that the reduction in chromatic contrast still preserved the maximal spacing of the colors.

Table 7
Chromaticity Coordinates and Contrast Ratios for the AAS Color Set Under 1% Sunlight

Color	x	y	Contrast Ratio w/ 1% Sunlight
blue	0.257	0.235	1.15
cyan	0.275	0.346	1.92
green	0.315	0.438	1.82*
magenta	0.351	0.311	1.26
red	0.388	0.356	1.22
white	0.340	0.356	1.70
yellow	0.371	0.384	1.60

* the discrepancy between the contrast ratios for green between the two color sets is due to the assumptions made about the luminance in for each set.

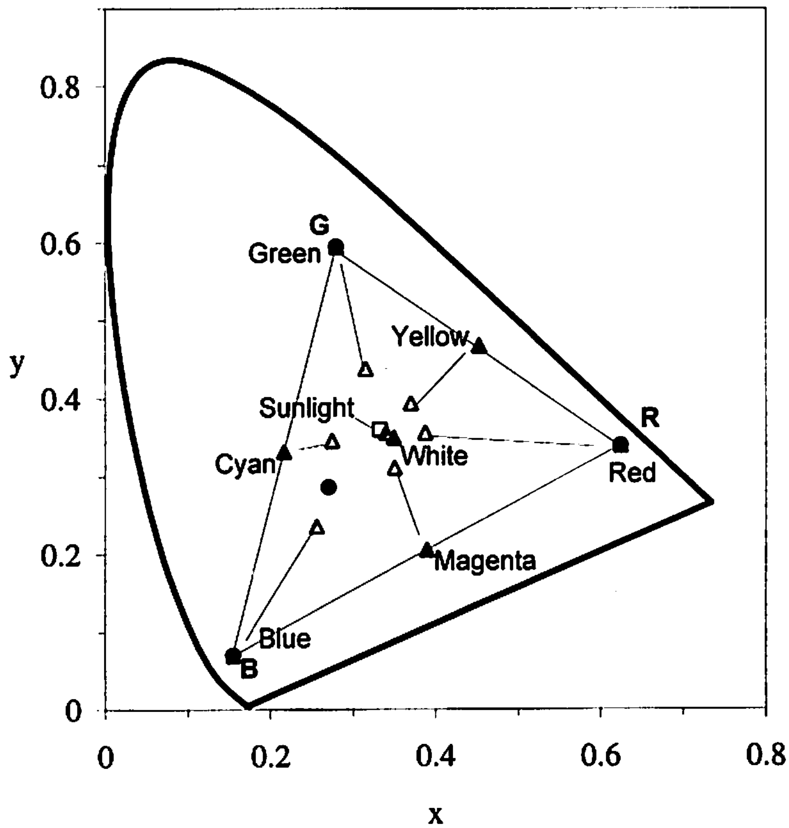


Figure 8. Estimated Chromaticity Coordinates of AAS Color Specifications Under 6300(fc) Illumination and 1% Screen Reflectance

4.2.4 Discussion

A comparison of the values in Tables 6 & 7 indicates that the AAS endured a similar reduction in luminous contrast to the ENO9 color set. Although the contrast ratios for the ENO9 set, in general, are higher than the values for the AAS set, it is not safe to draw the conclusion that the ENO9 set results in higher luminous contrast. The contrast ratios are dependent upon the assumptions made about the luminances of the target colors. Since the AAS luminances were estimated and not measured, it is not safe to make any further comparison between the two sets.

Chromatic contrast does appear to be better for the AAS color set. The larger distances between the points in Figure 8, when compared to Figure 7, indicate that the AAS color set maintained a higher degree of chromatic contrast than the ENO9 set. This difference in performance between the ENO9 and AAS color sets can be traced directly to the chromaticitycoordinates used in each set. The saturated colors of the AAS set resulted in better predicted chromatic contrast under conditions of 1% sunlight reflection.

Assuming that color identification is important, the AAS color scheme may be superior to the ENO9 set due to the greater chromatic contrast that was achieved under sunlight conditions. However, the low luminous contrast values indicate that even with the superior chromatic contrast, the displayed information will still be difficult to see with as little as 1% reflected sunlight bouncing off of the display.

4.3 A Color Scheme Designed for High Ambient Conditions

The advent of the glass cockpit airplane introduced color CRTs into the high ambient environment of commercial aviation. Sunlight readability problems were encountered and solutions were actively pursued. The color scheme chosen by The Boeing Commercial Airplane Company for airborne CRTs is one example of a color scheme devised for high ambient conditions. (Although special systems have also been designed for military applications, the commercial avionics industry provides a closer approximation to the requirements of the FAA in terms of procurement practices than the military. Therefore, solutions determined for military applications are not considered here.) In this section, the colors chosen for the Boeing electronic flight instrument system (EFIS) display are provided as a comparison to the ENO9 and the AAS color sets. Due to lack of information, it is not possible to estimate the performance of the Boeing color set under the same sunlight conditions as in section 4.2. However, there are a number of worthwhile issues to be considered in the comparison that is possible.

Table 8 provides the chromaticity coordinates for the Boeing colors. Figure 9 plots these points relative to the R, G, & B phosphors on the SONY monitor. Inspection of Figure 9 reveals that the Red, Magenta, Green and Blue (estimated) colors of the Boeing set plot slightly outside of the color gamut of the SONY monitor. This indicates that the Boeing color set cannot be produced on the SONY display. However, a comparison of Figure 9, with Figures 7 and 8 reveals that the Boeing set is more similar to the AAS color set than the ENO9 set. This similarity suggests that the Boeing set would perform in a similar manner to the AAS set under sunlight reflection conditions, given that luminance levels were equal. It is also worth noting that the Boeing color set does not include Blue. Blue was determined to be difficult to see, as was purple (Merrifield & Silverstein, 1986). Cyan and Magenta were chosen as replacements to these colors, and ultimately only six were selected for the EFIS display.

Table 8
Chromaticity Coordinates of Boeing EFIS Display (from Merrifield & Silverstein, 1986)

Color	x	y
cyan	0.192	0.207
green	0.300	0.590
magenta	0.321	0.149
red	0.653	0.323
white	0.315	0.274
yellow (amber)	0.468	0.463

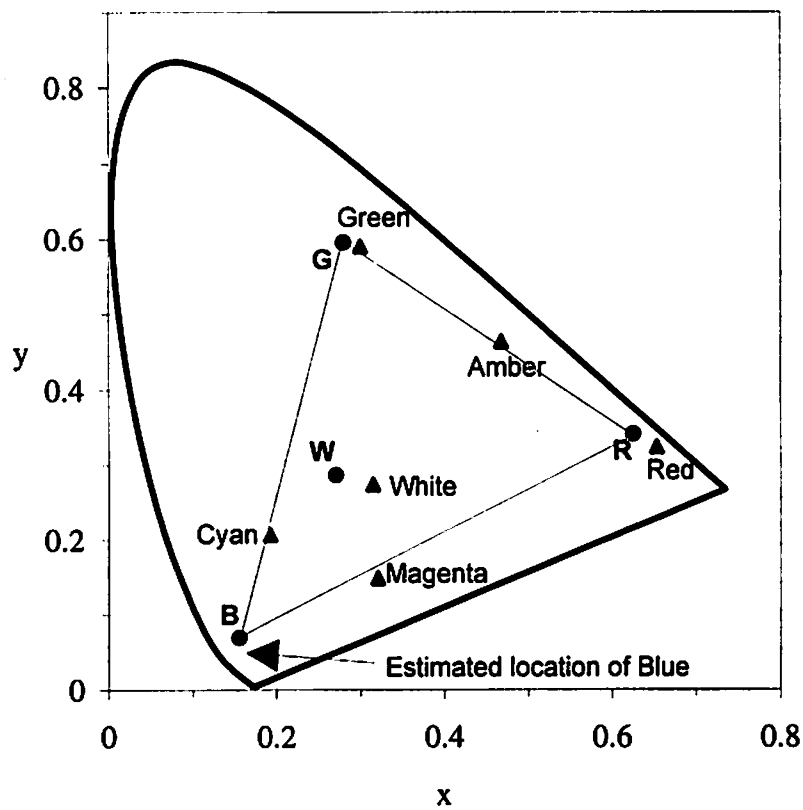


Figure 9. Chromaticity Coordinates for Boeing EFIS Display

Requirements for displays onboard Boeing aircraft include a maximum reflection of 1.25%, with lower values desired. Minimum contrast levels of 2.0 for white light are also required. Even at this level, color discrimination is not optimal (Merrifield, 1995). However, further improvements to this system are not likely. The added cost, weight, and heat of CRTs with higher luminous output prohibit their incorporation into the cockpit. Further reductions in reflectance are also expensive to achieve, and can diminish the viewing angle of the display. For these reasons, the current levels have been established in the Boeing cockpit.

The Boeing values were derived from empirical testing of the equipment prior to installation. The specific needs of the ATC tower will vary from the cockpit and deserve similar attention. It is evident from the preceding discussion, however, that the problem of sunlight readability is not simple to overcome when economic factors are considered as well. It may be determined that the controllers in the tower may have to endure a certain amount of reduction in contrast.

4.4 Choosing Color Beyond the Basic Seven

The number of different colors that can be used in a display of color coded information has been estimated by several authors, with values ranging from four to ten. The choice of six or seven seems to be most popular, based on the belief that beyond this number, color confusions become increasingly likely. Seven colors also represents a somewhat optimal number based on the properties of the CRT. A review of Figure 3 will illustrate this point. The three phosphors, R, G, & B are maximally spaced in the diagram at the corners of the triangle and represent the foundation for the creation of all other colors. The mid-point of each line segment represents a binary combination of each pair of phosphors. Yellow is a combination of Red and Green. Cyan is a combination of Green and Blue. Magenta is a combination of Blue and Red. With the inclusion of white, a seven member color set has been defined.

The addition of any other color to the basic seven starts to create difficulties for color discrimination. For example, orange would be located between yellow and red. However, orange may be easily mistaken for red. Brown is another possibility, however, the perception of brown is highly dependent upon surrounding colors, and it can easily appear closer to green, yellow or orange under some lighting and display (configuration) conditions. Gray can sometimes be used. However, the low luminance required make it difficult to see under high ambient conditions.

Spiker, Rogers, and Cincinelli (1986) provide a detailed account of their color selection process and summarize the effectiveness of their final product. The Society of Automotive Engineers also provides a good checklist for the selection of colors for electronic display screens (SAE ARP 4032). In addition, there are general guidelines for selecting additional colors:

- Specify the name of the color to be chosen (e.g., pink, brown, orange, etc.). Be certain that the name is consistent with the information that is being coded (e.g., use brown for terrain information).
- Determine appearance of the color to be used. Confirm that the appearance has a high correlation with the color name (e.g., is it a good looking red?). Determine that small variations in color appearance will not affect the name given to the color by an observer (i.e., red still looks red and not orange or pink).
- Measure the chromaticity coordinates and plot in the color space along with all other display colors. Determine separation from all other colors. Estimate potential color confusions.
- Predict color appearance under most adverse conditions that are likely to be encountered.
- Assess impact of potential color confusions. For example, what will be the problem for the user if pink is mistaken for red? What will be the problem if red is mistaken for pink?
- Determine if potentially confusable colors can be constrained to not appear on the display simultaneously (e.g. never show pink and red at the same time). Determine the conditions necessary to insure that two colors are not confused.

Ultimately, all colors selected should be evaluated in the operational environment before final confirmation.

5. Enhancements for Improving Display Readability in High Ambient Conditions

The problem of poor sunlight readability of a display screen is essentially a problem of low contrast, both chromatic and luminous, resulting from a large amount of light being reflected from the display. The goal, therefore, in improving sunlight readability is to improve contrast while reducing the overall screen reflectance. Several methods have been devised to attack this problem. In this section, the nature of reflected light is first summarized, and then different methods of improving sunlight readability are presented.

5.1 Reflected Light

Display engineers consider light reflected from a display screen in terms of two components, diffuse and specular. These two components are inherent properties of the surface being illuminated. Diffuse reflections are a result of light being scattered in all directions at the surface. A surface that has a high degree of diffuse reflectance sends light in all directions. A surface that reflects light uniformly in all directions is referred to as a Lambertian surface.

Specular reflections are the result of a coherent reflection of the incident light. The image of the light source is reflected from the screen. A high degree of specular reflection is a characteristic of a smooth, polished surface.

All surfaces have some degree of both types of reflectance. Typically, when one is high the other is low. For example, CRT screens typically have a low specular reflectance, however the diffuse reflectance is quite high. Different media will have different types of reflections.

Diffuse and specular reflections create slightly different problems from the standpoint of display readability. Washout effects from sunlight are most often created by diffuse reflection. There is no angle of view that will reduce the amount of diffuse light. The specular reflection of a light source is typically confined to a portion of a screen and the degree of disturbance it creates, to some extent, is based on the angle of incidence and the orientation of the observer.

The largest problem created by specular reflectance does not come from light sources, as such, but from reflections from other surfaces in the cockpit. Most often, the image of the observer is reflected from the display. This is often referred to as 'white-shirt reflectance,' indicating that the image of highly reflective surfaces will be reflected from the display. The coherent image reflected to the eye of the observer is in a different focal plane than the image from the display, since it is primarily generated by the front surface of the display screen. Although observers are able to focus on the displayed image, their own movements create a motion in the reflected image on the display screen that is inconsistent with the displayed information. Observers find this distracting (Merrifield, 1995).

5.2 Sources of Reflection

The front surface of a display screen is not the only source of reflection. Incident light that is not reflected by the front surface is absorbed or transmitted into the display. This light is bounced around inside the screen by each subsequent surface that it contacts. Every surface that borders another with a different refractive index will create a reflection of some kind. These surfaces include, the front and rear glass surfaces, adhesion layers, the shadow mask on a CRT, internal filters, and the phosphors themselves. The problem of sunlight readability, and display screen reflectance in general, is not simply dealt with by considering the front surface properties of the display.

5.3 Display-based Solutions to the Problem of Sunlight Readability

There are several potential solutions to the problem of display screen reflectance. Display manufacturers have taken general steps to reduce the problem of reflections. Shadow mask areas are typically blackened to reduce the reflection from this surface. Internal filters are sometimes applied to further reduce reflections. Tannas (1985) reviewed some of the techniques that have been developed. They are listed below.

- Antireflective (AR) front surface coatings can be applied to all displays. These coatings serve to reduce the percentage of the light that is reflected. With a series of AR coatings, display reflectance can be reduced to 0.1%. A typical value for common AR coatings is 1% screen reflectance. The antireflective property of these coatings is typically achieved at the cost of increasing the diffuse reflection. The coating applied has a rough surface that scatters incident light rather than reflecting it to the observer. Interference of wavelengths within the applied layer also reduce the overall amount of reflected light (for a discussion see Parish & Arego, 1995).

- Neutral density filters can be combined with AR coatings to further improve display contrast. Neutral density filters absorb light of all wavelengths to an equal degree. When an AR coating is combined with a neutral density filter the result is contrast enhancement. The AR coating reduces the front surface reflectance. Light that passes through the AR coating must pass through the neutral density filter before it can reflect from any of the internal surfaces in the display. Before that light is reflected out of the display it must pass through the neutral density filter a second time. Light emitted by the display only passes through the neutral density filter once. Although the neutral density filter reduces the emitted light onefold, it reduces the remaining incident light twofold. The result is greater contrast, but reduced luminance.

- Narrowband, or notch filters can be combined with neutral density filters to improve display contrast even further. These filters only pass light in a limited range of wavelengths. To the extent that they can be matched to the wavelengths emitted by the display device, reflections of light that are not of the same wavelengths as the emitted light can be reduced. The result is improved chromatic as well as luminous contrast.

- Circular polarizing filters are also used to enhance display readability, although their application is limited to particular types of internal reflections. Linear polarizers are integral parts of liquid crystal displays. In combination with other types of filters, these also provide some degree of contrast enhancement. Linear polarizers, however, tend to reduce viewing angle in one dimension, and therefore, are not widely used.

The application of filters to displays, in general, brings about a reduced reflection of the incident light and an increase in the overall contrast. In all cases, however, filters placed over the display reduce the luminance of the emitted light. Spiker, Rogers, and Cincinelli (1986) caution about the effectiveness of these filters. Even though reflected illumination is reduced, the ambient environment is still very bright. The bright environment sets an adaptation level for the observer. The reduction in luminance of the emitted light caused by the filters may reduce the visibility of the information below a useful level for the working conditions. Therefore, in conjunction with the application of filters, display designers often overcome the problem of sunlight readability by increasing the luminous output of the display device. So called high-bright monitors that are capable of generating 100 fL or more of emitted light, with appropriate filtering, are at present the most widely used solution to the problem of sunlight readability (Stauffer, 1994).

6. The Impact of the Viewing Angle on Display Appearance

Most of the specifications that are provided for display devices in terms of visual performance, refer to the appearance of the display when viewed at an angle perpendicular to the screen. The viewing environment of the ATC tower, however, does not always afford the controller this viewing geometry. The viewing geometry of the environment in which a display is viewed, and the off-angle viewing performance of display devices must be considered in the installation of color CRT displays in the ATC tower. Properties of the environment and display that impact the visibility of displayed information include the distance of the observer from the screen and the size of screen, the height of the screen from the floor, the angle of the observer to the screen and the luminous output of the display at different viewing angles, and the flatness of the screen. In this section, the impact of these issues on the visibility of information displayed on CRTs is considered.

6.1 Display Size and Distance

The distance of the observer to the display has an impact on the visibility of information at two levels. First, the amount of the observer's field of view that is consumed by the display is directly dependent on the observer's distance. The farther away the observer, the smaller the display will appear. This phenomenon is generally discussed in terms of visual angle, or the angle at the eye which is subtended by the image of the display. There is a range of distances that are optimal for viewing display devices, and observers can be both too close, and too far away. For example, concern has been expressed over the nearness of the TRACON controller to the display in terms of the large visual angle subtended by the display. Color coded information in the peripheral regions of the display may be outside of the area of the visual field of the human eye that is sensitive to color. This has lead to the call for redundant coding of information in that environment (Campbell, White, & Hamilton, 1990). In the tower, however, the controller often is farther from the screen than in the TRACON environment. This has the potential benefit of keeping all of the color coded information in the color sensitive field of view.

The second level of impact of viewing distance on the visibility of displayed information is at the micro scale of the individual pixels. Each pixel on the display has a finite size, and, therefore, each pixel subtends a very small visual angle at the eye. (For a discussion of visual angle, see section 3.1.2 of Chapter 3, Visual Perception.) Human acuity, or the ability of the human eye to resolve fine detail, is also limited. Normal human acuity is one minute of arc. Developments in the electronic display industry have focused on improving the resolution of the display screens. For a given viewing distance, smaller pixels result in smaller visual angles for the perceiver. Once pixel size is below one minute of arc, however, the differences between adjacent pixels becomes difficult to discern. Although image quality improves with increased resolution, there is a limit to the improvements in performance that will be obtained with a display with high resolution. Tannas (1994) has argued that the optimal viewing distance from a display device is one that sets the visual angle of a pixel at one minute of arc. A large distance of the controller from the display device, therefore, will effectively result in wasted information. Although the information displayed can be adjusted in size to be visible at fairly long distances, the pixels in the display will not be utilized in the most efficient manner. Since high resolution on a display device increases the cost, it is wise to consider the viewing conditions to determine the actual resolution needs of the observer.

The size of the displayed information is correlated to the size of the display and the viewing distance. Controllers viewing the same display from different distances may perceive the information differently. Information that is optimized in size for the controller who is farther from the screen will appear jagged and perhaps distorted to the near observer. Conversely, information that is optimal for the near observer may be too small for the observer viewing from a greater distance. Even if the information is not below the acuity limit for the more distant viewer, smaller symbols lose their color appearance due to a phenomenon known as small field tritanopia. (For an example of small field tritanopia, see Figure 3-18 in Chapter 3.) If color coded information on a display screen is to be viewed from different distances, then care must be taken to ensure adequate color perception at all distances, or effective redundant coding must be employed.

6.2 Display Position

The position of the display in an environment can have a large impact on the visibility of the information. The viewing geometry in the ATC tower at Logan airport was also measured recently. The average distance of a controller from the display was approximately 6.5 feet. Some of the monitors being viewed were suspended from the ceiling at a height of 7 feet. Other monitors were viewed from a height of 40 inches from the floor. The ceiling mounted monitors were moveable in both position and angle to the observers. The lower monitors were fixed in position. Although the suspended monitors were generally farther from the controllers, their orientation offered protection from some of the ambient reflections. The position of the display, particularly the height, therefore, can be used to control ambient reflections.

6.3 Viewing Angle

The angle of view to a display has a direct impact on the visibility of the displayed information. The angle of the controllers to the monitors at the Logan tower ranged from 90 degrees (i.e., perpendicular) to 51 degrees off angle. The light output of a CRT display is a decreasing function of the angle of view. Most CRTs are considered to be Lambertian emitters. That is, they emit light uniformly in all directions. The luminance off angle, therefore, falls off with the cosine of the angle. At a viewing angle of 51 degrees, the luminous output of the ATC tower displays would be 63% of the luminance measured perpendicularly. At a viewing angle of 60 degrees, the resulting decrease in luminance would be 50% of the value measured perpendicularly. Decreases in luminance result in decreases in contrast as well as darker appearing colors. Controllers may be able to overcome this problem by changing their position relative to the monitor.

Display manufacturers are sensitive to the need for a wide range of viewing angles, particularly in installations that support multiple viewers. However, most commercially available equipment begins to degrade markedly beyond 45 degrees. That is, the Lambertian emission properties of many display devices cannot be maintained beyond 45 degrees from the perpendicular to the screen. Therefore, if more extreme viewing angles are required, actual measurements from the display devices should be gathered to insure adequate luminance and contrast for the viewer.

6.4 Flat Display Screens

A final requirement of ATC viewing environments is that display screens be flat. Flat screens are desirable because they appear to improve the off angle viewing of the display screen. The rounded glass of conventional CRTs creates optical distortions for the viewer at extreme angles. Fortunately, flat-panel displays, including CRTs, are being manufactured in increasing numbers and it should be possible to procure adequate equipment for the ATC tower.

7. Visual Performance Characteristics of Different Display Media

The electronic display industry has been making great strides in the development of flat panel displays. Although CRTs are the most widely used display devices, significant advancements have been made with different media that utilize matrix addressing techniques. Unfortunately, from the perspective of acquiring COTS equipment for the ATC tower, the choices are still somewhat limited. The primary market forces that drive the development of display devices are the entertainment and video game industries. Niche markets such as avionics displays, or ATC displays that have specific requirements, are difficult to accommodate into the production lines of the major manufacturers. The challenge, therefore, is to define the requirements of the ATC displays such that existing technologies can be utilized. Of the current display technologies, the potential contenders for ATC applications include CRTs, liquid crystal displays (LCD), and gas plasma displays. A brief description of the visual properties of each is considered in this section.

The basis of color rendition in all of the electronic display devices to be considered is the additive light mixture of three primary colors. In this respect, all display technologies are similar. Inherent properties in the different technologies, however, create large differences in other respects.

7.1 CRTs

The CRT is the oldest of the display technologies. Currently, large sizes on the order of 36 inch diagonal displays are being manufactured. Resolution is limited by the spacing of the shadow mask and the luminous output of the phosphors. Although very small pixels can be created, the light output would not be sufficient for human viewing. The larger the CRT display area, the larger the depth requirement of the installation site. Larger screens generally require longer picture tubes. Color performance on CRTs is generally quite good, and reflection of sunlight is typically diffuse (Tannas, 1994).

7.2 LCDs

LCD are one of several matrix addressed devices. Matrix addressing refers to electronic control of the individual pixels through an electrode matrix. Liquid crystal displays utilize a layer of liquid crystal material sandwiched between two linear polarizers to control the passage of light. A backlighting system typically provides the light source for these displays. LCDs are flat and require very little depth in their installation. The use of polarizers tends to limit the viewing envelope around the LCD in one dimension more than the other. Usually, the horizontal dimension is maximized. Specular reflections tend to be a larger problem for LCDs than diffuse reflection. Color appearance is controlled by the application of colored filters over individual subpixels. Recently, color quality has begun to rival CRTs. The biggest obstacle to LCDs at the moment, is the limited size of these devices. At present, it is difficult to manufacture an LCD that will suit the size needs of the ATC environment. Predictions in the electronic display market, however, suggest that LCDs will become more popular, and larger and lower in cost in the near future.

7.3 Gas Plasma

Gas plasma displays also rely on matrix addressing. In a plasma display, light emitting gas is sandwiched between two layers of glass and a matrix of electrodes. Gas plasma displays have recently implemented full-color models and are capable of high luminances. The viewing angle is relatively large, and luminous output is peculiarly better off angle than perpendicular to the screen. A target market of the gas plasma technology is large displays. Gas plasma displays also are flat, and are relatively inexpensive compared to comparably sized CRTs and LCDs. In the past, sunlight readability has not been particularly good. However, improvements in design and the application of contrast enhancement techniques have improved performance (Friedman, 1995). The newest color gas plasma displays may be worth considering as potentially lower cost alternatives to CRTs.

7.4 Miscellaneous Display Issues

Although other types of flat panel displays exist (e.g., electroluminescent (EL) displays) current limitations in the technology preclude their consideration. An alternative to flat panel and CRT displays that was also not covered was projection displays. Projection displays are particularly useful in certain applications. Further investigation is needed to determine if there is a feasible ATC tower application for a projection display.

8. Alternative Methods for Controlling High Ambient Lighting Problems

An additional method of improving sunlight readability on CRT displays and enhancing contrast is to reduce the amount of sunlight that is incident on the display. Transparent window shades are currently used in the control tower to reduce the amount of sunlight in the tower and to absorb the heat from the sun. In this section, the impact of transparent window shades on color displays in the ATC tower is presented. Both the ENO9 and AAS color sets are considered. Difficulties with the current sunshades in the ATC tower are also discussed. Finally, an alternative form of window shading device is presented.

8.1 Improvements in Sunlight Readability with Transparent Window Shades

The transmission and colorimetric properties of the window shades in the Logan Airport ATC tower were also measured on a recent visit. The shades were on pull-down rollers that hung from the ceiling. Due to the taper of the windows in the tower, the shades did not completely block all incoming light. The transmission of the shades was approximately 3%.

The impact of the window shades on the ENO9 and AAS color sets can be predicted in a similar fashion to the calculations for the sunlight performance. Table 9 contains the chromaticity coordinates of the sunlight through the Logan ATC tower window shade, as well as the chromaticity coordinates of the AAS color set and resulting contrast ratios assuming 3% and 1% reflectance. Luminance values were based on specified values for the SONY monitor.

Table 9
Chromaticity Coordinates and Contrast Ratios for AAS Color Set Under Sunlight Attenuated by a Transparent Window Shade

Color	x	y	Contrast Ratio 3% Reflectance	Contrast Ratio 1% Reflectance
blue	0.194	0.138	1.48	2.43
cyan	0.231	0.310	3.98	9.93
green	0.272	0.439	3.66	8.99
magenta	0.329	0.231	1.86	3.58
red	0.390	0.288	1.71	3.14
white	0.318	0.316	3.28	7.85
yellow	0.362	0.367	2.95	6.85
			illuminance	
sunlight	0.263	0.259	199 (fc)	

Figure 10 plots the chromaticity coordinates for the AAS color set with and without 3% sunlight added through the window shade. Figure 11 plots the chromaticity coordinates for the 1% sunlight case. Comparable values are provide for the ENO9 color scheme in Table 10 and Figure 12 respectively for the 3% reflection case only.

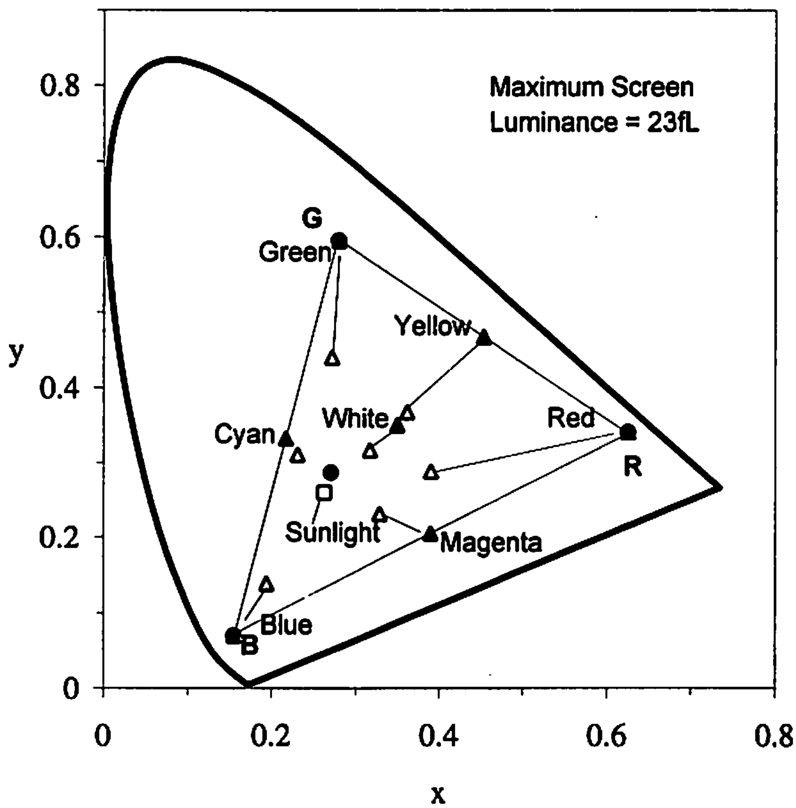


Figure 10. Estimated Chromaticity Coordinates of AAS Color Specifications Under 199(fc) Illumination through Control Tower Window Shade and 3% Screen Reflectance

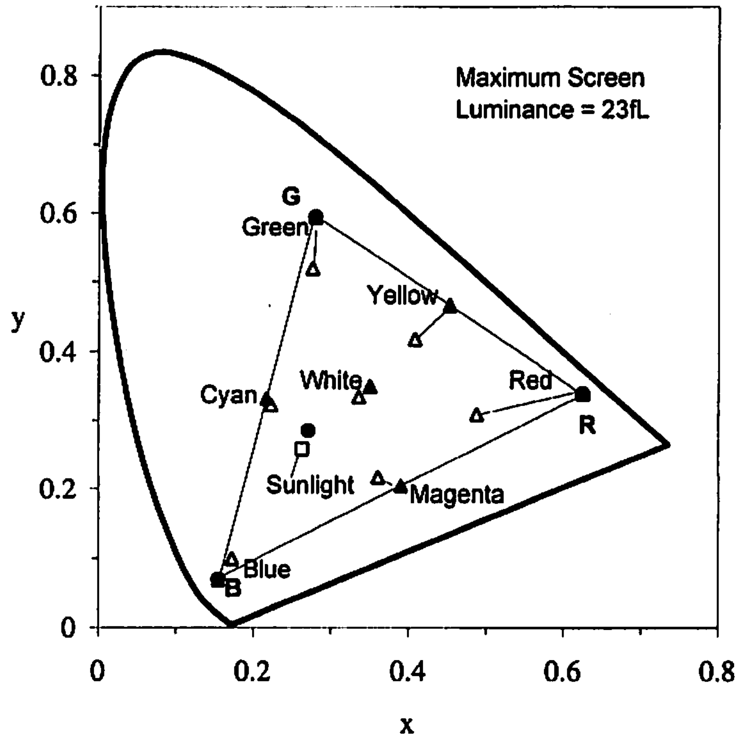


Figure 11. Estimated Chromaticity Coordinates of AAS Color Specifications Under 199(fc) Illumination through Control Tower Window Shade and 1% Screen Reflectance

Table 10
Chromaticity Coordinates and Contrast Ratios for ENO9 Color Set Under Sunlight Attenuated by a Transparent Window Shade

Color	x	y	Contrast Ratio 3% Reflectance
blue	0.208	0.215	4.27
cyan	0.216	0.269	6.19
green	0.270	0.415	2.99
magenta	0.274	0.264	4.75
red	0.344	0.284	2.98
white	0.270	0.302	6.78
yellow	0.331	0.413	5.94
			illuminance
sunlight	0.263	0.259	199 (fc)

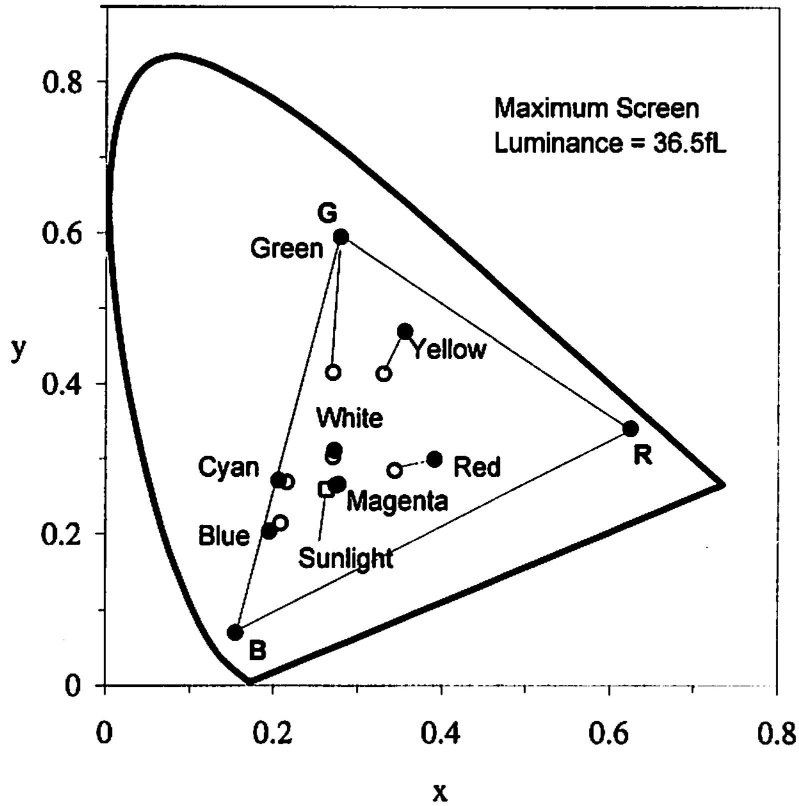


Figure 12. Estimated Chromaticity Coordinates of ENO9 Specified Colors Under 199(fc) Illumination through Control Tower Window Shade and 3% Screen Reflectance

A comparison of Figures 8 and 11 reveals that the effect of the window shades is to greatly improve the chromatic contrast of the display. As can be seen in Figure 10 and Table 9, even with a 3% reflectance, the chromatic and luminous contrast are still quite good with window shades used to block the sunlight. The plot of the sunlight is indicated by the filled square. All of the display colors in Figures 10 and 11 shift toward the sunlight. In this case, the window shades impose a specific blue appearing shift to the displayed colors. The results for the ENO9 color set are also dramatically improved with the window shades in place.

Perhaps the most startling result from this consideration of the impact of the window shades on the color sets is that lower luminances than are normally required in the tower were employed in the calculations. The apparent improvements in chromatic and luminous contrast with the window shades indicates that there is a significant benefit to be gained in the visibility of color CRT displays in the ATC tower with the use of the window shades. Conversations with personnel in the Logan tower, however, revealed that not all controllers prefer to use the shades. In some instances, the shades have not operated properly due to mechanical difficulty. Some controllers simply prefer to wear sunglasses instead of using the shades. This solution does not improve sunlight readability since filtering the sunlight at the eye does not effectively reduce the sunlight incident on the display. The large improvement in luminous and chromatic contrast that are predicted with the window shades suggests that the use of the shades (or other effective method for reducing external light) with color CRT displays should be considered as a requirement.

8.2 Electrochromic Devices

Alternatives to pull down window shades have been developed that may improve conditions in the ATC tower. In recent years, electrochromic windows have been developed that automatically reduce their transmission with the application of a voltage. These devices have been installed in office buildings to regulate heat loss and solar heating, and in automobile mirrors to adjust the reflection of incident headlights (Flannagan, 1995). Electrochromic windows are being developed for aircraft and automobile windows.

Installation of these windows into an ATC tower potentially would solve the sunlight readability problem and allow for use of lower luminance, lower cost CRTs. Lower luminance CRTs would also result in lower overall power consumption and the generation of less heat. The heat absorbing properties of these windows would also help to insulate the tower and further reduce costs. These windows would provide a perfect fit to the window opening so that no unwanted light could enter. With the use of appropriately placed sensors, the transmission of the windows could automatically be adjusted to negate the external ambient illumination. Each window panel could actually be independently tuned to the outside conditions.

Electrochromic materials have the property that their light transmission changes with an applied voltage. When sandwiched between two electrodes and two layers of glass, electrochromic materials create an electrically controllable filter. Transmission values between 70% and 10% can be set. Several different types of windows have been developed. Some windows have a long color memory. Once a signal is applied and the transmission has been adjusted, the signal can be removed and the window will remain in its current state for several hours (Truong, Ashrit, Bader, & Girouard, 1990). Other devices have been developed with a rapid response time. Transmission returns to 70% within a few seconds of removal of the signal.

The selection of electrochromic materials for ATC towers must be carefully considered to ensure that the windows will have the desired properties. Although the ambient conditions outside of the tower vary slowly over the day, it is not clear that slow responding electrochromic materials are desirable. The long color memory of some electrochromic windows could be problematic if a loss of power were to make the window remain in a darkened state. The rapid response of the materials used in automobile mirrors may not be necessary for ATC applications, however, they have the desirable property that the windows clear when the power is removed. Electrochromic materials, like the tower window shades, also have specific colorimetric properties. These will have to be evaluated before specific windows are selected.

9. References

Advanced Automation System: System Level Specification. FAA-ER-130-005H-AP. Updated through December, 1990.

Aschenbach, J. W., Kopala, A. J., & Douglass, L. N. (1982). *Evaluation of Four-Color Plan View Display Console.* DOT/FAA/RD-82/46. Included in this publication is: Kinney, G. C. & DeCicco, M. J. (1982). *Discussion of Trials of Four-Color Plan-View Displays at Washington Air Route Traffic Control Center.*

Campbell, P., White, R., Hamilton, A. (1990). *The Use of Colour on Air Traffic Control Displays.* Transport Canada Publication No.: TP 10532E.

Connolly, D. W., Spanier, G., Champion, F. (1975). *Color Display Evaluation for Air Traffic Control.* FAA-RE-75-39.

EN09. Controller Interaction and Task Analysis. CDRL EN09. Revised draft August 30, 1993. FAA AP1993-1437.

Flannagan, M. (1995). Personal communication. Michael Flannagan is a human factors psychologist at the University of Michigan working on the application of electrochromic mirrors in automobiles.

Friedman, P. (1994). Personal communication. Peter Friedman is a chemist, and president of the Photonics Company, a leading manufacturer of plasma, flat panel displays.

Kinney, G. C. & Culhane, L. G. (1978). *Color in Air Traffic Control Displays: Review of the Literature and Design Considerations.* FAA Contract: DOT - FA78WA - 4075.

Merrifield, R. (1995). Personal communication. Robin Merrifield is a senior display engineer at Boeing.

Merrifield, R. & Silverstein, L. D. (1986). *The Development and Evaluation of Color Systems for Airborne Applications: Fundamental Visual, Perceptual, and Display Systems Considerations.* NADC - 86011 - 60.

Narborough-Hall, N. C. (1985). Recommendations for applying colour coding to air traffic control displays. *Displays*, July, pp. 131-137.

Parrish, M. & Arego, D. (1995). Creating Higher-Contrast Color STN-LCDs. *Information Displays*, 11(2), pp. 10 - 13.

SAE ARP 4032 Human Engineering Considerations in the Application of Color to Electronic Aircraft Displays. *Society of Automotive Engineers*, Warrendale, PA.

Spiker, A., Rogers, S. P. & Cincinelli, J. (1986). Selecting colour codes for a computer-generated topographic map based on perception experiments and functional requirements. *Ergonomics*, 29(11), pp. 1313-1328.

Stauffer, T. (1994). Personal communication. Tim Stauffer is a display engineer with Battelle Memorial Institute.

Tannas, L.E. (1994). Personal communication. Larry Tannas is an electrical engineer, display designer, and industry consultant on displays.

Tannas, L.E. (1994). *Flat-Panel Displays and CRTs.* New York: Van Nostrand Reinhold Company.

Truong, V. Ashrit, P.V., Bader, G. & Girouard, F.E. (1990). *Electrochromic Materials and Devices for Switchable Window Applications.* DSS Contract No. 23283-8-6052/01-SZ. Mines and Resources Canada, Ottawa, Ontario.

APPENDIX B-MASTER CHECKLIST

This checklist is a compilation of the checklist items at the end of the preceding chapters.

Note: Checklist items marked with an "E" indicate items that must be assessed with equipment and/or by referring to the specifications documentation.

I. GENERAL

1. Separately developed subsystems are effectively integrated into the operational environment so that they are compatible with existing equipment and procedures (6.1.2).

2. With this design, the controller can find the necessary information quickly so that the computer does not delay the controller in any way (6.1.2).

3. This design provides all the information needed for planning purposes (5.2.1).

4. This design provides the controller with all the necessary information for a specific task when it is needed/in the appropriate sequence (5.3.2).

5. Information is provided to help the controller recognize situations that require control action (5.2.3).

6. Visual and auditory coding techniques help the controller maintain productive scanning and problem-detection strategies (5.3.2).

7. Perceptual displays help the controller in building and maintaining situational awareness, i.e., in perceiving, integrating, and projecting information about the ATC situation (5.6.1, 5.4.2).

8. Information presentation is split between auditory and visual displays such that neither mode is overused or cluttered (7.3.9).

9. If predictive displays are provided, they assist the controller in projecting the combined effects of many situational factors (5.2.3).

10. If predictive displays are provided, they do not place additional memory demands or other information-processing burdens on the controller (5.2.3).

II. VISUAL DISPLAYS

A. General

1. Information that the controller needs does not disappear from the screen without being deleted or suppressed by the controller (7.1.20).

2. The computer responds quickly so that the controller is not kept waiting for information (5.1.2).

3. Essential ATC information is never blocked or obstructed by other information (7.2.17).

4. All information that a controller needs to accomplish a task that is essential and time-critical is located on a single page or in a single window (7.2.17).

5. Visual displays provide necessary information in a usable form when it is needed (5.3.2, 7.2).

6. Display clutter is not a problem (7.2.20).

7. The meaning of each icon is immediately apparent to the controller or it is labelled (7.2.1).

8. Symbols chosen for the display are intuitive so that the controller can interpret them quickly and accurately (7.2.9, 5.1.5).

9. Controllers can change the amount of task-related detail that is presented (7.2.20).

10. When the meaning of the color is critical, color is used redundantly with another type of visual cue, such as shape, text, or size. For example, all yellow objects have a triangular shape (7.2.12, 3.2.3).

11. The controller is able to recognize and differentiate between color codes under all anticipated lighting conditions (9.6.1).

12. The controller will not need to identify more than five colors (i. e., to interpret the meaning of the color when it stands alone) (7.2.13, 3.2.4).

13. Color displays are readable and adequately bright under all anticipated lighting conditions (9.6.1).

14. When the controller must distinguish between the color of characters and symbols, small blue characters and symbols are not used (7.2.11, 3.2.3).

15. Saturated (i.e., vivid) red and blue are never presented next to each other (7.2.14, 3.1.8; also see Figure 3-11).

16. Colors are far enough apart in perceptual terms that they are not confusable even when "washed out" by sunlight, if applicable (3.2.3).

17. Characters and symbols can be read easily under all anticipated lighting conditions (e.g., from dim light to direct sunlight, if applicable) (9.3.4, 9.6.1, 9.6.2, 7.2.8).

18. Computer displays and controls are clearly visible and easy to use under all anticipated lighting conditions (e.g., from dim light to direct sunlight, if applicable) (9.3.4, 9.6.1, 9.6.3, 7.2.8).

19. To acquire needed information, the controller only needs to look at a single, localized display i.e., switching back and forth between two or more displays is not necessary to perform an individual task. (6.1.2).

20. The position and form of displayed objects appear the same to the controller while seated directly in front of the object as they do from other anticipated viewing angles (7.2.7).

21. If windows are used, the controller can scroll the underlying data set (7.2.17).

22. If windows are used, the controller can move windows (7.2.16).

23. If windows are used, the controller can resize windows (7.2.17).

24. If windows are used, the controller can iconify display pages (7.2.17).

25. If windows are used, the controller can open and close windows (7.2.17).

26. The active window is highlighted to distinguish it from inactive windows (7.2.17).

27. The relationship between different windows is clear to the user (7.2.17).

28. All information that a controller needs to accomplish a given task is located in a single window or within a small number of related windows (7.2.17).

29. Abnormal data are emphasized effectively so that it attracts the controller's attention (7.2.11).

30. Updated data are emphasized effectively so that it attracts the controller's attention (7.2.11).

31. Acronyms in the new display system have the same meanings as in the previous system (7.2.2).

32. Terms in the new display system have the same meanings as in the previous system (7.2.2).

33. Symbols in the new display system have the same meanings as in the previous system (7.2.3).

34. Symbol size can be adjusted by the controller (7.2.10).

35. Visual displays and their labels are sufficiently visible under all anticipated lighting conditions (9.3.4).

36. If size coding is used, it is limited to two widely different sizes (7.2.11).

37. Graphic displays are used only to present information that is naturally pictorial and to present dynamic data (7.2.5).

38. Placement of standard data fields is consistent from one display to another (7.2.18).

39. Formats used within data fields are consistent from one display to another (7.2.18).

40. Labels, terms, and abbreviations are used consistently across the display set (7.2.15).

41. Only one abbreviation is used for each word or item and abbreviations are used consistently on all visual displays (7.2.15).

42. Punctuation is used conservatively and consistently (7.2.15).

43. Continuous text is presented in mixed upper-and-lower case (7.2.15).

44. Computer printouts (in upper and lower case) are available for lengthy text (7.2.15).

45. Visual displays maintain good image quality even at the dimmest possible setting (7.2.8).

E 46. According to the display monitor manufacturer's report, the display refreshes at a rate of 65 cycles (or more) per second so that the display does not appear to flicker (7.2.6, 7.2.22, 3.1.6).

E 47. According to the display monitor manufacturer's report, a displayed object moves no more than .0002 times the viewing distance (in inches) in one second so that no display jitter can be detected (7.2.22, 7.2.23).

E 48. The heights and widths of characters appearing at the center and the four corners of the displays do not vary by more than 10 percent (7.2.22).

E 49. When the center of the display is compared to an edge, brightness uniformity does not vary by more than 50 percent (7.2.22).

E 50. The luminance of dynamic text and symbols are eight times that of the static background (7.2.22).

E 51. All colors are 8 times brighter than the static background symbology (7.2.14).

E 52. When the controller must distinguish between the color of characters, character height is at least 21 minutes of arc (7.2.10, 3.2.4).

B. Visual Alerts

1. Information that the controller must read and understand quickly, such as alarms or critical error messages, never blinks or flashes rapidly (greater than 3 Hz) (5.1.3, 5.1.5, 7.2.11).

2. High-priority alerts and other critical information are located within the central display area (i.e., the central 15 degrees of the area where the controller normally looks, given the normal viewing position) (7.2.11, 9.3.4).

3. Highlighting and blinking are used sparingly (7.2.11).

4. Alerts have a low incidence of false alarms (7.2.11).

5. The same color coding strategy is applied to every display used by the same controller (7.2.12).

6. The color red is used only for warning/danger (7.2.12).

7. Yellow is used to indicate caution (7.2.12).

8. Green is used to indicate for normal/ready status (7.2.12).

9. No more than two levels of blinking are used (7.2.11).

10. If blinking is used, it is cancelable by the controller (7.2.11).

11. For a time-critical warning system (such as a conflict detection or resolution advisory), the controller response time that is assumed by the algorithm has been measured (5.1.5).

12. This design effectively directs the controller's attention by means of alerting, coding, and emphasis techniques (5.3.1, 7.2.11).

E 13. Information that is blinking, has an "on" period that is at least as long as the "off" period (5.1.5).

E 14. If blinking is used, the blink rate is between 2 and 3 Hz (5.1.5, 7.2.11).

III. AUDITORY ALERTS

A. General

1. Auditory alerts are used only when necessary (7.3.2, 7.3.3).

2. The number of auditory alerts is sufficient, that is, auditory alerts are included wherever they are needed (7.3.2).

3. The meanings of auditory alerts are readily apparent (7.3.7).

4. All proposed auditory alerts have been tested and evaluated in a realistic environment by a representative set of controllers (7.4.3).

5. The auditory alert does not nag, or otherwise annoy, the controller (7.3.5, 7.3.6).

6. Auditory signals (and speech messages) are not masked by other auditory alerts or background noise (7.3.7, 7.3.8, 4.1.3).

7. For any situation, it is impossible for more than a few auditory alerts to be presented simultaneously (7.3.7).

8. The number of auditory signals (e.g., warnings, alerts) that the controller may need to identify is fewer than five (7.3.8).

9. Auditory alerts are easily discernible from other signals or noise (7.3.7, 7.3.8, 4.1.3).

10. Auditory alerts do not provide more information than is necessary (7.3.7).

11. The same auditory signal always indicates the same information (7.3.7).

12. Auditory alerts are consistently implemented throughout the system (7.3.7).

13. The information contained in an auditory alert is also displayed visually (7.3.8).

14. Auditory alerts are only used when immediate action is required (7.3.8).

15. Auditory alerts terminate automatically when the problem is corrected (7.3.8).

16. Auditory alerts are cancelable by the controller (7.3.7, 7.3.8).

E 17. A modulated signal emits from one to eight beeps per second (7.3.8).

E 18. A warbling sound varies from one to three times per second (7.3.8).

E 19. The frequency of all auditory signals is between 500 and 3000 Hz so that they are well within the band of frequencies that humans are most sensitive to. (7.3.8).

E 20. Auditory alerts sound for at least a 0.5 second duration (7.3.8).

E 21. The pause between repetitions of an auditory signal is less than or equal to three seconds (7.3.8).

E 22. Auditory alerts are at least 10 dB above ambient noise or have been demonstrated to be sufficiently intense for a specific working environment (7.3.7, 7.3.8).

B. Speech Messages

1. A detection signal display (for example the sound of static on the line) precedes a voice warning, unless a distinctive synthesized voice is used (7.3.3).

2. Speech messages are short enough to be easily remembered. (7.3.4).

3. Brief speech messages are available to the controller when there is the need to explain the specific nature of alarm and warning signals (7.3.4).

4. Speech displays are distinct from and not easily confused with other voices in the control room (7.3.4).

5. The controller does not need to remember more than one or two speech messages at a time in order to accomplish any of his or her ATC tasks (7.3.5, 7,3,7).

6. Speech messages are not masked by other auditory alerts or background noise (7.3.7, 4.1.3).

E 7. If important messages are produced by synthetic speech, they are at least 8 db above the surrounding noise (4.2.5).

IV. COGNITIVE WORKLOAD

A. General

1. With this design, the controller will be able to build and maintain sufficient situational awareness (6.2.1, 6.3.3, 8.3.4).

2. The design assists the controller in detecting errors in data entry (5.2.4, 6.3.3).

3. The design assists the controller in correcting errors in data entry (5.2.4, 6.3.3).

4. This design presents all information in usable form; the controller is never required to transform data from one unit to another or to perform mental calculations in order to use the data (5.1.1, 7.5.3).

5. This design helps the controller to integrate information from multiple sources, if the information is not already integrated before it is presented to the controller (5.1.2, 5.4.2, 5.6.1).

6. The design allows the controller sufficient time to perceive and act upon new information (5.1.1, 5.1.5).

7. The design allows the controller sufficient time to project potential outcomes of optional control actions (5.1.1).

8. This design requires little or no unaided recall of information (5.4.7, 8.2.4).

9. This system provides appropriate memory joggers (e.g., prompts, cues) (5.2.1, 5.4.4, 8.2.4).

10. This design does not require the controller to recall infrequently used data-entry commands (8.2.5).

11. This design does not place greater demands on memory than the previous system did (8.2.4).

12. This design does not increase the amount of data entry for controller tasks (8.2.5).

13. In comparison to the established baseline (e.g., the previous system), controller workload stays about the same with this design (8.2.3).

14. Controllers will be able to make this design work without having to invent ways around design flaws (8.1.5).

15. This design supports timesharing of information processing activities, that is, visual, auditory, and decision-making processes can be performed together without overloading the controller. (5.2.1).

16. This design supports complete, accurate awareness of the ATC situation (5.2.1).

17. This design is not likely to overload the controller's working memory (5.2.1, 5.4.4).

18. The information that is selected and presented supports the controller in making judgement calls and decisions (5.2.1).

19. Information from subsystems is integrated and presented in a way that minimizes the need to switch from one display to another (5.2.1, 5.4.2).

20. This design alerts the controller to critical situations with enough lead time to formulate and execute appropriate responses (5.3.2).

21. This design allows the controller to keep some information processing resources in reserve for unexpected events (5.3.2, 5.4.4).

22. This design calls attention to situations that depart from what the controller would normally expect (5.4.5, 5.4.6, 5.6.2).

23. This design provides adequate support for achieving aircraft separation and for detection of potential conflicts (5.5.1).

24. The design allows sufficient time for the controller to perceive, integrate, project, and act upon ATC information (5.6.1, 5.6.3, 5.1.5).

25. Workload evaluations have considered both observable and perceived effects of task demand on the controller (8.1.1).

26. System demands do not overload or underload the controller for prolonged periods of time (8.1.2, 8.3.3).

27. Participants in workload assessments represent the range of experience, skills and abilities that are present in the controller workforce (8.1.3).

28. Workload has been assessed with an appropriate battery of measures (8.1.3, 8.3.3).

29. This design fosters an active, yet comfortably manageable, role for the controller (8.2.1).

30. When tasks are performed together, workload remains manageable with this design (8.3.1).

31. This design does not contribute to increased information-processing workload (8.3.2).

32. With this design, the controller is able to maintain the highest standards of safety, without having to exert extreme effort (8.3.4).

33. The design supports the controller in making projections about the near-future traffic situation (8.3.4).

34. This design does not require timesharing of many moderately difficult tasks (8.3.4).

35. When timesharing is necessary, the tasks to be timeshared are spread across the controller's resources, that is, visual, auditory, and manual capacities, instead of loading up on just one or two capacities (8.3.4).

36. When timesharing demands are heavy, this design helps the controller remember to execute intended actions (8.3.4).

37. Procedural task sequences are interruptable at any point (8.3.4).

38. With this design, the controller does not experience abrupt changes in normal task loading (8.3.4).

39. With this design, timing of tasks can be flexible (8.3.4).

40. When certain tasks must be completed at specific times, their initiation is at the controller's discretion (8.3.4).

41. Use of this design over time will not have a negative effect on job satisfaction (8.3.5).

42. This design will have positive effects on the ways in which ATC team members interact and communicate with each other (8.3.6).

43. This design provides appropriate information to all members of ATC teams (8.3.6).

B. Automation

1. Automated features behave in ways that are consistent with controller expectations (6.3.3).

2. After system recovery from degradation or failure, a smooth return to automated operations will be possible (6.2.1).

3. Automated features provide explanation of their intentions, recommendations, and actions in ways that are readily understood by controllers (6.3.2).

4. With this design, reversion to manual control will be easy; that is, the controller will have no problem stepping in when the automation fails (6.2.1, 6.3.3).

5. Increased ATC automation results in better integration of data from multiple sources (6.1.2).

6. The limitations of the computer's information and advice are clear to the controller (6.2.1).

7. This design provides an active, involved role for the controller (5.3.2, 5.6.2, 6.2.1, 6.3.3).

8. This design does not require the controller to perform purely monitoring tasks for more than 20 to 30 minutes at a time (5.3.2, 6.2.1).

9. Automated aids are adequately integrated with each other (5.5.1).

10. Decision aids don't need to be monitored continuously (5.6.1).

11. Decision aids benefit the controller (5.6.1).

12. This design supports the controller's development of strategies for dealing with short-term (tactical) and long-term (strategic) situations (5.6.2).

13. Changes in the situation or unusual events are clearly indicated and are not easy to miss (5.6.2).

14. This design will not induce complacency (6.2.1).

15. Provisions have been made to help controllers maintain operational skills and efficiency (6.2.1).

V. DATA ENTRY PROCEDURES

A. General

1. The number of keystrokes (or other control actions) necessary to input data is kept to a minimum and the amount and complexity of data entry is about the same as was required in the previous system (6.2.2).

2. With this system, data-entry errors can be caught and corrected before they propagate through the system (6.3.3).

3. The design assists the controller in detecting and correcting errors in data entry (5.2.4, 6.3.3, 7.5.3).

4. This system makes it easy to recover from data-entry errors (6.3.3, 7.5.3).

5. Keystrokes or other data-entry actions are echoed immediately on the screen, that is, there is no delay in providing a legible representation of what has been entered (7.5.3).

6. The data entry method helps to minimize errors and provides for quick, simple data editing and correction (7.5.3).

7. This user interface system queries the controller at critical choice points, e.g., "Are you sure you want to delete this flight plan?" (6.3.3).

8. A particular data item, such as assigned altitude, must be entered only once; the computer can retain this value and enter it in other fields, as appropriate (7.5.3).

9. The controller receives appropriate feedback on data acceptance or rejection (7.5.3).

10. The computer does not erase all or part of any erroneous data entry (7.5.3).

11. The controller controls the pace of data entry, that is, the computer does not impose time limits or time outs (7.5.3).

12. The computer does not restrict the order in which data items are entered (7.5.3).

13. The computer prompts the controller for data that have been deferred for entry (7.5.3).

14. Data processing is initiated only after an explicit command from the controller (7.5.3).

15. Boundaries indicate where to enter the data and show maximum field length (7.2.17).

16. A cursor appears to indicate data-entry mode and location (7.5.3).

17. The controller can edit all or part of a data field (7.5.3).

18. The controller is not required to enter leading zeroes for numeric entries (7.5.3).

19. When delimiters, such as punctuation, are required to partition long entries, the computer provides the required format and prompts for the order of data entry (7.5.3).

20. Field labels use accepted ATC terminology and are used consistently (7.5.4).

B. Commands and Command Execution

1. Command execution requires minimal controller action (7.5.2).

2. The consequences of destructive commands are explained (7.5.2).

3. Destructive commands require controller confirmation of intention before they are executed (7.5.2).

4. Command execution always occurs by explicit controller action, never as a by-product of another action (7.5.2).

5. The controller can suspend/interrupt or cancel/undo a transaction in progress (7.5.3).

6. Command ordering is consistent from screen to screen/window to window (7.5.2).

7. Command labels use accepted ATC terminology and are used consistently (7.5.4).

8. The relevant command set is displayed to show the controller which commands are currently available (7.5.2).

9. Commands are consistent in their placement across multiple screens, panels, or windows; in their wording; and in their method of activation (7.5.2).

10. The computer indicates the current operational mode (7.5.2).

11. Entry of long sequences of command parameters is not required (7.5.2).

12. Upper- and lower-case letters are accepted as equivalent when the controller is entering a command or command parameter (7.5.2).

13. Feedback is always given to indicate that the computer has initiated a command (7.5.2).

14. Commands should be stated in the affirmative; that is, they should tell the controller what to do, rather than what not to do (7.2.15, 5.1.5).

C. Menus[1]

1. Menu options are phrased to reflect the action executed and worded in user vocabulary (7.2.18).

2. Options that perform opposing actions are not placed adjacent to each other (7.2.18).

3. The number of menu options is between three and ten (five to six options is optimal) (7.2.18).

4. If an option, or set of options, is never available to the user, the option(s) is not in the menu (7.2.18).

5. If an option is temporarily unavailable, it is displayed in the menu, but dimmed (7.2.18).

6. Menu options are organized in logical or functional groupings with clear titles (7.2.18).

7. If not in logical groups, order is by frequency of usage, with most frequently used options at the top (7.2.18).

8. If not in logical groups or by frequency, options are in alphabetical or numerical order (7.2.18).

9. Less frequently executed options and destructive commands are at the bottom of the menu (7.2.18).

10. If similar options are in different menus, the options are ordered in a consistent manner (7.2.18).

[1] from User Interface Specifications for the Joint Maritime Command Information Systems, Version 1.3, by Kathleen Fernandes, November 1993.

11. Each word in the menu is presented in upper and lower case with the first letter capitalized (7.2.18).

12. Cascading submenus appear to the right of the parent menu (below, if space to the right is limited) (7.2.18).

13. When a menu is displayed, the location cursor is in the first available option (7.2.18).

14. When a pop-up menu appears, it appears near the element with which it is associated (7.2.18).

15. A window containing a pop-up menu provides an indication that the menu is available (7.2.18).

16. If they are presented in a vertical list, menu options are left justified (7.2.18).

17. Menu organization supports specific controller tasks (7.2.18).

18. Graphical or textual aids are provided to assist controllers in navigating through menu structures (7.2.18).

19. The controller is required to traverse no more than four levels in a menu structure (7.2.18).

20. When a trade-off is required between menu breadth (i.e., number of options at a level) and menu depth (i.e., number of levels), the design increases breadth rather than depth (7.2.18).

D. Error Messages and User Guidance

1. Error messages are provided whenever needed (7.5.5).

2. Each error message briefly summarizes the specific problem and proposes a specific solution (7.5.5, 7.5.6).

3. Error messages are direct and precise (7.2.14, 7.5.6).

4. Error messages are presented immediately after an error's occurrence (7.5.7).

5. Error messages are not redundant (7.5.6).

6. Guidance messages are presented in mixed upper and lower case (7.5.7).

7. Messages about limits not met or exceeded specify the appropriate range for data entry (7.5.8).

8. Questionable data entries elicit cautionary messages (7.5.8).

9. Feedback regarding processing delays specifies the process, the length of the delay, and completion of the process (7.5.8).

VI. DATA ENTRY AND CONTROL DEVICES

A. General

1. Input devices work in ways that are compatible and consistent with the controller's tasks (7.4.1).

2. The overall design of input devices does not require frequent switching between devices (7.4.2).

3. The input device(s) is/are appropriate for performing the necessary functions (e.g., alphanumeric data entry; selection of displayed objects; cursor positioning) (7.4.3).

4. Input devices have been compared not only for speed and accuracy, but also for factors such as induced fatigue, resolution capability, and space requirements (7.4.4).

5. Controls and their labels are sufficiently visible under dim lighting conditions (9.3.4).

B. Keyboards

1. Alphanumeric keys are arranged consistently on all keyboards that the controller will use (7.4.3). (The preferred arrangement is the QWERTY layout.)

2. Keyboards are readable under all operating conditions and backlit, if necessary (7.4.3).

3. If a numeric keypad is provided, it is visually separated from the main keyboard and arranged in a 3 X 3 + 1 matrix (7.4.3).

4. Function keys are provided for frequently used commands (7.4.3).

5. Function keys are clearly labeled to indicate their function (7.4.3).

6. The functions invoked by the function keys are consistent throughout the system (7.4.3).

7. Keys on keyboards and keypads have no more than two functions (9.2.2).

8. Nonactive keys are left blank (i.e., not labeled) (7.4.3).

9. The key used to initiate a command is clearly labeled "Enter" (7.4.3).

10. Keyed data are displayed quickly (echoed) on the screen (7.4.3).

11. Tactile and auditory feedback are provided in response to keystrokes (7.4.3).

12. The main keyboard is located directly in front of and below the associated visual display, at a comfortable distance from the seated controller's position (7.4.3).

13. Forearm and wrist supports are provided (7.4.3).

14. Alphanumeric keys meet standards for dimensions, displacement, separation and resistance (7.4.3).

15. Guards have been considered for any key that would present a problem if inadvertently activated (7.4.3).

16. If alternative keyboards are featured, they have been tested for usability and operational suitability (7.4.3).

E 17. The slope of the keyboard is adjustable between 15 and 25 degrees from the horizontal (7.4.3).

E 18. Keyboard height is adjustable between 23 and 32 inches (7.4.3).

C. Touchscreens

1. If a touchscreen is used, it is suitable for the task(s) to be performed by the controller (7.4.3).

2. Controllers can achieve sufficient touch accuracy with the touchscreen (7.4.3).

3. Touchscreen displays can be read easily under all anticipated lighting conditions (7.4.3).

4. The touch input strategy (e.g., land-on, first contact, or lift-off) is compatible with the controller's task objectives (7.4.3).

E 5. Touchscreen displays meet standards for required finger pressure (displacement), separation of touch areas, and resistance (7.4.3).

D. Trackballs

1. The trackball can move the cursor in any direction without causing cursor movement in the opposite direction (7.4.3).

2. The trackball allows the controller to move the cursor quickly across relatively large distances and also to precisely position the cursor within a small area (7.4.3).

E 3. The trackball meets standards for physical dimensions, resistance, and clearance (7.4.3).

E. Control Grip Devices

1. Any input device meant to be held and operated by a standing controller can be held comfortably for a period of three to four hours (7.4.3).

F. Mice

1. If a mouse is part of the design, it can be used compatibly with all of the tasks the controller is supposed to perform (7.4.3).

2. Controllers can easily and smoothly position the cursor with the mouse (7.4.3).

3. Movement of the mouse produces cursor movement in the same direction on the display. For example, if the mouse is moved to the left, the cursor moves to the left on the display (7.4.3).

4. The mouse is equally usable with the left or right hand (7.4.3).

E 5. The mouse has no sharp edges and meets standards for width (1.6 to 2.8 in.), length (2.8 to 4.7 in.), and thickness (1.0 to 1.6 in.) (7.4.3).

G. Graphics Tablets

1. Movement of the stylus in any direction on the tablet surface produces smooth movement of the cursor in the same direction (7.4.3).

2. When the stylus is placed at any point on the tablet, the cursor appears at the associated coordinates on the display screen and maintains that position until the stylus is moved (7.4.3).

3. If the stylus and tablet are to be used for free-hand drawing, the device generates a continuous line as the stylus is moved (7.4.3).

4. If a graphics tablet is used, frequent switching to the keyboard is not necessary (7.4.3).

5. The graphics tablet can be located on the workstation within a comfortable distance from the controller (7.4.3).

H. Pushbuttons (Actual and Virtual)

1. Mechanical pushbuttons are sized and spaced to support activation but to prevent accidental activation (7.4.3).

2. The surfaces of "hard" pushbuttons are rough or concave (7.4.3).

3. Labeling of virtual pushbuttons is consistent (7.4.3).

4. The active and inactive states of virtual pushbuttons are visually distinct (7.4.3).

5. The on-off status of software-generated togglebuttons is made clear through the use of labels and graphic indicators (7.4.3).

E 6. Mechanical pushbutton resistance and separation are in the ranges recommended for single-finger operations (7.4.3).

I. Foot Switches and Pedals

1. Positive feedback is provided to indicate activation of the foot switch (7.4.3).

2. The controller is not required to operate more than one switch or pedal with the same foot (7.4.3).

3. Foot switches are positioned for operation by the toe or ball of the foot (7.4.3).

E 4. Foot switches/pedals meet requirements for dimensions, resistance, and displacement (7.4.3).

VII. ERGONOMICS AND WORKSTATION DESIGN

A. User-Centered Workstation Design

1. Controls are designed and arranged to be consistent with the controller's natural sequence of operational actions (9.3.3).

2. Frequently used controls are easy to see and to reach (9.2.2).

3. High-priority controls are centrally located and placed as close as possible to the controller (9.3.4).

4. Workstation dimensions are adequate for the extremes of the controller workforce (i.e., the 5th to the 95th percentile on applicable dimensions) (9.3.1, 9.4.1).

5. Keyboards and workstation controls are consistently backlit if necessary (9.2.2).

6. Adjacent controls are arranged so that sufficient space is between them to operate them easily and minimize the chances of accidental activation (9.3.3, 9.3.4).

7. Controls are spaced far enough apart that they can be easily grasped and manipulated (9.3.4).

8. Workstation controls are sufficiently visible to the controller while seated and standing (9.3.3).

9. Controls are equally accessible and usable by left- and right-handed controllers (9.3.3).

10. Controls can be reached without excessive shoulder movement or back bending/stretching (9.3.3, 9.3.4).

11. Controls used in sequence are located close to each other (9.3.3).

12. Controls used to adjust visual displays are located near the display set (9.3.3).

13. The workstation provides sufficient space for three-person teams (9.3.2).

14. The workstation design concept considers the operational needs of the particular ATC environment for which it is intended (ATCT, TRACON, ARTCC) (9.3.2).

15. Workstations are arranged and spaced to allow ready access by Airway Facilities personnel (9.3.2).

16. It is easy to open or remove equipment covers on racks (9.3.2).

17. Easy access is provided to workstation components (9.3.2).

18. Mirror-image layouts are avoided within and across ATC facilities (9.3.3).

19. The needs of individual facilities have been considered in determining space requirements (9.3.3).

20. Video display monitors are adjustable; they can be tilted in the vertical plane and swiveled in the horizontal plane (9.3.3).

21. Ceiling-mounted display monitors can be swiveled and tilted (9.3.3).

22. Indicators and their labels are sufficiently visible under dim lighting conditions (9.3.4).

23. Workstation labeling is visible and understandable (9.2.2).

24. Computer screen(s) and work surfaces are free from glare and reflections under all anticipated lighting conditions (9.6.2, 7.2.21).

E 25. Levels of electromagnetic radiation emitted by video display terminals have been minimized (9.3.3).

<u>B. Design of Control Room Seating</u>

1. Chairs provided for controller workstations are comfortable and support the lower back (9.4.2).

2. The seat and backrest are adequately cushioned (9.4.2).

E 3. Seat height is adjustable within 15 to 21 inches (9.4.2).

E 4. The seat's backrest reclines between 100 and 115 degrees (9.4.2).

E 5. Arm rests are 2 inches wide and 8 inches long (9.4.2).

C. Design of Communications Equipment

1. Transmitted speech is sufficiently intelligible (9.5.1).

2. Messages conveyed over loudspeakers are adequately intelligible (9.5.2).

3. There are no noticeable squeal problems or echo effects (9.5.1).

4. Listeners can differentiate between multiple channels fed into headphones (9.5.2).

5. Communication is unaffected by delays due to satellite transmission (9.5.2).

6. The speaker is not distracted by his/her own side tone (9.5.2).

7. Headset design helps to maximize intelligibility (9.5.2).

8. No bare metal parts of the headset come into contact with the controller's skin (9.5.4).

9. Controllers who wear glasses can comfortably wear headphones or other communication equipment (9.5.4).

10. Hands-free operation of communication equipment is possible under normal working conditions (9.5.4).

11. Telephone handsets are readily accessible (9.5.4).

12. For multiple telephone handsets, the most frequently used or the most urgently needed handset is the most readily accessible (9.5.4).

13. Volume/gain controls are separate from on-off controls (9.5.5).

14. Volume/gain controls are limited to an audible level (9.5.5).

15. Squelch control is provided to suppress channel noise during inactive periods (9.5.5).

16. The controller can manually deactivate the squelch control (9.5.5).

17. Foot pedals are provided as alternatives to hand-activated microphone switches (9.5.5).

E 18. Appropriate provision has been made for the calibration of microphones and headphones (9.5.1).

E 19. The location of the foot pedal's fulcrum, placement of the foot pedal in relation to other controls, and placement of the foot pedal in relation to the seated controller are satisfactory (9.5.5).

E 20. System-input devices are designed for optimal response to the range of frequencies between 200 and 6,100 Hz (9.5.1).

E 21. Across the frequency response bandwidth, amplitude variation is at or below plus or minus 3 dB (9.5.1).

E 22. Microphones used with amplifiers have a dynamic range that permits them to pick up variations of at least 50 dB in signal input (9.5.1).

E 23. Appropriate techniques are used to minimize the effects of noise during speech transmission (9.5.1).

E 24. If intelligibility testing is conducted, scores are at or above the following cutoff points for the various testing methods (ANSI phonetically balanced (PB) - 90%; Modified Rhyme Test - 97%; Articulation Index - 0.7) (9.5.2).

E 25. When two earphones are in use, sound pressure level can be increased to at least 100 dB overall (9.5.5).

D. Environmental Design

1. Ambient lighting of the workstation is adequate (9.6.1).

2. The controller can easily recognize all key labels on the keyboard, electronic display keypads, and trackball (9.6.1).

3. Lighting at the console shelf is adequate for reading and writing (9.6.1).

4. Under the proposed lighting conditions, the controller can readily locate switches, controls, headset jacks, connectors, handles, and display recess mechanisms (9.6.1).

5. Control labels are readable (9.6.1).

6. The controller is not distracted by shadows, glare, or reflections (9.6.1, 9.6.2).

7. Adequate display contrast is maintained for textual and graphic information (9.6.1).

8. Lighting is adequate for emergency and maintenance purposes (9.6.1).

9. Maintenance lighting does not interfere with controller tasks at the ATC console (9.6.1).

10. Design of the work environment limits sources of distraction (5.2.4).

E 11. The brightest area in the workplace is no more than three times brighter than the darkest area (9.6.1).

E 12. Ambient noise is at or below 65 dB (9.6.4).

E 13. Provisions have been made to reduce ambient noise caused by vibration (9.6.4).

VIII. HUMAN FACTORS TESTING AND EVALUATION

A. Test Methods

1. The proposed method of testing is appropriate (10.4.1).

2. If field observations are used, they are obtained under conditions that are representative of the full scope of actual operations (10.4.1).

3. If questionnaires are used:

    The questions are simple and direct (10.4.1).

    The questionnaire is unbiased (10.4.1).

    The questionnaire is administered as soon as possible after the task (10.4.1).

    The questionnaire is given to a representative set of users (10.4.2).

4. The test variables are operationally defined in the test plan (10.4.2).

5. The test participants (subjects) are representative of the user population (10.4.2).

6. The test design controls for subject bias (10.4.2).

7. The conditions (scenarios) included in the test plan are representative of actual operations (10.4.2).

8. If there is more than one condition, the order of the conditions is counterbalanced or randomized (10.4.2).

B. Analysis of Test Results

1. The proposed methods of data analysis are appropriate for the test design (10.5).

2. The measure of central tendency (i.e., mean, median, or mode) used is the most appropriate for the test results (10.5.1).

3. A measure of variability is included in the analysis (10.5.1).

4. If a measure of correlation is used or proposed, it is appropriate for the data (10.5.1).

5. If a t-test is used or proposed, it is appropriate (for a comparison between two groups) (10.5.2).

6. If an Analysis of Variance (ANOVA) is used or proposed, it is appropriate (for a comparison between more than two groups) (10.5.2).

7. In any report of test results, the level of statistical significance is cited appropriately (10.5.2).

8. If a regression analysis is used, all of points that are predictions and projections are close to the actual data (10.5.2).

9. In the interpretation of test results, the distinction between statistically and operationally significant results is clear (10.5.3).

C. Human Factors Plan

1. Responsibilities for who will develop and control the human factors work are specifically designated (2.3.1).

2. Methods for coordinating human factors concerns and considerations among integrated product team members and contractor personnel are established (2.3.1).

3. The processes and procedures for how the government will direct, control, and monitor the human factors efforts are described (2.3.1, 2.4.1, 2.5.1).

4. The operators and maintainers of the system are described in the plan (2.3.1).

5. The user functions/tasks are described in detail (2.3.1).

6. The system objectives for personnel resources, training, workload, ergonomics, and safety are identified (2.4.1).

7. Key design goals are operationally defined (i.e., described in terms of how they will be measured and evaluated) (2.3.1).

8. Parameters to be used as criteria against which the system will be evaluated are identified (2.3.1, 2.4.1).

9. The tasks and analyses that need to be conducted to support the definition and evaluation of system performance requirements are specified (2.4.1).

10. The system constraints on personnel resources, training, ergonomics, and safety are described (2.5.1)

11. Critical known issues and work to be done to address system performance requirements are identified (2.4.1, 2.5.1).

12. Critical "unknowns" are listed (to be answered/assessed as more information becomes available) (2.3.1, 2.5.1).

13. A feasible schedule is proposed for accomplishing the human factors work (2.5.1).

APPENDIX C-GLOSSARY OF TERMS

Elizabeth D. Murphy, Kim M. Cardosi, and Mary L. Lozano

absolute identification
In auditory perception, identifying a particular signal from a signal set when only one signal at a time is presented. See also **discrimination.**

absolute threshold
The lowest level of intensity at which an individual can always detect a stimulus.

acceptance
In relation to a new computer system, the consensus of the workforce that a new system is workable, useful, and worth the effort required to learn it.

accommodation
The automatic process of adjusting the lens of the eye to bring objects into sharp focus. Occurs with a shift in visual focus from near to distant objects and distant to near objects. Affected by age as the lens gradually loses flexibility.

acoustic confusion
Confusion of similar sounding letters and numbers with each other when they are transmitted aurally; for example, the letters "M" and "N" are easily confused as are the numbers "five" and "nine." ATC phonology (e.g., "niner" for "nine" and "mike" for "m") is designed to reduce acoustic confusions.

active window
In a visual display system with windowing capabilities, the window that is currently accepting input from the user, usually the window that is fully displayed in the foreground. See also **window.**

adaptation
In auditory perception, a temporary lessening of the auditory system's sensitivity after continuous exposure to a sound or set of sounds; in visual perception, the adjustment of the retina's sensitivity to light, which changes as lighting conditions change; in general, a reduction of the perceived intensity of sensation in any sensory modality as the result of continuous presentation of the same stimulus. See also **dark adaptation.**

adaptive user interface
A user interface that is capable of adapting or adjusting its selection and presentation of information on the basis of user characteristics, goals, and plans. A user interface that employs techniques from artificial intelligence to enrich human-computer interaction. See also **artificial intelligence, intelligent interface, user interface.**

aiding	See **decision aiding.**
Air Traffic Workload Input Technique (ATWIT)	An ATC workload measurement tool adapted from the Pilot Objective/Subjective Workload Assessment Technique (POSWAT). See also **POSWAT.**
algorithm	A precise computational procedure for transforming input data.
allocation of functions	See **function allocation.**
alphanumerics	Characters presented on a visual display as letters, numbers, or combinations of letters and numbers (e.g., a callsign).
ambient illumination	See **ambient light.**
ambient light	Light originating from sources other than the controller's visual displays, i.e., the general level of illumination in the control room due to sunlight, fluorescent lights, or lamps at the workstation. Can produce substantial changes in the appearance of a visual display.
ambient noise	All of the background sounds in the work environment; the general level of background noise in the control room. Can alter the detectability of auditory signals.
Analysis of Variance (ANOVA)	A statistical technique used to analyze data collected in behavioral studies and experiments. A method of identifying the relationships (effects and interactions) between test conditions and results.
anthropometry	The science of describing a population (i. e., all individuals in a group) according to physical measurements taken on a representative sample of the population. Provides a basis for the design of products and environments for human use.
application	A specific software program, software environment, or software tool.

artificial intelligence (AI) A concept and a technology for producing "intelligent" behavior (e.g., decision making) in computers; called "artificial" because it is man-made. Generally relies on a data base of encoded knowledge acquired from human experts. Some AI systems are capable of case-based learning, which extends the knowledge base. See **case-based reasoning, expert systems.**

attention As a technical term in psychology, refers to an ability for allocating cognitive resources that allows an individual either to be selective and focus on a specific item or to timeshare resources and attend simultaneously to more than one item or activity.

audiogram A graphic depiction of a person's ability to hear different frequencies, as measured by an audiometer.

auditory display Presentation of information to the ears by signal or speech; aural alerts, alarms, and warnings. See also **signal display, speech display**. Compare with **visual display**.

auditory icon An auditory "image" constructed of single or multiple sound elements which are associated with a specific meaning and/or action; sometimes called an earcon. See also **icon.**

automated aiding Software tools and capabilities that support the computer user's information-processing and decision-making activities. Examples of automated aids in the current ATC system include velocity vectors and range rings. See also **decision aiding.**

automation The process of implementing, in software, tasks that were formerly performed by human operators. See also **software.**

backlighting Illumination directed from behind or beneath a surface that diffuses the light energy.

bench test An evaluation or assessment conducted to establish a point of reference, known as a benchmark, for later measures or judgments of quality.

binaural unmasking The effect of making a masked sound in one ear audible by presenting the masking noise to both ears.

binocular depth cues	Indicators of depth obtained by automatic analysis of information from both eyes. Stereoscopic images resulting from the fusion of the slight difference in visual field for each eye. See also **retinal disparity** and **stereopsis.**
binocular rivalry	Occurs whenever the images presented to each eye are too different to be combined or fused. Results in an alternation between the image presented to one eye and the image presented to the other eye or the perception of a mosaic that combines portions of the two images.
blink rate	The frequency of blinking or flashing for a visually displayed symbol. Should be between 0.1 and 5 hertz (Hz), with 2 to 3 Hz used to generate a strong sense of urgency. See also **hertz.**
brightness	The perceived appearance of a display or indicator along a dimension from dim to bright. Along with hue and saturation, a component of perceived color. Influenced by the surrounding illumination of the room. Contrast with **luminance.**
brightness control	A capability to turn the brightness of a visual display up or down. Should be possible along a continuous scale, not only in discrete steps. Should be limited at both ends so that the display cannot be made too bright or too dim.
brightness uniformity	The extent to which different parts of a computer screen are equally bright.
cascading menu	In the context of computer-human interaction with menu-based systems, a submenu that is displayed when the user selects an option from a higher-level menu. Usually displayed to the right of the parent menu. See Figure 7-23, p. 382. See also **menu, pop-up menu, pull-down menu.**
case-based reasoning	Derivation of conclusions from individual occurrences of events; one basis for computer learning. See also **artificial intelligence.**

ceiling effect In statistical analysis, the likelihood that any manipulated factor will have little or no observable effect because the error rate is too low (close to zero). The effect of a very high baseline when performance is nearly perfect, that is, near the top or "ceiling" of the scale. Compare with **floor effect.**

character size Measured by the height of a displayed character in terms of its visual angle. Recommended character sizes differ for monochromatic and color displays. See also **visual angle.**

check button (or check box) A type of pushbutton implemented in software. Used in dialogue boxes to obtain instructions from the user. In a panel or group of check boxes, one, some, or all boxes can be selected; contrast with radio button. See also **dialogue box.**

chromostereopsis See **color stereopsis.**

coding techniques In visual display design, the structured use of visual cues such as capitalization, bold, underlining, consistency of location, blinking, color, or blank space, to convey specific information, to call attention to information, or to impose an organizational scheme on displayed information. In auditory display design, the systematic use of various sounds and combinations of sounds to convey information. See also **emphasis techniques, highlighting, redundant coding.**

cognition Human thought processes and their components, such as perception, memory, and decision making. Natural processes (e.g., learning and language) by which we come to know about the world and act upon it. See also **information processing.**

cognitive engineering The process of designing an automated system and its user interface for compatibility with the user's natural, human information-processing abilities and limitations.

cognitive interface The hypothetical mental "workspace" where information from the environment, specifically from the computer, is processed and analyzed as the basis for problem solving and decision making.

cognitive psychology The systematic study of human cognition. See also **cognition.**

cognitive task	An unobservable activity by which the human selects, processes, integrates, and manipulates information. Examples include planning, remembering, formulating a concept or instruction, acquiring information and integrating it into situational awareness, projecting future possible outcomes of action or inaction, and decision making. Sometimes contrasted with overtly behavioral or observable tasks.
cognitive task analysis	The process of decomposing information-processing tasks into their component parts for various purposes, such as identifying training requirements, and designing human-computer dialogues. May include a set of techniques such as observation, verbal protocol analysis, and videotaping of operational activities; the documentation produced by such a process. See also **dialogue, task analysis, verbal protocol.**
color	A perceptual experience, not an inherent property of an object; the sensation that results when light waves of various wavelengths strike the retina of the eye. Varies with hue, saturation, and brightness.
color contrast	A comparison between a foreground image and its background based on their relative hue/saturation ratios.
color discrimination	The task of discerning the difference between two or more colors. The ability to differentiate between colors that are simultaneously presented. Compare to color identification, e.g., magenta vs. red.
color identification	The task of associating a unique color name with a specific display color. A pre-requisite for successful color coding, the difficulty of this task increases with the number of display colors.
color stereopsis	An appearance of depth produced when different, highly saturated colors are displayed together; for example, a saturated red object will appear closer to the viewer than will a saturated blue object displayed next to it. Caused by retinal disparity in imaging short and long wavelengths. See also **retinal disparity, saturation, stereopsis.**

command — In the context of computer systems, an instruction from the user specifying a process or procedure to be performed by the computer. Entered via an input device. See also **input device.** Contrast with **menu.**

command button — See **pushbutton.**

compiler — A computer program that acts on software written in higher-level programming languages, translating these higher-level instructions into machine-readable language.

computer-aided conflict resolution — An expert system under development for ATC, whose purpose is to generate and display to the controller resolutions to computer-detected, potential conflict situations. A feature of the Advanced Automation System (AAS). See also **expert system.**

computer-human interface (CHI) — Also known as the human-computer interface; often referred to simply as the user interface. In physical terms, an aspect of a computer system, implemented in both hardware and software, which permits interaction between the user and the computer's coded computational procedures. The visual displays, input devices, and dialogue, taken together as a whole. In conceptual terms, the set of features that support communication between the user and the computer. See also **dialogue, hardware, input device, software, user.**

conceptual model — A simplified representation of system functions designed to convey a basic understanding of how the system works.

consistency — A basic principle of computer-human interface (CHI) design, which asserts that maintaining a similar "look and feel" throughout the user interface supports usability and human performance. A characteristic of the user interface achieved by adherence to design rules. See also **computer-human interface, design rules, human performance, look and feel, usability.**

construct — In the context of cognitive psychology, a hypothesized function or entity, such as long-term memory, workload, or situation awareness, which cannot be directly observed.

continuous device	A control mechanism, such as a mouse or trackball, which operates smoothly across a continuous range of values; contrast with discrete device.
contrast	The difference in luminance between foreground objects and their background or, generally, between any two areas of a visual display. Also called display contrast. See also **contrast ratio.**
contrast ratio	A measure of contrast; the luminance of the foreground divided by the luminance of the background. Also called **luminance ratio.**
contrast sensitivity function	A graphic representation of an individual's ability to perceive a difference in lightness between two areas. A representation of the reciprocal of contrast threshold. A measure used to complement measures of visual acuity. See also **contrast threshold, sine-wave grating, spatial frequency, visual acuity.**
contrast threshold	The level of contrast required for an observer to see a bar pattern, rather than a uniform gray field, when presented with a sine-wave grating stimulus. See also **contrast sensitivity function, sine-wave grating, spatial frequency.**
control grip device (CGD)	A hand-held input device designed for use by tower controllers; attached to a long tether, the CGD allows freedom of movement and head-up input to the computer.
control group	In an experiment or test, a group of participants who do not receive the experimental treatment. Provides a basis of comparison for the effect of the experimental condition.
controls	See **input devices.**
correlation	A statistical technique used to detect the presence and strength of an association between two variables. Does not provide a basis for drawing conclusions about cause and effect.
crew size	The number of people in a work group. For example the crew size of most air carrier cockpits is two or three.

critical flicker fusion (CFF)	The point at which the experience of flicker disappears, (e.g., as a function of the rate of successive flashes of light from a stationary source of light or the rate at which a computer screen is refreshed). See also **flicker.**
cursor	A highlighted, moveable indicator on the computer screen that shows the current location for data entry, editing, or selection of a displayed object.
cursor control ratio	In continuous control, the relationship between distance moved by the control device and distance moved by the cursor; also called control/display ratio.
dark adaptation	The process of becoming accustomed to dim lighting conditions after being adapted to normal or bright lighting conditions. A phased visual process that takes about seven minutes to reach initial adaptation and up to one hour for full adaptation. See also **adaptation.**
data	Raw, untransformed values, often given in engineering units, such as 0.08, 2.1X, or 9:1. As contrasted with information. Sometimes used interchangeably with information. See also **information.**
data entry and control devices	Workstation components (e.g., keyboards, joysticks, trackballs, touch screens, mouses, etc.) used to input information into a system or to make an input to a system function; also called input devices.
data-entry dialogue	The content of the process by which the controller communicates with the computer via input devices (e.g., a sequence of commands entered by keyboard or touches on touchscreen menus). See also **dialogue.**
data link	A technology for automating some controller-pilot communications so that they are transmitted directly to an aircraft display or printer (rather than by verbal communication of controller to pilot).

decision aiding	Automated tools and capabilities to support generation of options and choices among options. See also **automated aiding** and **computer-aided conflict resolution.**
default	A value displayed automatically when a visual or auditory display is accessed or activated. A value or quality subject to change by the user; for example, the default background color on a particular application may be gray, but the user can change it to another color.
delimiter	An element of punctuation or formatting (e.g., comma, slash, space) used to break up data strings or sequences of data entries.
design	The process of arranging elements or components of the whole in such a way that they satisfy the user's or consumer's functional and/or aesthetic requirements.
design guidelines	Generic recommendations for implementing the "look and feel" of user interfaces (such as Avery and Bowser, 1992). The basis for developing detailed, specific design rules for the application of interest. See also **application, design rules, look and feel.**
design product	A preliminary or final result or outcome of the process of deciding what a system will look like and what functions it will perform; for example, may be a prototype user interface, a set of procedures, a facility-layout plan, or a design specification.
design rules	Detailed, application-specific guidance for user-interface developers derived from general human factors guidelines. A set of agreed-upon conventions for implementation of the look and feel of the user interface. See also **design guidelines, human-computer interface, look and feel.**
detection	In auditory perception, deciding whether a given signal is present. In visual perception, recognizing the presence of a visually displayed object or "target" item.
developmental controller	An Air Traffic Control Specialist (ATCS) in training.

developmental test and evaluation (DT&E) — Formal assessment, often of prototypes, that is conducted iteratively during the development stage of the system life cycle; precedes operational test and evaluation.

dialog (or dialogue) — The interaction or "conversation" that takes place between the user and the computer. Design of this interaction may be based on the traditional dialogue types: command-driven and menu-driven, or on some variant or hybrid of these types. See also **command, menu.**

dialogue box — A small window whose function is to obtain information and/or instructions from the user of a computer system. See also **window.**

direct manipulation — A form of interaction which allows the computer user to select visually displayed objects, resize them, or move them about on the screen in ways that mimic direct physical interaction with objects in the real world.

discernability — Ease of identifying the information of interest among other signals, noise, or visually displayed objects. Related to distinctiveness.

discrete device — An input device with non-continuous settings or a specified set of selectable entries, such as a keypad. Contrast to **continuous device.**

discrimination — The perceptual process of distinguishing one stimulus from another; for example, in color discrimination, distinguishing one shade of blue from another; in auditory perception, distinguishing one sound from another.

displacement — A measure of the space that controls move through when they are activated. Certain minima and maxima are recommended for, e.g., the downward movement of keys during data entry.

display — A visual or auditory vehicle for conveying information; examples of visual displays include cathode ray tubes (CRTs), D-Brites, and approach plates. Examples of auditory displays include voice messages and tones that serve as alerts or warnings. See also **auditory display, signal display, speech display, visual display.**

display contrast — See **contrast.**

display factors — Characteristics of visual or auditory displays that affect human performance. For visual displays, these factors include location in the visual field, presentation format, brightness, and contrast. For auditory displays, these factors include loudness and frequency of the signal or message.

display formatting — A design process of laying out and grouping selected symbols on visual displays.

distinctiveness — In auditory or visual display design, characteristics of the signals and symbols displayed that cause them to stand out from other displayed information.

divided attention — The process by which people actively switch their attention between two or more tasks. In ATC, the tasks that require divided attention may include, for example, aircraft separation, communicating with pilots, and planning instructions to be issued later; also called timesharing. Compare to **selective attention.**

duration — Elapsed time between the onset and offset of a stimulus, such as a light or sound.

Dvorak keyboard — An alternative to the standard QWERTY keyboard, developed by A. Dvorak in the early 1940s. (See Figure 7-13, p. 351). Also called the Dvorak Simplified Keyboard. Supports faster typing and less muscle fatigue as compared to results for the QWERTY arrangement. Specified in American National Standards Institute (ANSI) standard ANSI X4.22-1983. Compare to **QWERTY keyboard.**

dynamic function allocation The moment-to-moment process by which system functions are allocated to system personnel or machine components on the basis of current and expected human workload levels. Under low workload, the human would be given more to do. Under high workload, the computer would relieve the human of some routine tasks.

earcon See **auditory icon.**

echo A form of auditory feedback in which either the speaker hears his or her own speech following a slight delay or the listener hears the speech of another person, with a slight delay between the presentation to each ear. Causes distress and disruption of the speaker's ability to proceed and of the listener's ability to understand the message.

embedded control devices "Soft" controls located in window borders, such as zoom boxes, size boxes, and scroll arrows. Activated via cursor positioning and selection. See also **virtual controls.**

emphasis techniques In the context of the computer-human interface, various uses of visual cues to call attention to visually displayed information or objects. Includes highlighting, use of color, brightness, reverse video, blinking, and flashing in addition to the basic coding technique. See also **coding techniques, highlighting.**

ergonomics The study and design of working environments (e.g., workstation, cockpit, automobile) and their components, work practices, and work procedures for the benefit of the worker's productivity, health, comfort, and safety. See also **human factors.**

error See **human error.**

error messages Guidance presented to the user (either visually as a message presented on a screen or auditorially as a voice message) to convey that the computer cannot accept the data as entered. Should convey the nature of the error and a resolution for the error.

error rate	The ratio of incorrect responses (or lack of response where one is called for) to the total number of responses. The ratio of errors actually committed to the total opportunities for error.
error tolerance	A system's capacity to capture or trap input errors before they are propagated through the system.
expectation	Anticipation of the occurrence of a certain stimulus or event; a powerful determinant of human performance.
expert system	Software that incorporates knowledge acquired from subject-matter experts base and draws conclusions about some current issue or situation in the known environment. See also **artificial intelligence, computer-aided conflict resolution.**
eye reference point	The location of the center of the pupil of the operator's eye in the working environment, as when seated at a control console, from which measures are taken of viewing angle, viewing distance, and other related dimensions.
fast-action menu	A menu of frequently used options that allows the user to by-pass levels of the menu hierarchy to reach an option quickly; see also menu.
feedback	See **kinesthetic feedback, subjective feedback, tactile feedback.**
feel	See **look and feel.**
field	1) In the context of the human-computer interface, a space reserved for a specific data entry or set of entries; for example, the data block is made up of multiple fields (e.g., aircraft ID, assigned altitude, ground speed). 2) The operational environment, e.g., air traffic control towers, TRACONs, en route centers, and flight service stations.

field testing Assessment and evaluation of new systems or system components that occurs at field sites, that is, in the operational environment instead of in the laboratory. Benefits include operational realism. Drawbacks include lack of control over variables that may limit what can be studied and may influence the results in unknown ways.

fill-in-the-blank form One of several computer-human interaction styles; requires the computer user to enter character strings, such as commands or identifiers, into labeled fields displayed on the screen. Contrast to **menu.**

first-contact strategy An approach to touchscreen control, which activates the portion of the screen first touched by the user. See also **touchscreen.**

flicker A perception of fluttering or waviness in a computer display, which results when the display is not refreshed frequently enough. Sensitivity to flicker varies with age. See also **critical flicker fusion (CFF), refresh rate.**

floor effect In statistical analysis, the likelihood that any manipulation will have no measurable effect because initial performance is so poor. The effect of a very low baseline of performance near the bottom or "floor" of the scale. Compare with **ceiling effect.**

font In the context of user interfaces, a type style, such as Courier, Helvetica, or Times Roman, used to display alphanumeric information on the computer screen. Sometimes used as an emphasis technique (for example, different levels of organization or importance may be associated with different fonts). See also **typography.**

foot pedal See **foot switch.**

foot switch A control located on the floor under the workstation and operated by foot pressure. In ATC environments, may be used for keying the microphone. Leaves one hand free for other tasks, such as data entry.

free-text entry Permits user-defined composition of messages, comments, notes or other character-based data entries.

frequency — 1) Rate of occurrence. 2) A characteristic of sound waves and electromagnetic waves. Measured in number of complete vibrations or cycles per second. See also **hertz.**

function — In the context of computer systems, a capability of the system, which may be implemented by its human or machine components or by some combination of components. Functions of the ATC system include, for example, monitoring flight progress, posting weather advisories, and resolving conflicts with flow restrictions. Functions are high-level groupings of related tasks.

function allocation — The process of assigning defined system functions or tasks to system personnel or machine components, or to some combination of human and machine components. See also **dynamic function allocation.**

function analysis — The process of defining the what the system has to do (in terms of high-level activities called functions) to achieve system objectives and specified performance requirements. Occurs prior to function allocation.

function keys — Keys labeled (as F1, F2, F3, or (preferably) with the function name, e.g., 'delete'), that serve as keyboard short cuts by combining in one key the actions of a sequence of individual keys. F1 is typically reserved for the Help function.

functionality — A software capability provided by a computer system. Sometimes used in reference to the total set of functions or capabilities implemented in software.

glare — Excessive demand for visual adaptation brought on by the retina's exposure to more light than it can tolerate. May be caused by excessive contrast in brightness between various portions of the visual field, by a light source so bright that adaptation is impossible, or by moving from a dark environment to a bright one; can be reduced by workplace design. See also **reflection.**

graphic tablet An input device made up of a touch-sensitive grid, which may be operated by contacting it with a finger, pencil, stylus, or other suitable instrument. Appropriate for interaction tasks that require accurate pointing, tracing, drawing, or continuous tracking of slow targets (Galitz, 1993; Greenstein and Arnaut, 1990).

graphics As opposed to alphanumerics, pictorial displays that include geometric shapes, lines of various styles, shadows, icons, and other elements used to convey information.

guidance messages See **user guidance.**

habituation In auditory perception, a lessening of response that cannot be attributed to adaptation. Results from a natural tendency to tune out repeated environmental cues in favor of new information. See also **adaptation.**

hardware System components such as workstation consoles, computers and computer terminals, and communications equipment. Distinguished from software.

harmonics In the context of auditory perception, frequency components that are higher than the fundamental frequency (i.e., the pitch of the lowest frequency component of a sound). Affect the quality or timbre of a sound. Produce the familiar differences in sound quality between different musical instruments, such as a trumpet and a piano. In general, the science of musical sounds.

head-coupled auditory display A sound-localization technology under development for military cockpits. By varying the intensity of sounds and their timing for display to each ear through headphones, causes the pilot to hear sounds as if they were coming from different locations; a technique to improve situation awareness.

hertz or Hertz (Hz) The standard unit of frequency, equal to one cycle (vibration) per second. Used internationally to measure the frequency of waves and vibrations. Named for Heinrich R. Hertz, a German physicist, who discovered electromagnetic waves. See also **frequency.**

hierarchy	A tree-like structure or framework with one node at the top connected, by branching, to other nodes at lower levels; for example, a structure of persons or things arranged in order of rank and responsibility, such as an organization chart or a computer menu system.
highlighting	In the context of user interfaces, an emphasis technique for calling the user's attention to some displayed area or information; for example, certain fields in the data block may be highlighted to indicate a change in value. Typically, involves an intensification and brightening of the background color behind the data or information of interest. Other highlighting or emphasis techniques include reverse video, blinking, bold, changes in color, and the use of lines, borders, and special characters, such as asterisks. See also **coding techniques, emphasis techniques.**
hot key	An alternative way of selecting a menu option. Usually a letter underlined or highlighted in the menu option to designate a key to press for that option, instead of clicking on the option.
hot spot	A location on a visual display, often a graphical object, that displays a pop-up menu upon selection of that spot by the user. See also **selectable object.**
hue	One component of the experience of color (for example, red, green, yellow). Other components are saturation and brightness.
human error	A mistaken intention or an action/lack of action that causes unwanted consequences. A natural, often random characteristic of human performance, which cannot be "automated out" of systems, but which can be reduced and controlled by design.
human factors/human factors engineering	A discipline concerned with the design and evaluation of products and tools that match people's needs and that support effective, efficient, safe human performance. See also **ergonomics.**

human performance Human responses to specified tasks. The extent to which goals for speed, accuracy, quality, and other criteria are met by people functioning in work environments. Measurable behaviors that occur in task situations.

hypermedia A technology that permits linkages between and coordinated display of text, graphics, video, and sound.

hyperstructure The framework of links between objects or nodes in hypertext or hypermedia.

hypertext A technology that allows the definition of links between different but related parts of an on-line document. By selecting a highlighted word in a page of text, for example, the user of hypertext is moved to a related portion of the document. Allows browsing among widely separated topics in large databases, such as on-line encyclopedias. If designed poorly, can create problems in navigation and loss of orientation for the user. See also **hypermedia, hyperstructure, navigation.**

icon In a computer-human interface (CHI) design, a graphical object (e.g., a picture or symbol) that can be selected, can represent a command, or can represent an application; selecting an icon issues a command to the application software or, if the application is iconified, opens the application. See also **application, computer-human interaction, graphics, software.**

illumination The amount of light striking a surface from a single source of light. Equals candlepower divided by the square of the distance of the observer (or the light measurement device) from the source, where candlepower is the output level of the light source. Compare with **luminance.**

image quality The clarity of a computer display as presented to the human visual system. Compatibility of a computer display with the basic characteristics of human vision and with the surrounding environment (Snyder, 1990).

implementation phase The stage of the system development cycle during which the software design is coded (programmed) or implemented by software developers.

information	Facts, knowledge, or concepts that reduce uncertainty; often produced by transforming raw, unintegrated data into meaningful units or measures.
information integration	The process of taking related data (information) from several sources and displaying the various elements together in a way that makes the task easier or more efficient to perform.
information processing	A set of activities and capabilities of the brain, including the perception of objects, manipulation and transformation of new and old information, development of knowledge structures in memory, problem solving and decision making. See also **cognition.**
information transfer	The exchange of information between sender and receiver; the process of communicating job-related information and instructions from speaker to listener, especially when the parties are located at some distance from each other.
information requirements	The facts or knowledge needed by an operator to perform a specific task; for example, route information, posted fixes, aircraft types, and flight identifications are some of the information elements required by the controller to project aircraft arrival flows. Must be identified as a basis for system design and development. Can be expected to evolve during the system design process as operators come to understand automated capabilities and their implications for operations.
input device	A workstation component used for data entry and display control; for example, keyboard, trackball, mouse.
intelligent interface	A user interface that incorporates techniques from artificial intelligence to adjust its selection and presentation of information depending on what it knows about the user. See also **adaptive user interface, artificial intelligence.**
intelligibility	The degree to which a speech signal can be understood by the listener.
intensity	The strength of an environmental cue: volume for auditory cues; brightness for visual cues.

interaction	The communication between a computer system and its user. The complex dynamics of cause and effect among experimental variables.
interaction devices	See **input devices.**
interactive control	User commands or actions that initiate, interrupt, or terminate computer processes. Also known as sequence control.
intuitiveness	A characteristic of a user interface that incorporates natural or learned relationships that are familiar to the operators/users.
invariance	Use of the same auditory signal or visual symbol to convey the same information throughout the user interface. Consistent mapping of information to signals or symbols. Violations of this principle can promote confusion and human error.
jitter	Movement of pixels in displayed objects that should remain in a fixed position. A departure from geometric stability.
keyboard	A kind of input device. See **Dvorak keyboard** and **QWERTY keyboard.**
kinesthetic feedback	Information about an object's movement that is provided by physical contact with the object.
knowledge base	In an expert system, the repository of coded information, acquired from human experts, that is used as the basis for conclusions or recommendations produced by the system.
knowledge-based behavior	Novel, creative approaches or solutions to problems that are not predefined by existing rules. As distinguished from skill- or rule-based behavior by Rasmussen (1983).
labelling	In the context of computer-human interfaces, the process and result of explicitly naming visually displayed objects, such as icons, fields, and pushbuttons.
land-on strategy	A technique for selecting items displayed on a touchscreen. With this approach, an item is selected as soon as the user touches the appropriate place on the screen. See **lift-off strategy.**

legibility In the context of computer displays, the extent to which alphanumeric characters and text are easy for the user to read. Varies as a function of such factors as brightness (luminance) contrast and color contrast (if colors are used).

levels of processing In the context of memory, refers to different kinds of information processing, which are associated with minimal or maximal recall. Superficial processing, such as counting the number of letters in a word, is associated with poorer recall for a list of words than is deeper processing, at a more semantic level, such as using the words in sentences (Craik and Lockhart, 1972). Deeper processing is thought to support development of knowledge structures in memory.

lift-off strategy A technique for selecting items displayed on a touchscreen by removing the finger from the screen; allows the user to "capture" a displayed item and drag it to another location on the screen. See **land-on strategy**.

line of sight An imaginary line extended from the plane of the viewer's eyes; the horizontal line of sight occupies the same horizontal plane as the center of the pupils. The normal line of sight declines 15 degrees below the horizontal; maintaining a horizontal or higher line of sight takes effort and can be fatiguing over time.

localization In auditory perception, identifying the directional source of a signal.

log-in/log-out The process and procedure of initially accessing a computer system (log-in) and of exiting the system when finished (log-out).

long-term memory A psychological construct representing the repository of knowledge in the brain. A critical component of human information processing. Capacity is thought to be infinite.

look and feel "Look" derives from consistency in the visual design of a computer-human interface; "feel" derives from consistency in the behavior of the computer-human interface.

luminance The amount of light reflected upon striking a surface; a measure of the physical properties of light, including wavelength and amount of energy transmitted. Luminance may remain constant while brightness varies with surrounding illumination. See also **illumination.** Contrast to **brightness.**

luminance ratio See **contrast ratio.**

masking In auditory perception, a decrease in the audibility of a particular tone when, in its presence, other tones of higher or lower frequency are sounded.

mean A measure of central tendency in a set of scores on behavioral measures. The arithmetic "average" of the sum of the scores divided by the number of scores. See also **median** and **mode.**

measures of central tendency Indicators of a point on the scale of measurement around which all the scores in a set tend to cluster. See also **mean, median, mode.**

measures of dispersion Indicators of the extent to which scores tend to diverge from "average." Include the standard deviation and the variance. Also called measures of variability.

measures of variability See **measures of dispersion.**

median The score at the 50th percentile in a set of scores, determined by counting an equal number of scores down from the highest score and up from the lowest score. Less affected than is the mean by extreme scores. See also **mean.**

menu A set of related options listed together for selection by the user. See also **cascade menu, menubar, menu options, permanent menu, pop-up menu, pull-down menu.**

menubar In menu-based computer systems, a narrow panel, usually at the top of a computer screen, that continually displays the highest-level menu options for selection by the user. See also **menu, menu options.**

menu navigation — The process of traversing the various levels of options in a menu structure; see also menu.

menu options — The alternative items or actions listed in a visually displayed menu. The user selects from the list of options.

metaphor — In the context of computer-human interaction, an extended comparison between two different objects or concepts that is useful for the purpose of unifying the user interface; for example, comparing the computer screen to a desktop is intended to help the user conceptualize the look and feel of the user interface in everyday terms (e.g., file, record, stack, in-box, mail, wastebasket).

mock-up — A preliminary version of a design product, implemented inexpensively (i.e., in paper, cardboard, or plywood) in order to demonstrate design concepts and elicit feedback from end users. The eventual product that is mocked-up may be a set of computer screens, an aircraft flightdeck, or an ATC workstation.

modality — A neural mechanism for acquiring information or responding to a stimulus; for example, visual, auditory, or tactile modalities (channels).

mode — 1) An internally defined state or condition of computer operation, such as keyboard input mode, help mode, edit mode, or save mode. Overuse of modes can confuse the user into thinking s/he is operating in one mode when, in fact, another mode is active.
2) In statistics, the most frequently occurring score in a set of scores. A measure of central tendency for a group of scores. See also **mean, median.**

monocular depth cues — Perceptual indicators of size and relative distance, when objects are viewed with one eye. These include width and rate of movement.

mouse An input device with one or more selection buttons used to control the movement of a cursor on a computer screen and to make selections by "clicking" on displayed objects. The mouse casing houses a rubber ball that allows the user to move the device on a flat surface. Movement of the cursor occurs lawfully according to the direction and speed of mouse movement.

multimedia In the context of a software environment, refers to capabilities for user interaction via two or more of the following methods: text, graphics, animation, video, or sound. In a hardware context, refers to the attachment of peripherals, such as CD-ROM, to a computer for purposes of importing sound, video, and so forth.

multiple resources theory A theory of attention, which asserts that people have several pools of information-processing resources (e.g., visual, auditory, tactile) instead of a single pool of resources to draw upon (Wickens, 1984; 1992).

navigation In the context of computer-human interaction, the process of maintaining a sense of orientation within an application as one moves from screen to screen, for example, up and down the menu hierarchy. Users may have a sense of getting lost within an application if it lacks adequate navigational aids. See also **application, hypertext.**

negative transfer A performance decrement that occurs when skills or experiences from one work environment contribute to human error in a new environment; that is, old skills interfere with learning and using the new skills required. See also **positive transfer.**

normal line of sight See **line of sight.**

objective data Generally used to refer to quantitative data (that is, numbers) collected by measures that are perceived to be less open to subjective influence than others, such as time to complete a task and number of errors. Subjectivity is present, however, in the data collector's choice of how to identify the beginning and end of a task and how to define what constitutes an error. Compare with **subjective data.**

objective workload — A quantified expression of system demands on the human operator. In ATC environments often expressed in terms of number of aircraft, amount of communication, and so forth. A concept limited by the fact that different operators experience and respond to the same objective workload in different ways, depending on training, experience, and skill. See also **perceived workload.**

observable workload — Contributors to operator workload that can be identified and counted, such as number of aircraft in the sector and pending handoff. Not to be equated with perceived workload. See **objective workload, perceived workload.**

operating system — The set of instructions that manage the internal operation of a computer system. Contrast with **application.**

operational suitability — The extent to which a new system and its components are judged as usable and supportive of critical mission goals, such as aviation safety in the case of ATC systems.

operational test and evaluation (OT&E) — The formal process of assessing operational suitability prior to fielding a new system or system component.

operator error — See **human error.**

parallax — Distortion of the true geometrical relationship between two objects as a function of viewing angle.

parameter — A defining dimension or attribute of an object or process. Cruising altitude is an example of a flight parameter.

parsing — In the context of computer languages, the process of either separating fields in the database or separating data into variable length fields. A means of fitting a limited number of characters into specific field sizes.

pattern matching — An automatic cognitive activity involving the comparison of a perception to information stored in long-term memory. Reading and recognition tasks are pattern-matching activities. See also **pattern recognition, recognition.**

pattern recognition	The outcome of pattern matching, if a match is found in long-term memory for the perception to be identified. See also **pattern matching, recognition.**
perceived workload	The level of effort experienced by the worker, which may vary from person to person for the same task. Also called subjective mental workload. Varies as a function of training, skill, and experience.
performance parameter	A defining characteristic of human performance, such as time to complete a task or level of task accuracy.
performance measures	Metrics used to assess human performance in an experimental or operational environment. May include error rate and speed of response.
performance testing	Evaluation of human performance given a particular set of automated tools and capabilities. Designed and intended to assess the operational suitability of system features.
peripheral vision	Detection of visual cues that occur around the edges of the visual field (that is, in the periphery) as opposed to the center of the visual field.
permanent menu	A menu whose options are continuously displayed on the screen.
pictorial displays	Graphical presentations of information needed to support task performance. Elements may include geometric shapes, lines, graphs, maps, or schematics.
Pilot Objective/Subjective Workload Assessment Technique (POSWAT)	An automated tool for collection of objective and subjective data on pilot workload, developed at the FAA Technical Center. See also **Air Traffic Workload Input Technique (ATWIT).**
pitch	The highness or lowness of a tone, as determined by frequency. The higher the frequency, the higher the pitch of the tone; distinguished from intensity (loudness) and timbre (resonance).

pixel — The smallest discrete picture element that can be addressed, illuminated and resolved by a computer display (ANSI/HFS, 1988; Thorell and Smith, 1990).

pixel spot — Size of a pixel ("picture element") in terms of width and height as measured under specific photometric test conditions (ANSI/HFS, 1988). See also **pixel.**

pointing devices — See **input devices.**

pop-up menu — A menu that is displayed when the user selects a specially-programmed graphical symbol or hot spot on the computer screen. A method of conserving screen space. See also **hot spot, menu, pull-down menu.**

positive transfer — An enhancement in performance that occurs because skills from a previous work environment are applicable in a new environment. Contrast with **negative transfer.**

predictive display — A visual display that simulates the situation at a user-specified point in the future based on trends in the current situation and interventions that may or may not be made.

presbycusis — Hearing loss that increases with age.

presbyopia — The loss in the ability of the lens of the eye to accommodate and bring close objects into focus, as a function of age. Can be corrected by prescription glasses. See also **accommodation.**

priming effect — A phenomenon in human performance that occurs when test participants with prior knowledge and expectations about a system respond faster in an experiment than do participants with no prior knowledge or expectations.

production phase — A stage in the system-development cycle, during which hardware and/or software are built according to specifications.

prompt — A message from a computer system to the user asking for specific input. May be implemented in various forms, such as a word or phrase, followed by blanks, or a dialogue box.

prototype An early, sometimes non-operational, version of a design product, such as a set of user interface screens or a workstation console, developed to try out design ideas. May be implemented on paper, in a cardboard mock-up, or in a prototyping tool that allows some level of apparent but "canned" interaction.

pull-down menu A sub-menu that is displayed when the user selects a menu title (such as File, Edit, View). See also **menu, pop-up menu.**

pushbutton A *physical* pushbutton is a control located in or on the surface of the workstation that is activated by a push of the finger. A *virtual* pushbutton is a software-generated object that is displayed on the computer screen and designed to look like a physical pushbutton. A virtual pushbutton is activated by moving the cursor to it and clicking on it or, in a touchscreen application, by physically touching it. Virtual pushbuttons are also called command buttons.

QWERTY keyboard An input device that uses the standard layout of keys for letters, numbers, punctuation marks, special characters, controls, and functions. See Figure 7-13, p. 351. Contrast with the **Dvorak keyboard.**

radio button A kind of electronically-displayed pushbutton. Indicates that the user should select one option from a set of options.

rapid prototyping Both a process and a set of software tools for quickly generating the "look and feel" of a computer-human interface. See also **look and feel.**

recall Unaided retrieval of facts or other information from long-term memory. Compare with **recognition.**

recognition The matching of information in memory to a stimulus. Identification of a perceptual cue or stimulus as known or unknown. Less demanding, in general, than recall.

redundant coding Using more than one cue (e.g., position and shape or color and shape) to convey the same information. Recommended especially when color is one of the codes because users may not be able to recognize colors under all conditions.

reflection A mirror image of the environment surrounding the computer screen (such as the keyboard, paper on the work surface, lights, or the controller) that is coincidentally superimposed on screen content. Partly reflected from the screen's glass surface and partly reflected from the CRT's phosphor layer (Grandjean, 1988).

refresh rate The rate in cycles per second at which the displayed contents of a computer screen are periodically regenerated by a scanning electron beam. If below 65 cps, flicker may be a problem. See also **critical flicker fusion (CFF), flicker**.

registration Placement and alignment of symbols and text on the display screen. Accurate registration prevents the overlapping of symbols and text on the display screen.

relative discrimination In auditory perception, differentiating between two or more signals presented in the same time frame.

resistance In the manipulation of input devices, the feedback the controller senses due to the amount of displacement of the control and/or the amount of force applied to the control.

reliability In the context of test results, the repeatability of results if the same test were to be given again. A factor in test validity.

resolution A characteristic of a visual display. Expressed in pixels per square inch. See **pixel**.

response time The time required to perform a task or a specified component of a task. In human performance measurement, a measure of the time it takes for a participant in an experiment to respond to experimental stimuli or cues. For example, if upon the sound of a tone, the participant is to press the Enter key, response time is the time elapsed between the onset of the tone and the exertion of force on the key, as recorded by a key-stroke collection mechanism. Sometimes referred to as reaction time.

retinal disparity The difference between the images produced by the two eyes because of their different positions in the head. A binocular cue to depth perception. See also **binocular depth cues, stereopsis**.

rule-based behavior Activity based on pre-defined procedures or rules. As distinguished from skill- or knowledge-based behavior by Rasmussen (1983).

sans serif A French phrase referring to a plain letter type, that is, letters formed as simply as possible, without additional strokes or lines. Used to describe the non-ornamented fonts typically used in computer applications.

saturation A component of perceived color, along with hue and brightness; the extent to which a color is "pure." Varies along a dimension from vivid to "washed-out", depending on how much black, grey, or white is mixed with the pure hue.

selectable objects On a computer screen, items that can be chosen by means of an input device and that, when selected, display further information. Might be a word, an icon, or a shape, for example.

selective attention The process by which people focus on an object, activity, or task to the exclusion of others from fully conscious awareness. Compare with **divided attention.**

sensory store In human information processing, a capacity that supports the brief retention of sensations, which are lost if not attended to immediately.

sequence control See **interactive control.**

shadowing 1) In the context of research on auditory perception, a technique which requires participants to repeat a message while listening to it.
2) In the context of computer displays, the technique of shading at the right and bottom edges of a visually-displayed object to give it a three-dimensional appearance.

shift keying Holding down the shift key while striking other keys. Often requires two-handed operation of the keyboard.

short-term memory The currently active contents of consciousness and the processes for conscious manipulation of information, also known as short-term memory. Constrained by capacity limitations to holding and manipulating just a few items under severe time pressure or stress. Also known as working memory.

signal display A kind of auditory display. May take the form of tones, beeps, buzzing sounds, ringing bells, and so forth. See also **auditory display**.

simultaneous contrast In color perception, a sense of enhanced saturation that results from simultaneous presentation of two colors, one in the foreground and one in the background. For example, a red image appears to be most highly saturated on a green background. As a result of this effect, grays take on a pastel appearance depending on the background color (e.g., gray on red may look light green; gray on blue may look yellow). See also **saturation, successive contrast.**

sine-wave grating A visual stimulus in which luminance is varied systematically, as illustrated in Figure 3-5, p. 66. A set of such bar patterns is used to measure contrast sensitivity. Also called acuity gratings. See also **contrast sensitivity function, spatial frequency, visual acuity.**

situation awareness The extent to which an operator, such as the controller or pilot, has a detailed, integrated understanding of the operational environment.

skill-based behavior Usually, sequences of physical actions that occur "automatically," without the need for conscious thought, such as riding a bicycle. Can also refer to sequences of cognitive activity that underlie expert problem solving and decision making. As distinguished from rule-based or knowledge-based behavior by Rasmussen (1983).

software Instructions, written in a programming language, that implement computer functions.

speech display — A kind of auditory display. A verbal message delivered by a "live" speaker, a recordings, or a voice synthesizer. See also **auditory display.**

speech rate — Speed of both spoken and synthetic (i. e., computer-generated) speech. Preferred speech rates vary between 120 and 200 words per minute for different listeners.

speech-to-noise ratio — The degree to which the speech signal occurs with other, unrelated sounds (i.e., "noise"). If this ratio is high, it is relatively easy to filter the speech from the noise; if it is low, it is more difficult to detect the speech signal.

standards — Sets of recommendations and detailed guidance on various design issues, that have been adopted by formally established and recognized professional groups, such as the American National Standards Institute (ANSI), and government agencies, such as the Department of Defense (DoD).

startle reflex — A person's automatic reaction to an intense, unexpected stimulus, such as a sudden, loud noise.

stereopsis — The ability to judge depth by fusing the slightly different retinal images from each eye. See also **binocular depth cues, color stereopsis, retinal disparity.**

style guide — A collection of the detailed design rules adopted by the design team to guide the development of a user interface. See also **design rules, design guidelines.**

subject — In research design, a participant in a study or experiment.

subject matter expert (SME) — A person who has worked in a particular domain long enough (at least 10 years) to be a skilled practitioner in that area. Someone who can guide non-experts through the intricacies of a technical domain, such as ATC, for such purposes as defining system requirements, documenting an operations concept, or reviewing and evaluating design products.

subjective data — Information collected by self report from participants in a study or experiment. Might include responses to questionnaires and rating scales. Perceived to be less "objective" or more qualitative than are directly measured numerical quantities but can be quantified and used effectively when considered along with data from other measures. Subjective data are among the best indicators of perceived workload.

subjective feedback — Self-reported observations or other qualitative data provided by participants in an evaluation. May include workload ratings and responses to questionnaires. See also **objective data, subjective data.**

successive contrast — An illusory, temporary visual phenomenon in which extensive viewing of one color, such as green, produces afterimages in the complementary (contrasting) color (e.g., pink), when the observer's gaze shifts to a white surface. See also **simultaneous contrast.**

suitability — See **operational suitability.**

symbol — A visual or verbal image which substitutes for an entire concept or entity. For example, the American flag is a symbol of the United States; the White House symbolizes the American government; dotted lines on the controller's display symbolize sector boundaries; a trashcan icon represents the process of deleting files from a word-processing system.

symbology selection — A design process of selecting the symbols to be used on visual displays.

synthetic speech — Machine-generated speech produced by a voice synthesizer.

system — A set of functions designed to meet a goal or a set of related objectives. Key components of automated systems are hardware, software, people (operators, managers, maintainers), procedures, and reward structures. See also **operating system.**

system design — The process of devising automated systems to support attainment of mission goals and objectives. The engineering principles and practices employed for this purpose.

system development phase The stage of the system life cycle during which hardware and software are produced or acquired according to specified system requirements. Follows the system design phase. See also **system life cycle.**

system life cycle The iterative process of conceptualizing, designing, developing, testing, maintaining, and modifying an automated system.

systems approach An outlook or perspective that considers system components not in isolation, but in relation to each other, as parts of the whole.

tactile feedback Information sensed through touch about an action just taken.

task A set of goal-related, operational actions with an identifiable beginning and end. See also **cognitive task.**

task analysis Any of several systematic procedures for identifying the components of tasks, their information requirements, and other attributes, such as performance requirements and necessary skills.

task demand The amount of effort required to perform a task. May include some combination of visual, auditory, analytic, and/or response requirements. Differs between people depending on their relative skill and experience. A component of perceived workload.

task load See **task demand.**

testing See **developmental testing, field testing, operational test and evaluation (OT&E), performance testing.**

thrashing In computer-human interface (CHI) terms, a problem in window management when too many windows are in use simultaneously. Spending time finding windows and making them active. See also **window.**

timesharing Devoting attention to more than one item or activity at the same time; for example, chewing gum while walking, reading the paper and listening to music, monitoring the plan view display (PVD) and listening to voice communications. See also **attention.**

togglebutton A pushbutton that can be in one of two states, on or off; the user "toggles" between the two states by pressing the button to turn it on and pressing it again to turn it off. May be implemented in software to mimic mechanical toggles. On/off status should be coded so that the user is aware of the current state.

touchscreen A visual display technology that allows the user to select objects or locations on the screen by touching the surface of the display.

trackball An input device consisting of a small ball that rotates freely in a socket. Manipulated by the user's fingertips to produce controlled cursor movement.

tutorial Step-by-step instructions, provided with a computer program, that walk the user through the procedures associated with using the program. An on-line training aid.

typography In the sense of digital typography, the available selection of fonts (type sizes and type styles) and their use in organizing the contents of visual displays. See also **font**.

usability The extent to which a system and its components are easy to learn and easy to use.

usability testing The process of evaluating a computer-human interface for ease of learning and ease of use. May be conducted in a usability laboratory. Seeks evidence that the user interface meets pre-defined goals for user performance.

user In the context of automated systems, a person who uses or operates a system by interacting with the internal computational procedures through the computer-human interface. Also referred to as the end user. In ATC environments, the user is generally the controller or traffic management specialist.

user confidence In the context of computer systems, trust in the automation and its components to function reliably, as expected.

user guidance Computer-generated error messages, alarms, prompts, and instructional materials designed to aid the user in navigating through the user interface.

user interface See **computer-human interface (CHI).**

validity The extent to which a measure actually measures what it is designed to measure. See also **reliability.**

variable A factor of interest, such as workload, fatigue, or equipment status, that is defined in terms of units that can be measured in a research study.

variability The extent to which individual measures (e.g., scores) differ from the mean (average).

verbal protocol A sequence of task-related statements, comments, questions or other verbal material collected from a participant in a behavioral study or experiment. A concurrent verbal protocol is collected while the participant is performing the assigned tasks. A retrospective verbal protocol is collected after completion of assigned tasks.

vigilance decrement A rapid decline in attention and accuracy that occurs after about 30 minutes in pure monitoring tasks.

virtual control A software-generated object or widget that is displayed visually and that behaves like a physical control, such as a pushbutton. See also **widgets.**

virtual reality A technology that allows a participant to experience and explore some domain, such as an architectural design, the structure of molecules, or conditions of flight, via a three-dimensional simulation, which the person typically "enters" by wearing a head-mounted display. May permit virtual manipulation of objects within the simulated environment.

visual acuity One's ability to detect fine detail, which depends largely on the ability to accommodate or shift focus from near to far objects and back. See also **accommodation.**

visual angle The angle that a displayed symbol or character forms with the viewer's line of sight.

visual comfort zone The area of the visual field with the best visual acuity; a cone extending from the normal line of sight with a radius of 15 degrees. The optimal area for placement of visual alerts. See also **visual field.**

visual display Presentation of information to the eyes by text, numbers, and/or graphics. May be static, as a map or poster, or dynamic, such as a movie or computer display that updates periodically. Compare with **auditory display.**

visual field The area of the surrounding environment that can be seen without moving the head. Each eye has a slightly different visual field, which combine to form the total visual field.

visual momentum A high level of consistency in the location of objects from one display to the next, which allows the user to build up a natural rhythm in moving between visual displays. For example, the menu bar is always at the top of the screen with options listed in the same order. See also **consistency, look and feel.**

visual scope The user's ability to integrate information displayed in multiple windows into a unified concept or whole. Can be increased by various approaches to window layout. See also **window.**

widgets Software-generated, virtual controls that are displayed visually for user activation. Also called interactive control objects or, simply, objects. Examples include window controls, radio buttons, check boxes, and soft pushbuttons. See also **virtual controls.**

window In the context of computer-human interaction, a rectangular area displayed on a computer screen that provides the user with a set of related data or queries the user for additional input information. A window may cover all or part of the displayed area.

window animation In the context of computer-human interaction, a method for simulating the movement of windows on the computer screen. Allows the user to follow the dynamics of change based on his/her input (e.g., shrinking a window into an icon through several stages, instead of going directly from full-sized window to icon). See **window, computer-human interaction.**

window controls In the context of window management, controls that alter what is displayed on the screen, e.g., close box, size box, zoom box, scroll bars, scroll arrows. See **window, window management.**

window management The coordinated computer-human interface (CHI) process of opening, closing, moving, and resizing windows as well as shrinking displayed pages into icons. Can be excessively time consuming. See also **window, window animation.**

workforce acceptance See **acceptance.**

working memory The currently active contents of consciousness and the processes for conscious manipulation of information. Constrained by capacity limitations to holding and manipulating just a few items under severe time pressure or stress. Also known as short term memory.

workload The combined physical and psychological demands experienced by a worker. The amount of effort required of a human operator by a task. Varies as a function of ability, skill, training, and experience. See also **task demand.**

workstation The set of all job-related items located within an individual's work space (e. g., radar console, communications equipment, input devices, work surfaces, storage, seating).

zoom in/zoom out In the context of computer-human interaction, a capability to focus in on an image or to broaden the scope of the image. In zooming in, more details of a small portion of the image are made discriminable to the user; zooming out gives the user a broader perspective on the situation, but fine-grained details are lost.

REFERENCES

ANSI/HFS. (1988). *American national standard for human factors engineering of visual display terminals* (ANSI/HFS 100-1988). Santa Monica, CA: Human Factors Society.

Avery, L. W., and Bowser, S. E. (Eds.). (1992). *Department of Defense human-computer interface style guide* (Version 2.0). Fort Lewis, WA: Pacific Northwest Laboratory.

Craik, F. I. M., and Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Galitz, W. O. (1993). *User-interface screen design*. New York: Wiley-QED.

Grandjean, E. (1988). *Fitting the task to the man* (4th ed.). New York: Taylor & Francis.

Greenstein, J. S., and Arnaut, L. Y. (1990). Input devices. In M. E. Helander (Ed.), *Handbook of human-computer interaction* (pp. 495-519). New York: North-Holland.

Rasmussen, J. (1983). Skills, rules, and knowledge: signals, signs, and symbols and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13*, 257-267.

Snyder, H. (1990). Image quality. In M. E. Helander (Ed.), *Handbook of human-computer interaction* (pp. 437-474). New York: North-Holland.

Thorell, L. G., and Smith, W. J. (1990). *Using computer color effectively*. Englewood Cliffs, NJ: Prentice Hall.

Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles E. Merrill.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

APPENDIX D-ACRONYMS LIST

2D	two dimensional
3D	three dimensional
4D	four dimensional
AAP	advanced automation program
AAS	advanced automation system
ACID	aircraft identification
ADITS	automated design issue tracking System
AF	airway facilities
AI	artificial intelligence
AI	articulation index
AM	airspace manager
ANOVA	analysis of variance
ANR	active noise reduction
ANSI	American National Standards Institute
ARTCC	air route traffic control center
ASRS	aviation safety reporting system
ATC	air traffic control
ATCS	air traffic control specialist
ATCT	air traffic control tower
ATR	air traffic plans and requirements
ATWIT	air traffic workload input technique
CA	conflict alert
CAMI	Civil Aeromedical Institute
CAR	conformance assessment report
CDR	critical design review
CD-ROM	compact disk-read only memory
CDRL	contract data requirements list
CFF	critical flicker fusion
CGD	control grip device
chi	computer-human interface
CIE	Commission Internationale d'Eclairage
cm	centimeter
COTS	commercial off-the-shelf
cps	cycles per second
CRA	conflict resolution advisory
CRT	cathode ray tube
CSERIAC	Crew Systems Ergonomics Information Analysis Center

CTAS	Center TRACON Automation System
DARC	direct access radar channel
dB	decibels; a logarithmic unit of sound pressure
DCP	design competition phase
df	degrees of freedom
DoD	Department of Defense
DoT	Department of Transportation
DT&E	developmental test and evaluation
ECP	engineering change proposal
FAA	Federal Aviation Administration
FDE	flight data entry (electronic flight strips)
FPL	full performance level
GPWS	Ground Proximity Warning System
HFP	human factors plan
HFS	Human Factors Society (became the Human Factors and Ergonomics Society in 1993)
HUD	head-up display
Hz	hertz
ICAO	International Civil Aviation Organization
ID	identification
IES	Illuminating Engineering Society
in.	inch
IOC	initial operating capability
IPT	integrated product team
IQ	intelligence quotient
ISSS	Initial Sector Suite System
IVSI	instantaneous vertical speed indicator
KDP	key decision point
LCD	liquid crystal display

MIL-STD	military standard
mm	millimeter
mN	milliNewton
mrad	milliradian
MRT	modified rhyme test
msec.	milliseconds
N	newton
NAILS	National Airspace Integrated Logistics Support
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
NCP	draft change proposal
NDI	nondevelopmental item
NIOSH	National Institute of Occupational Health and Safety
nm	nanometer
ORD	operational requirements document
OT&E	operational test and evaluation
oz.	ounce
p	probability
PB	phonetically balanced
PDC	pre-departure clearance
PDR	preliminary design review
PIREP	pilot report
POSWAT	pilot objective/subjective workload assessment technique
PTT	push to talk
PVD	plan view display
RA	resolution advisory
RPD	recognition-primed decision making
RFP	request for proposal
RT	response time
S^2	standard deviation squared
s.d.	standard deviation

SDP	system development phase
SID	site implementation plan
SLS	system level specification
SME	subject-matter expert
SOW	statement of work
SWAP	severe weather avoidance program
SSRVT	sector suite requirements validation team
TA	traffic alert
TCAS	Traffic Alert and Collision Avoidance System
TCCC	tower control computer complex
TEMP	test and evaluation master plan
TIM	technical interchange meeting
TRACON	Terminal Radar Approach Control (Facility)
VCR	video cassette recorder
VSCS	Voice Switching and Control System
WIMPs	windows, icons, menus and pointing devices

INDEX