

DOT/FAA/TC-07/22

Federal Aviation Administration
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

Human Factors Analysis of Safety Alerts in Air Traffic Control

Kenneth Allendoerfer, Human Factors Team – Atlantic City, ATO-P
Ferne Friedman-Berg, Ph.D., Human Factors Team – Atlantic City, ATO-P
Shantanu Pai, Engility Corporation

November 2007

Technical Report

This document is available to the public through the National Technical Information Service (NTIS), Springfield, VA 22161. A copy is retained for reference at the William J. Hughes Technical Center Library.



U.S. Department of Transportation
Federal Aviation Administration

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof. The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report. This document does not constitute Federal Aviation Administration (FAA) certification policy. Consult your local FAA aircraft certification office as to its use.

This report is available at the FAA, William J. Hughes Technical Center's full-text Technical Reports Web site: <http://actlibrary.tc.faa.gov> in Adobe[®] Acrobat[®] portable document format (PDF).

Technical Report Documentation Page

1. Report No. DOT/FAA/TC-07/22		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Human Factors Analysis of Safety Alerts in Air Traffic Control				5. Report Date November 2007	
				6. Performing Organization Code AJP-6110	
7. Author(s) Kenneth Allendoerfer, Human Factors Team – Atlantic City, ATO-P Ferne Friedman-Berg, Ph.D., Human Factors Team – Atlantic City, ATO-P Shantanu Pai, Engility Corporation				8. Performing Organization Report No. DOT/FAA/TC-07/22	
9. Performing Organization Name and Address Federal Aviation Administration Human Factors Team – Atlantic City, ATO-P William J. Hughes Technical Center Atlantic City International Airport, NJ 08405				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address Federal Aviation Administration Human Factors Research and Engineering Group 800 Independence Avenue, S.W. Washington, DC 20591				13. Type of Report and Period Covered Technical Report	
				14. Sponsoring Agency Code ATO-P	
15. Supplementary Notes					
16. Abstract <p>Controllers receive several types of alerts from Air Traffic Control (ATC) automation systems that warn of potentially hazardous situations, including Conflict Alerts, Mode-C Intruder alerts, and Minimum Safe Altitude Warnings. This report provides a human factors analysis of ATC alerts and recommends changes to the alert algorithms and presentation that should increase controller effectiveness and overall system safety. We collected automation data from en route, approach control, and tower facilities that show how often alerts occur, how controllers respond to alerts, and when controller actions occur relative to the alerts. Of all the alerts examined, the majority received no response from controllers; many were so brief that controllers must have resolved the situation prior to the activation or the alert situation resolved itself without action by the controller. Of the alert situations where actions were taken, controllers most often took action <i>before</i> the alert activated. The results suggest that (a) many alerts are valid according to the alert algorithms but do not provide useful information to controllers, (b) these “nuisance alerts” are extremely common in the field, and (c) high nuisance alert rates may desensitize controllers and lead to poor performance. We recommend that the Federal Aviation Administration address the problem of nuisance alerts by improving safety alert algorithms, improving alert presentations, and providing better alert suppression functions. To improve safety alert algorithms, we recommend using data from this study to obtain better measures of critical reaction time parameters for alert algorithms. To reduce the impact of nuisance alerts, we recommend using alert presentations with multiple levels of urgency.</p>					
17. Key Words Air Traffic Control Alarms Alerts Warnings			18. Distribution Statement This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161. A copy is retained for reference at the William J. Hughes Technical Center Library.		
19. Security Classification (of this report) Unclassified		20. Security Classification (of this page) Unclassified		21. No. of Pages 74	22. Price
Form DOT F 1700.7 (8-72)			Reproduction of completed page authorized		

Table of Contents

	Page
Acknowledgments	vii
Executive Summary.....	ix
1. INTRODUCTION.....	1
1.1 Background.....	1
1.1.1 Program Office Support.....	2
1.2 Objectives and Focus.....	2
1.3 Designing Effective Alerts	4
1.3.1 Effective Alerts Accurately Detect Potential Hazards.....	4
1.3.2 Effective Alerts Do Not Activate When Potential Hazards Do Not Exist	4
1.3.3 Effective Alerts Provide New Information.....	4
1.3.4 Effective Alerts are Perceptible and Salient	4
1.3.5 Effective Alerts are Distinguishable From One Another	5
1.3.6 Effective Alerts are Easy to Locate	5
1.3.7 Effective Alerts are Legible.....	6
1.3.8 Effective Alerts Communicate Urgency.....	6
1.3.9 Effective Alerts Communicate Appropriate Actions.....	6
1.3.10 Effective Alerts Draw Attention but Do Not Disrupt Operations	7
1.4 Alerts and Decision Making.....	7
1.4.1 Signal Detection Theory	7
1.4.2 SDT and Time.....	10
1.4.3 Nuisance Alerts.....	12
1.4.4 Automation and Trust.....	15
2. METHOD.....	16
2.1 Initial Field Visits	16
2.2 Data Collection and Reduction.....	17
2.2.1 Facility	17
2.2.2 Request for Data	19
2.2.3 Overall Sample	19
2.2.4 Sector and Time Period Selection.....	21
2.3 Audio Requests.....	22
2.4 Transcription.....	23
2.5 Time Synchronization.....	24
2.6 Response Analysis Sample.....	25
2.7 Response Categorization	28
2.7.1 Control Instructions	28
2.7.2 Advisories and Alerts	29
2.7.3 Visual Separation.....	30
2.7.4 Repeats and Due Diligence.....	30
2.7.5 Coordination with Other Controllers.....	31

3. RESULTS.....	31
3.1 Alert Sample Data.....	31
3.1.1 Operational Errors	31
3.1.2 Duration	32
3.2 Controller Responses to Alerts	33
3.2.1 Time Between Alert and Controller Action.....	37
4. DISCUSSION AND RECOMMENDATIONS	39
4.1 Nuisance Alerts.....	39
4.2 Recommendations to Reduce Nuisance Alerts.....	41
4.2.1 Provide Better Suppression Functionality	41
4.2.2 Develop Graded Alert Presentation	44
4.2.3 Base Alert Times on Human Factors Research	45
4.3 Recommendations to Improve Alert Presentations	46
4.3.1 Collocate Speakers with Displays	46
4.3.2 Restrict Audible Alert Volume Settings.....	47
4.3.3 Improve Salience and Readability of Alert Data blocks	49
4.4 Summary of Recommendations.....	51
5. FUTURE ACTIVITIES AND RESEARCH.....	51
5.1 Causes of Nuisance Alerts	51
5.2 ARTS Algorithm Development.....	52
5.3 Incorporating Intent into Alert Algorithms.....	52
5.4 Examination of Relationship Between Handoffs and Alerts.....	53
5.5 Analysis of Controller Awareness and Responses	56
5.6 User-Centered Design and Evaluation.....	57
References	58
Acronyms	63

List of Illustrations

Figures	Page
Figure 1. Signal detection curves showing a low criterion placement.	9
Figure 2. Signal detection curves showing a high criterion placement.	9
Figure 3. Signal detection curves showing increased sensitivity.....	10
Figure 4. Sequence of project phases and possible future activities.....	16
Figure 5. A typical handoff. The shaded oval indicates a time period in which the automation system believes that Sector B is responsible for the aircraft, but Sector A is actually in communication with the aircraft.....	55
Tables	Page
Table 1. Signal Detection Theory Decision Categories.....	8
Table 2. Signal Detection Theory Applied to Conflict Alerts.....	8
Table 3. Alert Activations by ARTCC.....	20
Table 4. Alert Activations by TRACON.....	20
Table 5. Alert Activations in Requested Sector Periods by ARTCC.....	22
Table 6. Alert Activations in Requested Sector Periods by TRACON.....	22
Table 7. Sample Summary Transcript Compared to a Word-for-Word Transcript for the Same Recording.....	24
Table 8. Alerts Included in the Response Analysis Sample by ARTCC.....	27
Table 9. Alerts Included in the Response Analysis Sample by TRACON.....	27
Table 10. Frequency Distribution of Alert Durations.....	33
Table 11. Controller Responses to Alerts.....	34
Table 12. Controller Responses to Aircraft Involved with CAs.....	35
Table 13. Controller Responses to Aircraft Involved with MSAWs.....	36
Table 14. Time Between CAs and Responses.....	37
Table 15. Time Between MSAWs and Responses.....	38
Table 16. Example of a Graded CA Based on Urgency.....	45
Table 17. Reaction Time Parameters.....	46

Acknowledgments

The Federal Aviation Administration (FAA) Air Traffic Organization – Operations Planning (ATO-P) Human Factors Research and Engineering Group (HFREG) provided funding and management oversight for this project. The authors thank Mike McAnulty, Dino Piccione, and Earl Stein (ATO-P, HFREG) for their guidance throughout the project and for their comments on earlier drafts of this report.

We thank Rose Merchant Bennett and Yen Khahn Vu from the FAA Office of Independent Operational Testing and Evaluation and their staff. In particular, we acknowledge the efforts of Luis Guzman and Terry Griffin (Synoptek Corporation), who conducted the initial reduction and analysis of the automation data. We also thank our industry and academic collaborators, Christopher Wickens and Ken Leiden from Alion Sciences and Stephen Dixon from New Mexico State University, who provided comments on the data analyses contained in this report.

In addition, we acknowledge the efforts of the following people from the FAA and its support contractors who participated in the data collection, reduction, or analysis. Kate McDevitt (Hi-Tec Systems) and Tina DiProspero (Northrop Grumman Corporation) created many of the voice recording transcripts. Tom Dury and Ken Malott, Air Traffic Control Subject Matter Experts from the Indianapolis Air Route Traffic Control Center (ARTCC), were on detail to HFREG and lent us their expertise and technical knowledge by interpreting and helping transcribe the voice communications. Nick Marzelli (BARCO Corporation) created the data analysis and presentation tools for the project. John Dilks and Otto Smith (Engility Corporation) helped to set up and operate the audio equipment. Randy Sollenberger (ATO-P, HFREG) participated as an Engineering Research Psychologist in one of the field visits. Norm Hitchner and Mike Paglione (FAA) provided technical assistance regarding the terminal and en route automation systems. Theresa Payne and Michael Vendetti (FAA Information Technology Office) provided us with the data regarding operational errors. Eric Nadler (Volpe National Transportation System Center) provided information about the audio levels in towers.

Finally, we thank the controllers, managers, automation, and quality assurance staff at each facility included in the data sample; they provided us with the automation and voice recording data for the project. We also thank the staff at the Northern California Terminal Radar Approach Control (TRACON), Las Vegas TRACON and Airport Traffic Control Tower, and Indianapolis ARTCC for hosting us during the field visits.

Executive Summary

Federal Aviation Administration (FAA) controllers receive several types of alerts that warn of potentially hazardous situations, including Conflict Alerts (CAs), Mode-C Intruder (MCI) alerts, and Minimum Safe Altitude Warnings (MSAWs). The alerts are presented visually on the radar display and, in some environments, there is a corresponding audible alert. This report contains a human factors analysis of Air Traffic Control (ATC) alerts, focusing on how and when controllers respond to alerts. It provides recommendations for improving the human factors attributes of the alerts, with the goal of increasing controller effectiveness and safety. The FAA Air Traffic Organization (ATO) Vice Presidents of Terminal Services, En Route and Oceanic Services, and Safety Services requested this research in response to several recent incidents in which controllers seem to have not responded properly to alerts.

The project occurred in two phases. First, researchers visited several ATC facilities where we observed safety alerts during live operations, interviewed personnel about their experiences using alerts, and reviewed data and voice recordings. Second, the researchers examined automation data and voice recordings related to 607 CAs and 178 MSAWs from 5 en route and 17 terminal ATC facilities. They categorized controller responses to alerts, such as issuing traffic advisories and control instructions, and analyzed when those responses occurred relative to alert activation.

The results show that 62% of the CAs examined and 91% of the MSAWs examined in en route received no response from controllers. Similarly, 44% of the CAs examined and 61% of the MSAWs examined in terminal received no response from controllers. However, in none of these cases did an operational error or deviation occur. When controllers did respond to conflict situations, they made the response prior to the alert 67% of the time. For MSAW situations, controllers made the response prior to the MSAW 68% of the time. Furthermore, 31% of the CAs examined in en route and 36% in terminal lasted such a short time that controllers must have resolved the situation prior to the alert or the situation resolved itself without action. These results lead us to estimate that 81-87% of CAs and 87-97% of MSAWs are nuisance alerts or unnecessary; that is, the alerts are valid according to the algorithms but do not provide useful information to the controllers.

A large number of nuisance alerts can create serious human factors problems. By design, alerts cause controllers to interrupt their current tasks and focus attention on the aircraft involved in the alert situation. When these interruptions are frequent and unnecessary, controller workload is increased needlessly and overall performance may be reduced. In addition, a large number of nuisance alerts can desensitize controllers to the alerts, overall, which may lead to poor responses to genuine alerts. Furthermore, a large number of nuisance alerts can reduce controller trust in automation.

The FAA should make reducing nuisance alerts a top priority. One potential method for reducing the number of nuisance alerts is to provide more sophisticated alert suppression functions. For example, we recommend that the alert suppression functions operate like a *snooze* function in which a suppressed alert automatically reactivates when additional criteria are met. A second potential method for reducing the number of nuisance alerts is to base the alert algorithm parameters, such as the look-ahead time, on human factors data, such as how quickly controllers can identify and resolve hazardous situations. A potential method for reducing the impact of nuisance alerts is to provide graded alerts, in which the alert presentation becomes increasingly obvious as the situation becomes more urgent.

1. INTRODUCTION

In Federal Aviation Administration (FAA) Air Traffic Control (ATC) automation systems, controllers receive several types of alerts that warn of potentially hazardous situations. The alerts are presented visually on the radar display and, in some environments, there is a corresponding audible alert. Conflict Alerts (CAs) activate when the flight paths of two controlled aircraft are projected to imminently lose separation. CAs prompt controllers to “evaluate the reason for the alert without delay and take appropriate action” (FAA, 2006b, section 5-14-1). Aircraft separation minima are defined by FAA procedures and vary with domain, altitude, and aircraft type. CAs are intended to provide enough time for controllers to address situations before operational errors (OEs) occur. Mode-C Intruder (MCI) alerts are similar to CAs and warn of imminent conflicts between one controlled and one uncontrolled aircraft. Minimum Safe Altitude Warnings (MSAWs), also known as low altitude alerts, activate when the altitude of an aircraft is projected to be lower than a designated minimum for a geographic area. MSAWs prompt controllers to “immediately analyze the situation and take the appropriate action to resolve the alert” (FAA, 2006b, section 5-14-2). Minimum safe altitudes are defined to provide separation between aircraft and physical obstructions like terrain, buildings, and antennas.

This report provides an analysis of ATC safety alerts from a human factors perspective. The FAA Air Traffic Organization – Operations Planning (ATO-P) Human Factors Research and Engineering Group (HFREG) conducted the analysis in collaboration with other organizations. According to FAA definitions, human factors is a “multidisciplinary effort to generate and compile information about human capabilities and limitations and apply that information to equipment, systems, facilities, procedures, jobs, environments, training, staffing, and personnel management for safe, comfortable, and effective human performance” (FAA, 1995, p. 1). This report focuses on how and when controllers respond to alerts that occur in their facilities, and provides recommendations for improving the human factors attributes of safety alerts with a goal of increasing controller effectiveness and overall system safety.

1.1 Background

Several recent events provided the motivation for this human factors analysis. In June 2006, the FAA ATO Vice Presidents of Terminal Services, En Route and Oceanic Services, and Safety Services addressed a memorandum to the Acting Vice President of ATO-P, requesting a human factors study of ATC safety alerts. In particular, the memorandum asked that the study “consider how and when the alert is presented to the controller and offer solutions for improving this process” (FAA, 2006c). They made this request in response to one mid-air collision and one near miss with an antenna that occurred in early 2006, where the controllers involved seemed to have not seen, heard, or responded properly to the CA or MSAW. In response to this request, HFREG undertook the project described within this report.

In a related development, in July 2006, the National Transportation Safety Board (NTSB) recommended that the FAA “redesign the MSAW and CA systems and alerting methods such that they reliably capture and direct controller attention to potentially hazardous situations detected by the systems” (NTSB, 2006). The NTSB made this recommendation as a result of 11 accident investigations occurring between 2002 and 2006. In each of these cases, the NTSB concluded that either controllers did not see or hear the alert, or they did not take appropriate

action in response to it. Because the NTSB report and the memorandum from the ATO Vice Presidents contained similar findings and made similar recommendations, HFREG accepted the NTSB report as an additional impetus for this project.

Around the same time, ATO Safety Service Office of Independent Operation Testing and Evaluation (IOT&E) began its own analysis of safety alerts resulting from recent safety audits of field sites. IOT&E was concerned that some controllers may not be communicating safety alerts to pilots properly, even in cases where procedures call for them. IOT&E intends its analysis to determine the extent of this issue in the field and intends to develop recommendations as necessary. Because IOT&E and HFREG projects had overlapping goals, the organizations integrated their efforts. This also allowed them to share resources and reduce impact on field sites. Both organizations contributed to the other's activities and reports, and both provided the other with expertise and resources. IOT&E and HFREG expect to continue this collaboration in the future as recommendations are implemented.

1.1.1 Program Office Support

The researchers involved in this project provide occasional human factors support to the Automated Radar Terminal System (ARTS) and Standard Terminal Automation Replacement System (STARS) program offices. During the conduct of this project, both programs requested assistance from HFREG as part of ongoing efforts to improve safety alerts in their systems. These requests demonstrate the high level of interest throughout the FAA on the topic of safety alerts and emphasize the importance of involving human factors when studying this issue.

In February 2006, a case file related to the audible volume of safety alerts was opened for STARS. A case file is a formal problem description that requests a change to National Airspace System (NAS) equipment. This case file requested a national modification to STARS to "provide for a lower perceived volume level for audible alarms in STARS when the Terminal Controller Workstation volume level is set to the minimum value" (FAA, 2006a). The case file was written in response to complaints from the field that the STARS audible alerts are unnecessarily loud and disruptive, even when the volume control is adjusted to its lowest setting. The program office initially determined that STARS was performing within specifications and deferred implementation of the requested modification until human factors recommendations are available on this topic.

During November 2006, engineers from the Common ARTS program contacted HFREG as part of an effort by the program to improve the performance of its alert algorithms. The program identified a need for research-based parameters for use in algorithms and techniques currently under development. In particular, the Common ARTS program is interested in determining the length of time needed for controllers to recognize alerts and take appropriate action. Their goal is to ensure that any potential changes to the alert algorithms provide sufficient time for controllers to respond to the situation.

1.2 Objectives and Focus

The research requirements for this project originated in the ATO memorandum requesting a human factors study of safety alerts (FAA, 2006c). ATO-P HFREG, which includes the Human

Factors Team – Atlantic City, received the task to develop and execute a human factors study to meet these requirements. This study focuses on two human factors aspects of safety alerts (a) how alert algorithms affect controller responses and (b) how alerts are presented to controllers.

First, ATC automation systems use algorithms to determine when a potentially hazardous situation exists and whether to activate an alert. Different algorithms or parameter settings can have different effects on controllers. For example, algorithms that produce a high number of unnecessary alerts can reduce controller trust in the system and can lead to desensitization and poor controller performance. In this study, we examined how controllers respond to the alerts produced by the current safety alert algorithms. We examined how the current algorithms affect decision making, and we recommend features that may improve controller performance.

The main questions we answer in this study regarding the human factors aspects of alert algorithms are as follows.

- What really happens when a CA¹ or MSAW activates? What actions do controllers take in response to alerts? Do they take action immediately or do they wait? How should current algorithms be modified to ensure that controllers select the right responses in a timely manner?
- What portion of the current alerts are useful and necessary and what portion are nuisance alerts? What are the possible effects of nuisance alerts on controller decision making and performance? How should current algorithms be modified to reduce the number or impact of nuisance alerts?

Second, once an algorithm has determined that an alert is necessary, the alert is presented to the controller as visual information and, in some environments, with corresponding auditory information. Because CAs and MSAWs are safety related, the alert presentation should draw the controller's attention quickly and reliably across many different operational situations and physical environments. In the human factors literature, many studies exist that examine techniques for presenting effective alerts in a variety of domains.

The main questions we will answer in this study regarding the human factors aspects of alert presentation are as follows.

- What human factors standards and best practices apply to ATC safety alert presentation?
- What steps should be taken to ensure that alert presentations follow applicable standards and best practices? How should current alert presentation be modified to effectively draw attention and communicate information?

¹ For the remainder of the document, we group MCI alerts with CAs because both involve the potential loss of separation between aircraft. Technically speaking, they are different types of alerts with subtle differences in presentation.

1.3 Designing Effective Alerts

Warnings, alerts, and alarms are widely studied in the human factors domain. This study draws from research in visual and auditory perception, attention and vigilance, and decision making. An extensive literature exists on ways to design and implement effective alerts in many safety-critical domains, including aviation, military, nuclear power generation, and medicine. FAA human factors standards for alerts and alarms can be found in the *Human Factors Design Standard* (Ahlstrom & Longo, 2003). Most of the relevant standards appear in Chapter 7, *Alarms, Audio, and Voice*; with other useful standards in Chapter 3, *Automation*; Chapter 6, *Controls and Visual Indicator*; and Chapter 8, *Computer Human Interface*. In the sections that follow, we discuss attributes of effective alerts.

1.3.1 Effective Alerts Accurately Detect Potential Hazards

The quality of an alerting system is largely defined by how well its sensors and algorithms identify potentially hazardous situations. The best alerting systems reliably activate alerts whenever an action by the user is required to maintain safety. We discuss this topic more thoroughly in section 1.4.1, which examines the issue of hazard detection according to Signal Detection Theory (SDT), which is an important theory of human decision making.

1.3.2 Effective Alerts Do Not Activate When Potential Hazards Do Not Exist

Another critical component of the quality of an alerting system is how well its sensors and algorithms identify when potentially hazardous situations do not exist. The best alerting systems prevent alerts from activating when no action is required to maintain safety. The best systems also quickly deactivate alerts when potential hazards are resolved.

1.3.3 Effective Alerts Provide New Information

Alerts often occur in rich information environments. Users must pay attention to many sources of data simultaneously, as well as monitor, filter, and interpret the data quickly. An effective alert tells the user something that he or she does not already know. If the user already is aware of a potentially hazardous situation, an alert notifying the user about the situation is redundant and can be distracting. Redundant information can be beneficial in some circumstances, but eliminating or reducing redundant information is usually the best design solution. Redundant alerts are especially problematic when the user must take an action to deactivate or silence the alert because workload is increased needlessly.

1.3.4 Effective Alerts are Perceptible and Salient

Alerts are not useful if they cannot be seen or heard. Visual alerts must be large enough and bright enough that a person can determine when the alert is active. Audible alerts must be loud enough that a person can determine when the alert is active. Audible alerts must also be within the frequencies that humans are able to hear.

The perceptibility of alerts can be affected by other factors in the environment. Alerts can be made imperceptible by large distances, physical obstructions, or by other lights or sounds in the environment. For example, a bright blinking alert symbol might be completely perceptible when a controller is sitting in front of the display but imperceptible when a controller views the display from 5 feet away at a 45° angle. An audible alert designed and tested in a quiet laboratory may

be inaudible in a room full of equipment and other people talking. Alerts can also be obscured by other information presented on the same information display, such as a window overlaid on the alert window.

In addition to being perceptible, alerts need to be salient. Broadly defined, salience is the extent to which a stimulus stands out from other stimuli in the environment, thereby drawing attention. For example, a blinking red symbol is salient and draws attention when it occurs in a context of many static green symbols. However, the same blinking red symbol would be less salient if it occurred in a context of many other blinking red symbols. To be most effective, alerts should be so obviously different from other information in the environment that they are unlikely to be missed or mistaken as normal.

1.3.5 Effective Alerts are Distinguishable From One Another

In complex technical environments like ATC, multiple alerting systems are typically present. For example, at controller workstations in Air Route Traffic Control Centers (ARTCCs) and Terminal Radar Approach Control (TRACON) facilities, there may be alerts associated with automation systems, communication systems, and information display systems. In Airport Traffic Control Towers (ATCTs), there may also be alerts associated with weather systems and airport equipment. In all ATC environments, there may be alerts associated with systems at nearby supervisor or traffic management workstations. Furthermore, systems may have maintenance alerts that occur in locations where controllers can see or hear them (Ahlstrom & Panjwani, 2003). Finally, there may be alerts associated with environmental and workplace safety systems like fire alarms.

Human error is more likely to occur when multiple systems use similar indications to signal situations with different levels of urgency or that require different actions. Alerts should not use visual and audible indications that can be easily confused with other alerts in the same environment. For visual alerts, text or graphics (e.g., icons, shapes, colors) can be used to make alerts appear different from each other. For audible alerts, different volumes, frequencies, or periodicities (i.e., the on-off cycle) can be used. Audible alerts can include verbal information, to provide guidance about what action to take or to distinguish one class of alert from another.

1.3.6 Effective Alerts are Easy to Locate

Alerts should allow people to quickly determine where an alert is occurring and to which situation it applies. The alert should draw the user's attention to the location of the information needed to address the situation. In ATC automation systems, a common way to do this is to position the alert indication close to other information that applies to the alert. For example, in ARTS and STARS, the letters "CA" appear in red text on the first line of each data block involved in a potential conflict, and the sectors owning the aircraft receive a corresponding audible alert. This notifies the controllers that a potentially hazardous situation exists, informs them which sectors and aircraft are involved, and draws their attention to the position on screen where the aircraft are located. In ATCTs, where controllers move around and look away from their screens frequently, drawing controllers' attention to the correct location is especially important.

1.3.7 Effective Alerts are Legible

If a visual alert contains text information, users must be able to read and interpret the words. For example, when a CA activates, controllers need to quickly read text information about each aircraft from data blocks or tabular lists, such as altitudes and aircraft types. The presentation of this information should follow human factors guidelines and standards for making text easy to read and interpret, such as by using highly readable font sizes and typefaces. The information should also be in formats that controllers can immediately understand.

Perceptibility and salience do not guarantee legibility. For example, controllers might be able to determine which aircraft is involved in the MSAW because its data block is flashing and has changed color, but the controllers might be unable to read the altitudes because another data block is overlapping.

1.3.8 Effective Alerts Communicate Urgency

When alerts activate, controllers need to determine how to respond. One aspect of this is prioritizing the response among other actions that the controllers are currently taking or planning to take. People can prioritize responses more efficiently if the alert communicates the urgency of the situation. Research in other domains has shown that people respond more quickly and accurately to alerts that communicate urgency than to alerts that do not communicate urgency (Edworthy, Loxley, & Dennis, 1991; Guillaume, Drake, Rivenez, Pellieux, & Chastres, 2002).

People naturally associate certain visual and audible characteristics with situations that are more urgent. Some of these associations are historical and cultural, whereas others are inherent to human perception and psychology (Treisman & Gelade, 1980). For visual alerts, colors such as yellow, orange, and red as well as brighter symbols and faster blink rates are associated with urgency (Christ, 1975; Xing, 2006a). For audible alerts, louder sounds, faster repetition, and higher pitches tend to be associated with urgency (Guillaume et al., 2002). A mismatch between the perceived urgency of an alert and the actual urgency of the situation can increase confusion and lead to mistakes.

In practice, it can be difficult to design alerts that communicate urgency effectively. When multiple systems are located in a single working environment, each system may have been designed and built separately. Designers often decide how salient to make an alert (e.g., how loud should it be?) from the perspective of the individual system rather than from the environment as a whole. This may lead to mismatches between the perceived urgency of the alert and the actual urgency of the situation. This difficulty is magnified when people with different tasks and priorities work in the same environment or use the same equipment. An alert that may be urgent for one category of user may be routine for another.

1.3.9 Effective Alerts Communicate Appropriate Actions

When an alert activates, controllers can typically take several possible actions in response to the alert. When the situation is time critical, the controllers must choose an action as quickly as possible. They could make this choice more quickly if the alert provided information about which actions are appropriate. Some alerts in other environments, such as the Traffic Alert/Collision Avoidance System (TCAS) found in cockpits, explicitly instruct the user what to do in certain very serious situations (e.g., “CLIMB! CLIMB!”). In the case of TCAS, procedures and

training are in place so that pilots know when they are required to follow the instructions and when they have more leeway. In ATC, the appropriate response to an alert depends on the traffic situation, applicable procedures, and the controller's discretion. Some ATC systems, such as the User Request Evaluation Tool (URET), allow controllers to examine several possible resolutions to a situation, but no current FAA system explicitly tells a controller what to do.

In practice, it can be difficult to design alerts that communicate which actions are appropriate. The alerting system typically has less information than the user about the overall situation. The alerting system may have an unsophisticated or incomplete model of the relationships between different situations and actions. The system may be unable to determine an appropriate solution in the time available. If the users come to regard the guidance as inaccurate, not useful, or slow, they may begin to distrust the alerting system and disregard its advice (Wickens, & Dixon, 2005).

1.3.10 Effective Alerts Draw Attention but Do Not Disrupt Operations

By design, alerts draw people's attention away from what they are currently doing. However, except perhaps in extreme situations, alerts should not prevent people from completing what they are working on or cause people to make mistakes on their other tasks. For example, an MSAW visual indication should not cause such dramatic changes in the user interface that controllers are prevented from completing data entries they had started before the MSAW activated (e.g., by placing an alert window on top of the data entry window). In addition, alerts should not make it difficult to accomplish other high priority tasks, especially those that are necessary to respond to the alert. For example, an audible CA should not be so loud that it makes it difficult for controllers to hear radio transmissions, because air-ground communications are necessary to respond to the CA. Furthermore, alerts should not disrupt the tasks of other people who are working in the same environment, using the same system, or who are unaffected by the situation. For example, an audible CA occurring for the sector at the end of a row of workstations should not be so loud that it distracts the controllers working at the sectors located in the middle of the row. Finally, and obviously, visual or audible alerts should not create workplace safety hazards (e.g., no alert should be so loud that it could damage people's hearing).

1.4 Alerts and Decision Making

Safety alerts are intended to draw controllers' attention to potentially hazardous situations, especially those they may have overlooked or underestimated. Alerts are part of a human-machine system, and changes to either can affect the success of the overall system. How controllers respond to alerts is affected by aspects of the traffic situation, the visual and auditory characteristics of the alert, the environment in which the alert occurs, and the controllers' own abilities and expectations.

1.4.1 Signal Detection Theory

Signal Detection Theory is a well-known theoretical framework for understanding how people make decisions under uncertain and noisy conditions (Sorkin & Woods, 1985). SDT was originally developed for use in human perceptual research but has been applied to a variety of domains by cognitive psychologists and human factors researchers. Because we discuss SDT concepts and techniques throughout this report, this section provides an introduction to SDT.

In the SDT framework, a person is asked to look for specific information, such as a tone of a certain pitch, known as the signal. Other information that the person may encounter, such as tones at other pitches, are considered noise. People are imperfect detectors of signals, especially when the difference between the signal and the noise is small or subtle. People will occasionally misperceive signal as noise and vice versa. The ability of people to reliably distinguish signal from noise, known as sensitivity, depends on training, experience, and some inherent physical and cognitive abilities. In addition, the magnitude of the difference between the signal and noise influences the outcome. If the signal is only subtly different from the noise, people must be very sensitive to detect it reliably. Once a person makes a decision, it can be classified as a hit, miss, false alarm, or correct rejection (see Table 1). Alerting systems have been considered in SDT terms in a variety of domains (Sorkin & Woods, 1985). Table 2 provides an application of SDT decision categories to CAs.

Table 1. Signal Detection Theory Decision Categories

	Person Responds "No"	Person Responds "Yes"
Signal Present	Miss	Hit
Signal Absent	Correct Rejection	False Alarm

Table 2. Signal Detection Theory Applied to Conflict Alerts

	Alert Does Not Activate	Alert Activates
Imminent Conflict Present (i.e., if no action is taken, a loss of separation will occur)	Miss	Hit
Imminent Conflict Absent (i.e., if no action is taken, a loss of separation will not occur)	Correct Rejection	False Alarm

The overall goal for alerting systems, in SDT terms, is to achieve a high hit rate while maintaining a low false alarm rate. People set an internal criterion for how much evidence must be detected before they conclude that a signal is present or absent. The criterion affects all four decision categories. For example, if people set their criterion low when signal and noise are hard to tell apart, they will have a high rate of hits but also a high rate of false alarms (see Figure 1). If people set their criterion high, they will have a low rate of false alarms but also a low rate of hits (see Figure 2).

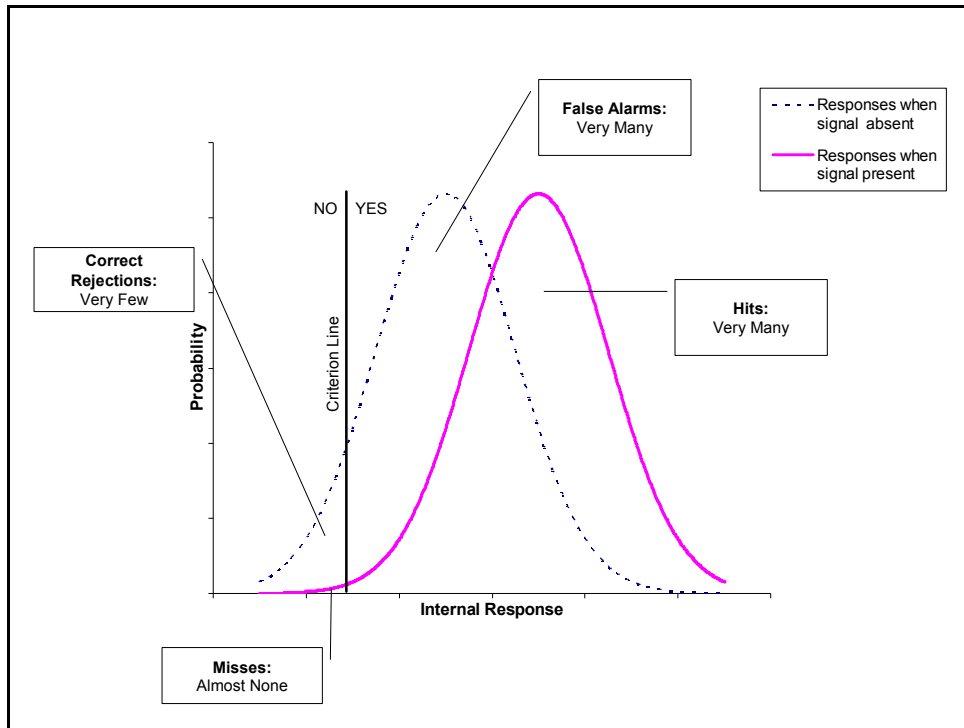


Figure 1. Signal detection curves showing a low criterion placement.

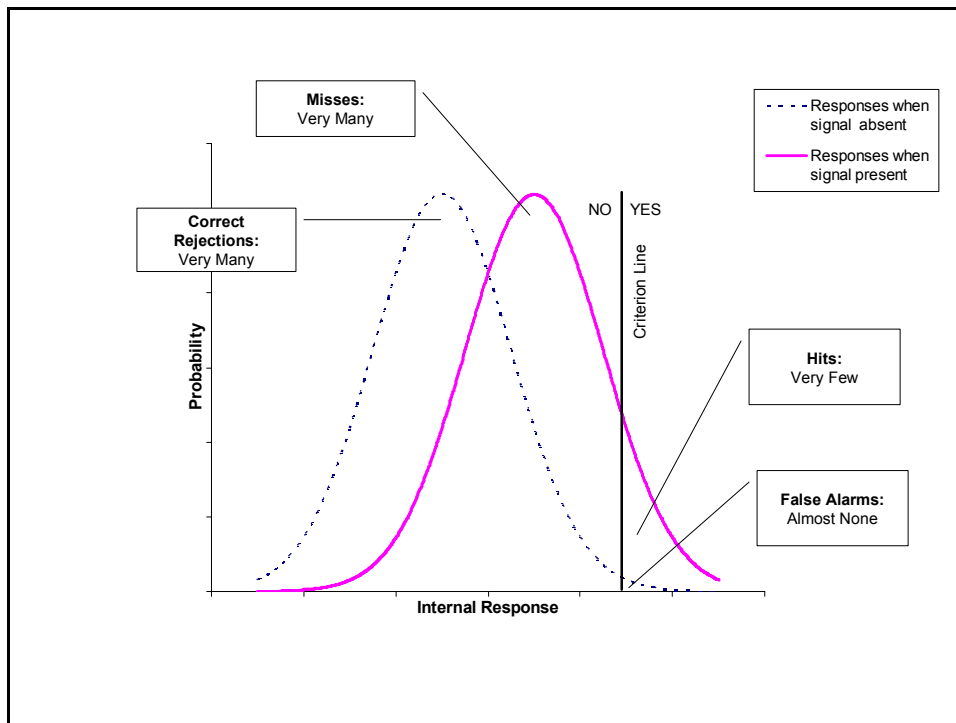


Figure 2. Signal detection curves showing a high criterion placement.

People set the criterion based on the perceived costs and benefits of each decision category. If the cost of a miss is high and the cost of a false alarm is low, people normally set a low criterion. If the cost of a miss is low and the cost of a false alarm is high, people normally set a high criterion. To improve the hit rate without also increasing the false alarm rate, people must become more sensitive to the signal (see Figure 3). When sensitivity increases, the curves become farther apart and overlap less. Sensitivity can be increased through training or experience, or the signal must be made more distinguishable from noise by making the signal louder, a different color, and so on.

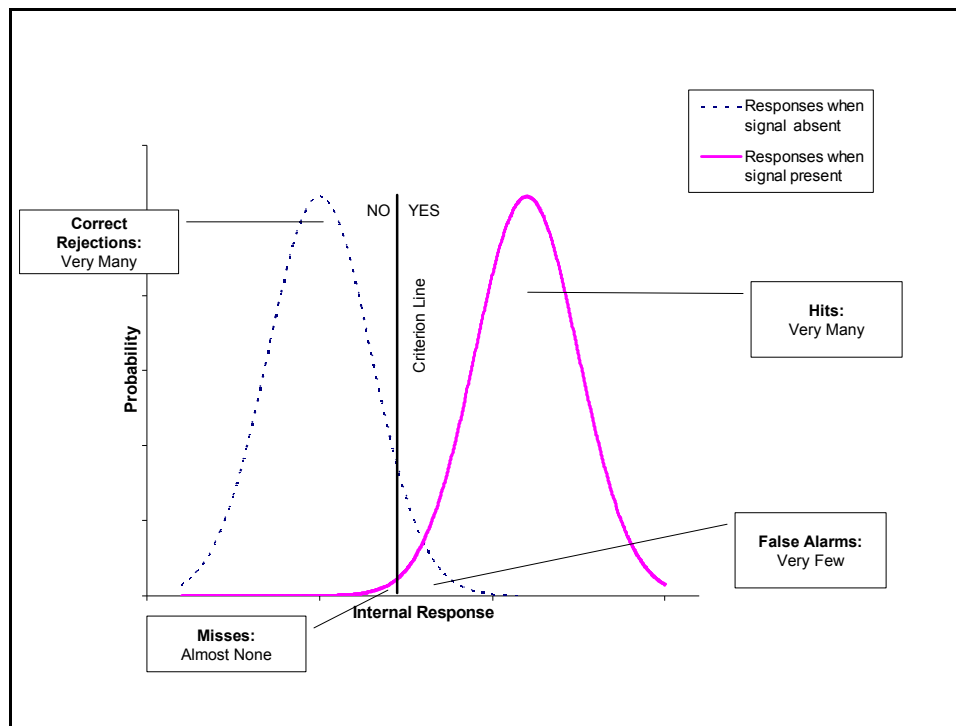


Figure 3. Signal detection curves showing increased sensitivity.

Unlike humans trying to detect signals, the sensitivity and criterion placement for an alerting system can be defined and manipulated by the designers. In ATC, the cost of a miss (e.g., the aircraft should have turned, but the turn was not issued and the aircraft lost separation or collided) is extremely high for everyone involved. The cost of a false alarm (e.g., the controller descended an aircraft when it actually did not need to descend) is not zero but is much lower than a miss. For this reason, ATC safety alerts are conservative and maintain a relatively low criterion for activation. That is, an alert activates when the system collects a relatively small amount of evidence that a potential hazard exists. This corresponds to an SDT curve similar to Figure 1, which shows both high hit and high false alarm rates.

1.4.2 SDT and Time

Despite their theoretical and explanatory value, the SDT categories do not always apply cleanly to real world applications. In particular, alerts are often intended to provide people with time to react to and possibly prevent a hazardous situation from occurring. In these cases, the timeliness of the alert is critical and must be considered when evaluating the effectiveness of the system.

All ATC safety alerts incorporate time parameters that are intended to activate the alert *before* a loss of separation occurs and provide controllers with enough time to prevent an OE from occurring.

The actual ATC safety alert algorithms are numerous and extremely complex. Conceptually, each algorithm contains the following or equivalent steps.

- Define a distance between two aircraft or between an aircraft and the obstruction that is necessary for safety, $D_{\text{separation}}$.
- Define a time that is needed to address most safety critical situations, T_{critical} and define a look ahead time, $T_{\text{lookahead}}$, that is equal to or greater than T_{critical} . This gives controllers and pilots sufficient time to respond in case the situation is actually hazardous.
- Project the future positions of each aircraft from the current time, T_{current} , into the future until $T_{\text{lookahead}}$.
- Calculate distance between the aircraft or between the aircraft and the obstruction for each projected position, $D_{\text{projected}}$.
- Activate CA or MSAW if $D_{\text{projected}} \leq D_{\text{separation}}$ for any projected position.

The algorithms also incorporate logic to account for noise in the surveillance data and for the effects of active maneuvers on the position projections. Incorporating time parameters changes the SDT categories somewhat. For example, a potential loss of separation might exist and CA might activate (i.e., an SDT hit), but too late to prevent the loss. How should we judge the performance of the alert in this case? It is successful from the perspective of detecting a hazardous situation, but it is unsuccessful from the perspective of helping controllers prevent an OE. Even so, the alert did activate in time to prevent a collision, and contributed positively to safety. We call this category Hit But Late (HBL). A large number of HBLs is undesirable, despite the high hit rate, and would signal that the alerting system needs improvement.

On the other hand, some HBLs are forgivable. For example, suppose two aircraft are perfectly separated and landing at closely spaced parallel runways. If one aircraft suddenly turns 20 degrees to the left, it would cause a loss of separation almost immediately and a corresponding CA would activate. This situation would be counted as a hit because a potentially hazardous situation did exist and the alert activated appropriately, but it also would be considered an HBL because the alert did not provide enough time for the controller to react. In this case, however, it would be difficult to fault the algorithm or the controller for not predicting the unplanned maneuver.

Furthermore, trying to prevent HBLs can affect the other decision categories. Suppose a potential conflict does not exist, but the system lacks certain critical information at the moment the decision to alert is made. For example, two aircraft are heading toward the same fix in opposite directions. The algorithm projects their positions and determines that they will lose separation within the look-ahead time, so it activates a CA accordingly. However, unknown to the system, one aircraft is actually planning to turn 20 degrees to the left in 12 seconds. Because the algorithm does not have access to the flight plan and does not project a four-dimensional trajectory, a CA activates for a situation that quickly resolves itself without further action. In this case, the false alert

occurred only because the time parameter forced the algorithm to make a decision before all the information was known. If the system were able to wait a little longer, the false alert could have been avoided. We call these situations False Alert But Early (FABE).

Like traditional SDT hits and false alarms, HBLs and FABEs trade off each other. A longer look-ahead time will yield fewer HBLs but also more FABEs. A shorter look-ahead time will yield fewer FABEs but also more HBLs. Choosing the right look-ahead time is essential to the effectiveness of the alert algorithm. The look-ahead time must allow sufficient time for each of the following actions.

1. Controllers react to the alert (e.g., shift attention from their current task to the appropriate screen).
2. Controllers obtain necessary information (e.g., read the altitudes for the affected aircraft from the data blocks).
3. Controllers decide what action to take.
4. Controllers communicate instructions to pilots.
5. Pilots hear and acknowledge the instructions.
6. Pilots implement the instructions (if necessary).
7. Aircraft execute the maneuver (if necessary).

Human factors research can provide parameter values for items 1 through 6 based on field studies, simulations, laboratory experiments, and research from other domains available in the literature.

1.4.3 Nuisance Alerts

Alerts should draw attention to situations that controllers may have overlooked or misinterpreted. The sequence of events in a model alert situation is as follows:

1. The controller is preoccupied with other operational tasks and is unaware that a potentially hazardous situation is developing or underestimates its severity or urgency.
2. The situation develops beyond a critical point and the alert activates.
3. The alert focuses the controller's attention on the situation. The controller immediately begins deciding how to respond, such as issuing a traffic or altitude advisory, coordinating with another sector, or giving a control instruction over the radio.
4. The controller decides how to respond, takes the necessary actions, and the pilot responds appropriately.
5. The situation is resolved, and the alert deactivates.

However, only a portion of real-world alert situations follow this model. In some cases, controllers are already aware of the situation when the alert activates, or the alert activates but no controller action is necessary and the situation resolves itself. These are examples of **nuisance alerts** (Sanquist, Thurman, & Mahy, 2005). A nuisance alert can fall into any of the following categories.

- **No Action Necessary** – An alert that activates for a situation where no OE would actually occur, even if no action is taken. Such cases would be categorized as false alarms in SDT and include FABEs.
- **Already Addressing It** – An alert that activates for a situation where the controller is already working to resolve the situation but the resolution has not completely played out. For example, a controller sees a potential altitude hazard, formulates a plan for action, begins communicating with the pilot, and then the MSAW activates. The MSAW is valid in the sense that a potentially hazardous situation exists, but it is also a nuisance in that the controller already knew about the situation, and no action is required beyond what the controller is already doing. If an alert literally activates while the controller is taking the action (e.g., communicating with the pilot), it can be distracting and cause delay or error in the completion of the action (Ahlstrom, 2003).
- **Somebody Else’s Problem** – An alert that activates in a location or for an aircraft that has no operational impact on the sector in question. This would be considered a hit in SDT terms, but, from the perspective of the unaffected controller, it is a nuisance. Controllers who are not affected by the situation can be distracted by flashing data blocks, loud audible alerts, and potentially having to acknowledge or silence the alert.
- **Obnoxious** – An alert that is more salient (e.g., brighter or louder) or that lasts longer than necessary to draw attention. For example, if controller attention is completely drawn within 5 seconds but the audible alert lasts for 10 seconds, controllers may come to regard the alert as a nuisance. This category also includes cases where the perceived urgency of the alert does not match the actual urgency of the situation. In the 1990s, some of the negative effects of obnoxious alerts were found in FAA facilities. An accident investigation revealed that someone at a facility had muted the MSAW aural speaker using paper and masking tape (NTSB, 2000). Presumably, the people responsible found the MSAW aural alert to be so obnoxious and disruptive that they were willing to risk ATC safety, not to mention their jobs, in an effort to make the alert quieter. This finding prompted the NTSB to issue safety recommendations regarding alert speakers and the FAA ordered facility management to inspect alert speakers to ensure that they were working properly and were not being muted.
- **Using Other Types of Separation** – An alert that sounds for a potential conflict where the controller is using a type of separation other than radar separation, such as visual or diverging courses separation. Current ATC systems cannot independently determine when controllers are applying these types of separation. In some cases, controllers may manually record this information, but may not do so every time or may do so only on a paper flight progress strip.
- **Repeat** – An alert that activates when it has already activated and the controller has already taken action or decided that no action is necessary. This often occurs with borderline alert situations. For example, due to noise in the radar data, an alert may activate for a few radar updates. The controller may act on the alert or may suppress it. If the controller does not suppress it, the alert might reactivate later.
- **Surveillance or Tracking Error** – The terms “nuisance alert” and “false alert” are not interchangeable in ATC. Occasionally, errors in surveillance or tracking can cause the

automation system to depict a target where none exists or two targets for one aircraft. In ATC terminology, a false alert is one that occurs between one real target and one false one. These are false alarms in traditional SDT, and they are also nuisances because they draw attention but require no action by the controller.

Nuisance alerts are troubling because they cause people to stop whatever they are doing and attend to the alert to determine whether the alert is informative (Sarter, 2005; Xiao & Seagull, 1999). There is also a resumption lag after the interruption, defined as the time it takes to gather one's thoughts and resume the task at hand (Altmann & Traflet, 2002). Because of this effect, nuisance alerts can be extremely disruptive to primary task performance and, when they occur too frequently, can lead to desensitization. A familiar example of desensitization is when people learn to ignore fire alarms when the alarm goes off too frequently or for no apparent reason.

When people experience frequent false or low-urgency alerts, they tend to respond less quickly and accurately to real and high-urgency alerts (Getty, Swets, Pickett, & Gonthier, 1995; Maltz & Meyer, 2001; Parasuraman & Riley, 1997). When there is a high incidence of false or nuisance alerts, people may suppress an alarm before they determine its actual status or may no longer treat the alert as mandatory. In both cases, their overall alarm compliance decreases (Meyer, 2001; Parasuraman & Hancock, 1999; Watson, Sanderson, & Anderson, 2000). Researchers have even related the high rate of nuisance alerts in critical care settings to nurse burnout (Topf & Dillon, 1988). In the ATC domain, a high rate of nuisance alerts can create other problems because frequent and irrelevant auditory interruptions can potentially disrupt visual task performance, known as the irrelevant sound effect (Jones, 1999; Watson, Sanderson, et al., 2000).

Xiao, Seagull, Nieves-Khouw, Barczak, and Perkins (2004) examined why people fail to respond to alarms in a health care environment. By interviewing hospital personnel, they determined that there were often so many alarms in the environment that staff stopped responding to every alert, although they sometimes missed critical alerts. Research has shown that in operating rooms, alarms can sound as often as every 4.5 minutes and up to 75 % of these alarms may be false alarms or nuisance alerts (Kesting, Miller, & Lockhart, 1988; as cited in Seagull & Sanderson, 2001). Watson, Russell, and Sanderson (2000) found that operating room anesthesiologists responded or made adjustments to only 3.4% of operating room alarms. Even when hospitals implemented programs to increase response rates, they often did nothing to address the problem of too many nuisance and false alarms. Xiao et al. noted that to increase the response rate, it is not sufficient to simply increase the perceptibility of alerts. It is just as important to address why so many unnecessary alerts occur and eliminate as many as possible.

Zabyshny and Ragland (2003) performed a literature review of the effect of false alarms on the use and acceptance of collision detection systems in automobiles. They noted that many times designers engineer systems to minimize misses, which results in a corresponding increase in false alarms, especially when the base rate of a dangerous event (in this case, a collision) is low. Even for highly reliable systems, if the base rate of an event is low, the posterior probability of an alarm indicating the actual occurrence of an event will be low (Parasuraman & Hancock, 1999). In addition, if there are events that appear to the system to be potentially hazardous but actually are not, this will increase the nuisance alert rate and decrease the posterior probability that an alarm signals an actual dangerous event (Sanquist et al., 2005). The research indicates

that high false alarm rates lead to an increase in driver response times and can lead people to ignore alarms. In some situations, high false alarm rates caused drivers to increase their risk-taking behavior. Zabyshny and Ragland argued that to design more effective warning systems, engineers should redefine what the system identifies as a dangerous event (e.g., use loss of separation rather than collision as the dangerous event) so as to increase the posterior probability of valid alarms (see also Parasuraman & Hancock).

Audible alerts seem especially prone to being nuisances. They are certainly effective at drawing attention, and they can serve as an interruption cue that allows users to perform some preattentive processing of alerts (Sarter, 2005). Evidence shows that people demonstrate a reduction in reaction time when an auditory signal accompanies a visual alert (Stokes & Wickens, 1988). However, audible alarm designers potentially diminish the benefits obtained by using audible alerts when a system has a high rate of irrelevant or uninformative nuisance alerts. Audible alerts are also the source of complaints from ATC field sites (FAA, 2006a).

1.4.4 Automation and Trust

Even well designed and effective alerting systems can cause human factors problems. In particular, if an alerting system is generally accurate, users will naturally come to trust and rely on it (Parasuraman & Riley, 1997). In some cases, trust can lead to over reliance on the automation, which can lead to declines in attention, vigilance, or problem-detection ability. In such cases, the user might be unprepared to take over when the automation system fails or might have difficulty detecting or resolving situations that the automation cannot handle. We do not believe that controllers currently over rely on safety alerts, but as alerting algorithms and related automation become more sophisticated and prevalent, over reliance may become an issue.

In addition, there are actually two detectors of potentially hazardous situations operating simultaneously: the alerting system and the controller (Sorkin & Woods, 1985). If the controller and the alert reach different conclusions for the same situation, the controller may learn to distrust the system because it *cries wolf* too often. Research in other domains has found that when the reliability of automation is less than .70, users who rely on it would actually achieve better overall performance if the automation were removed and users did all the work themselves (Wickens & Dixon, 2005).

Controller responses to alerts can be understood in terms of probabilities. When a controller recognizes an alert, he or she decides whether to respond based on factors such as the following:

- What is going on with the traffic? Is this situation serious enough to require a response from me now or can I wait for the situation to develop further?
- Has the alert been reliable for situations like this in the past? Is the current alert valid (in which case, I should act) or is it a nuisance (in which case, I should suppress it)?
- Do I have time to respond? What are the consequences if I make the wrong decision about whether to respond?

Errors in decision making can occur at any of these points. The controller may misinterpret the situation. The controller's level of trust in the automation may lead the controller to misjudge the past performance of the alert. The controller may misjudge the costs of different decisions. Other psychosocial factors like confidence and personality can also play a role.

2. METHOD

The current project occurred in multiple phases (see Figure 4). We also show possible future activities that could occur as a second project, in which we provide human factors support to projects seeking to design, test, and implement changes to the alert algorithms or presentations.

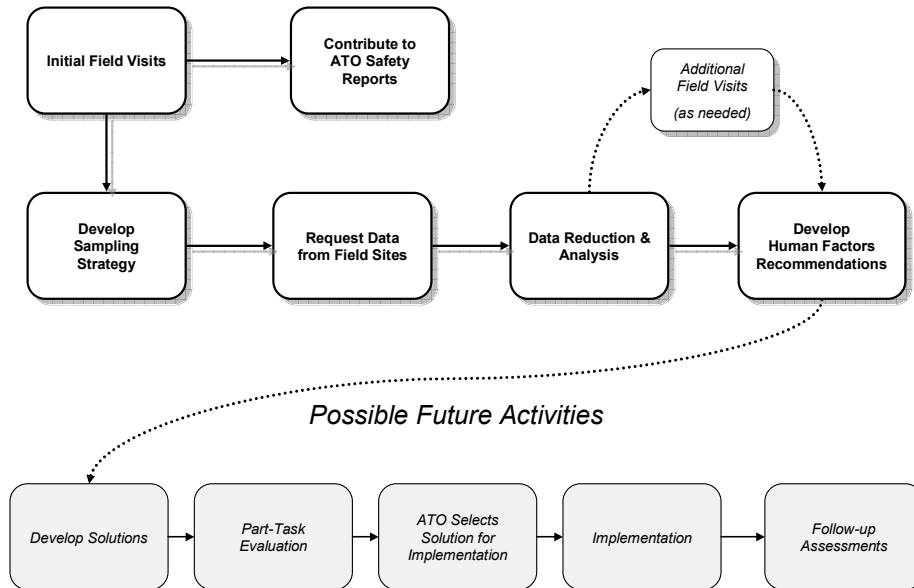


Figure 4. Sequence of project phases and possible future activities.

2.1 Initial Field Visits

In collaboration with IOT&E, Engineering Research Psychologists (ERPs) from HFREG Human Factors Team – Atlantic City visited one ARTCC, two TRACONS, and one ATCT. The initial visits served to increase our understanding of the real-world use of safety alerts, including the number of alerts that controllers typically experience during a shift, the urgency and duration of the alerts, and the common work practices that controllers have adopted to handle alerts.

During the visits, we observed operational controller positions and listened to live communications. We chose the sectors and time periods to observe based on discussions and information provided by Quality Assurance (QA) personnel, Air Traffic managers or supervisors, and automation specialists. In general, we selected sectors to observe where many alerts were known to occur and where traffic volumes and complexity were high.

We also discussed alerts with these personnel to obtain their opinions and learn about their experiences using safety alerts. When operational conditions allowed, we informally discussed safety alerts with the controllers working near our observation positions. Working with automation specialists, we collected automation and voice communication recordings that allowed us to review selected alert situations. These data are included in the main sample discussed later.

2.2 Data Collection and Reduction

The primary data for this report were derived from recordings provided by the automation and communication systems at each facility. The automation data include the activation time of the alert, the call signs and locations of the aircraft involved, and derived values such as the projected closest point of approach and the closing speed. Each automation system reports these data using its own format that first had to be parsed and reorganized.

The communication data includes the radio transmissions between controllers and pilots and the phone or intercom calls between controllers. The communication data also can include a time synch signal, which is useful for determining precisely when communications occurred. To determine whether an alert is a nuisance, an analyst must observe the alert live or review recordings later. To analyze audio recordings, analysts must first transcribe or summarize the communications by listening to the recordings. This is a labor-intensive process.

Because of the large number of alerts, the complexity of the data analysis, and our desire to include a range of facility and traffic characteristics, we determined that it was not feasible to conduct a detailed analysis of every alert. Instead, we sampled judiciously and focused our efforts on the areas where human factors problems were most likely to exist. The following sections describe our sampling strategy.

2.2.1 Facility

In collaboration with IOT&E and local Subject Matter Experts (SMEs), we selected the following facilities to include in the initial sample.

Eastern Service Area

- Atlanta Large TRACON (A80)
- Potomac Consolidated TRACON (PCT)
- Boston Consolidated TRACON (F90)
- Philadelphia (PHL) TRACON/ATCT
- Orlando (MCO) TRACON/ATCT
- Louisville Standiford (SDF) TRACON/ATCT (SDF)
- Burlington Vermont (BTV) TRACON/ATCT
- Atlanta ARTCC (ZTL)
- Washington ARTCC (ZDC)

Central Service Area

- Dallas-Ft. Worth TRACON (D10)
- Detroit TRACON (D21)/ATCT
- Indianapolis (IND) TRACON/ATCT
- Houston Hobby TRACON (I90)/ATCT
- Minneapolis-St. Paul TRACON (M98)/ATCT
- Oklahoma City (OKC) TRACON/ATCT
- Houston ARTCC (ZHU)
- Indianapolis ARTCC (ZID)

Western Service Area

- Denver TRACON (D01)/ATCT
- Southern California TRACON (SCT)
- Northern California TRACON (NCT)
- Salt Lake City TRACON (S56)/ATCT
- Phoenix TRACON (P50)/ATCT
- Las Vegas TRACON (L30)/ATCT
- Los Angeles ARTCC (ZLA)
- Salt Lake City ARTCC (ZLC)

The goal in selecting these facilities was to sample a broad range of facilities that would allow us to generalize conclusions across the NAS. We selected facilities based on the following considerations.

- Similar number of facilities from each service area.
- Different sizes and traffic volumes.
 - Example: ZTL was the busiest ARTCC in 2006; ZLC was 19 out of 21 (FAA, 2007a).
 - Example: SCT was the busiest TRACON in 2006; SDF TRACON was 47.
- Different automation systems and configurations.
 - Example: PHL TRACON is a large STARS site with one pacing airport and one satellite; NCT is a large Common ARTS site with several large airports and many small ones.
- Facility identified in the ATO memorandum or in the NTSB report.
 - Example: NCT, BVT, SCT, MCO.
- Mountainous terrain where MSAWs may be more common.
 - Example: S56, D01, NCT, BTV.
- Unusual equipment or configurations.
 - Example: SDF TRACON has an ARTS IIIIE system with ARTS Color Displays (ACDs). This configuration is normally found only in consolidated TRACONs or TRACONs associated with large airports.
- Previous positive experience obtaining data from the facility or connections with facility personnel that would facilitate obtaining data.

As the project progressed, we added more facilities to the sample. In particular, in reviewing the automation data, we found that a substantial number of alerts occurred in ATCT airspace or involved one aircraft in the TRACON and one in the ATCT. For these reasons, we added the following smaller ATCTs associated with larger TRACONs.

Additional Satellite ATCTs

- Dekalb-Peachtree Airport (PDK) ATCT (associated with A80)
- Dallas Love Field Airport (DAL) ATCT (associated with D10)
- Centennial Airport (APA) ATCT (associated with D01)
- Bowman Field Airport (LOU) ATCT (associated with SDF TRACON)
- Flying Cloud Airport (FCM) ATCT (associated with M98)
- St. Paul Downtown Airport/Holman Field (STP) ATCT (associated with M98)
- Northeast Philadelphia Airport (PNE) ATCT (associated with PHL TRACON)

2.2.2 Request for Data

Personnel from IOT&E sent a memorandum to the directors of each service area requesting automation data from the facilities in that area. From ARTS and STARS facilities, we requested the following information for 7 consecutive days:

- reports from the ARTS Monitor (SMON) or STARS Monitor & Control Workstation system logs that list general information about each alert occurring in the facility;
- reports from the Continuous Data Recording (CDR) that provide more detailed information about each alert from the busiest 2 hours each day of the sample; and
- site adaptation data that contain the locations of all radars and airports in the TRACON airspace.

From ARTCCs, we requested the following information for 7 consecutive days:

- reports from the Host Computer System (HCS) System Analysis Recording that provide detailed information about each CA, MCI, or MSAW from the busiest 2 hours each day of the sample; and
- site adaptation data that contain the locations of all radars and airports in the ARTCC airspace.

The request for data also indicated that we would be requesting audio recordings of selected sectors after we reviewed the automation data. They provided these data to analysts from IOT&E, who executed initial data reduction and analysis procedures that computed the number of alerts of each type occurring during each hour of the day and from each sensor.

2.2.3 Overall Sample

Twenty-two of the 25 facilities provided the requested automation data. ZDC, I90, and P50 did not provide automation data because of technical problems or, in the case of P50, the automation personnel were fully occupied at the time of the request transitioning the facility from ARTS to STARS and moving to a new building. We used these data to determine overall alert counts for each facility based on alert activations. The methods for determining when an alert activates are conceptually the same for each automation system, but each requires a different reduction technique based on the format of the data. The overall method is to find the first alert that was displayed to controllers for each set of CA or MSAW messages. A set of messages is an unbroken series of messages. If an aircraft or pair of aircraft goes into alert status, comes out of alert status, and then goes back into alert status, two sets of messages are created and are counted as two alert activations.

The first round of data reduction produced the sample of alerts listed in Tables 3 and 4. These numbers represent the total number of alert activations occurring at each facility during the sampled time periods. The totals for the ARTCCs include alerts occurring during only the 2 hours each day with the heaviest traffic volume. The totals for the TRACONs include alerts from the entire day.

Table 3. Alert Activations by ARTCC

ARTCC	CA (hours)	MSAW (hours)
Atlanta	4911 (14)	798 (14)
Houston	1212 (12)	828 (14)
Indianapolis	1468 (14)	667 (14)
Los Angeles	5269 (14)	1372 (14)
Salt Lake City	2108 (14)	211 (14)
Total	14968 (68)	3876 (70)

Table 4. Alert Activations by TRACON

TRACON	CA (hours)	MSAW (hours)
Atlanta	1538 (168)	1161 (168)
Boston	569 (120)	136 (120)
Burlington	24 (168)	43 (168)
Dallas-Fort Worth	1525 (168)	580 (168)
Denver	286 (168)	403 (168)
Detroit	213 (168)	91 (168)
Indianapolis	78 (216)	515 (216)
Las Vegas	2246 (168)	369 (168)
Louisville	259 (144)	164 (144)
Minneapolis	256 (168)	156 (168)
Northern California	2672 (168)	750 (168)
Oklahoma City	123 (192)	695 (192)
Orlando	1783 (168)	262 (168)
Philadelphia	317 (168)	304 (168)
Potomac	1915 (168)	718 (168)
Salt Lake City	540 (168)	35 (168)
Southern California	7393 (168)	1486 (168)
Total	21737 (2856)	7868 (2856)

2.2.4 Sector and Time Period Selection

It was clearly not feasible to analyze 36,705 CAs and 11,744 MSAWS in detail. We reduced the original sample by selecting individual sectors and time periods for analysis. Hereafter, the term “sector period” refers to a single combination of facility, sector, date, and time. For example, we selected the following sector period:

Facility: ZTL
Sector: Sector 18
Date: January 24, 2007
Time: 2100 to 2200 Coordinated Universal Time (UTC)

First, we identified one primary CA and one primary MSAW sector period for each facility using the following process:

1. We determined the day in the sample with the highest number of CAs or MSAWs,
2. We selected the day and determined the sector where the highest number of CAs or MSAWs occurred.
3. We determined the hour of that day where the highest number of CAs or MSAWs occurred in that sector and chose that as the sector period. In the analysis, we also included any CAs that occurred in the primary MSAW sector period and vice versa.

Second, we also identified one or two miscellaneous sector periods that were interesting or notable but did not have enough alerts to be the primary CA or MSAW sector period. In many cases, the miscellaneous sector periods had the second or third highest number of CAs or MSAWs for that facility. In other cases, the miscellaneous sector periods showed an unusual pattern, such as all the alerts occurring in a straight line when plotted on an airspace map. In other cases, QA personnel or supervisors identified sectors with complex traffic patterns or sectors where they believed safety alerts commonly occur.

Third, because CAs can occur between aircraft in two different sectors, we also typically requested one or more sector periods adjacent to the primary sector periods. It is often too difficult to determine which sector an aircraft was communicating with at the time of the alert. In particular, when handing off an aircraft from the Final sector in the TRACON to the Local sector in the ATCT, controllers do not typically make an accompanying entry in the ARTS or STARS. In these cases, when using automation data alone, we cannot determine which sector is communicating with the aircraft at the time of the alert. Therefore, we requested voice recordings from the Local sectors corresponding to the Final sectors we selected. In addition, the reports from SMON in ARTS did not provide information about which sector was controlling the aircraft at the time of the alert. For the ARTS facilities, we had to estimate which sectors were communicating with the aircraft based on the aircraft positions. In these cases, we requested voice recordings for adjacent sectors to increase the chances that we obtained audio recordings for the appropriate sectors.

The total number of sector periods, including adjacent sectors, requested from a single facility ranged from 3 (at numerous facilities) to 8 (at A80). Selecting the sector periods and filtering the alert data appropriately produced alert samples (see Tables 5 and 6).

Table 5. Alert Activations in Requested Sector Periods by ARTCC

ARTCC	CA (hours)	MSAW (hours)
Atlanta	93 (3)	20 (3)
Houston	29 (3)	7 (3)
Indianapolis	6 (3)	28 (3)
Los Angeles	104 (3)	7 (3)
Salt Lake City	67 (4)	21 (4)
Total	299 (16)	83 (16)

Table 6. Alert Activations in Requested Sector Periods by TRACON

TRACON	CA (hours)	MSAW (hours)
Atlanta	85 (3)	53 (3)
Boston	68 (3)	4 (3)
Burlington	4 (3)	14 (3)
Dallas-Fort Worth	104 (3)	55 (3)
Denver	29 (3)	29 (3)
Detroit	16 (3)	5 (3)
Indianapolis	3 (3)	11 (3)
Las Vegas	47 (1)	0 (1)
Louisville	52 (3)	18 (3)
Minneapolis	17 (3)	21 (3)
Northern California	145 (3)	57 (3)
Oklahoma City	2 (3)	12 (3)
Orlando	56 (3)	5 (3)
Philadelphia	6 (3)	6 (3)
Potomac	108 (3)	28 (3)
Salt Lake City	28 (3)	2 (3)
Southern California	504 (3)	75 (3)
Total	1274 (49)	395 (49)

2.3 Audio Requests

We requested audio recordings from each facility for the selected sector periods. We typically made this request to the QA or automation personnel who provided the automation data. In cases where the TRACON and ATCT are not collocated and do not share a single communications system, we coordinated with additional personnel at the ATCTs. Most facilities provided the

voice recordings as requested. However, when the alerts occurred in Local sectors and the ATCT were not collocated or closely associated with the TRACON, we sometimes had difficulty obtaining voice recordings because the request for audio was made to the TRACON. Audio recordings are kept for only 45 days and if there was any delay in the initial analysis or the requests, the recordings may be unavailable. In addition, we encountered technical difficulties at some facilities where the recorder malfunctioned or the recordings were made with incorrect parameters or settings.

Depending on the capabilities of the facility, we received audio recordings on compact disc or cassette tape. When facilities provided us with only cassettes, we used Audacity audio software (Mazzoni, 2006) to make digital audio files (WAV formats) of the recordings.

2.4 Transcription

The main results of the project are based on an analysis of recordings of radio transmissions between controllers and pilots (also known as air-ground communications) and telephone or intercom calls between controllers (also known as ground-ground communications) or simply “coordination.” From experience, we know that controllers working near each other sometimes coordinate face-to-face without using the phone or intercom. In these cases, no recording of the coordination is available.

Analysts used Transcriber transcription software (Barras, 2005) to listen to each digital recording and create transcriptions of relevant communications. For each alert identified in the automation data, the analysts listened to the recordings 5 minutes before and 5 minutes after the alert activation and transcribed all communications related to the affected aircraft. Due to the large number of alerts in many sector periods, this strategy often resulted in the analysts listening to the entire recording.

When questions arose as to the meaning of a communication, an ATC SME reviewed the recording and helped the analyst build an accurate transcript. In cases where the analyst understood the audio but not the meaning, the SME reviewed a word-for-word transcript to explain the significance of the communication. In addition, after draft transcripts were complete, other ERPs reviewed the transcripts to ensure quality.²

In accident investigations, the NTSB develops word-for-word transcriptions of ATC communications, which are very resource intensive to create. Because of resource constraints and the large number of alerts analyzed, our transcriptions were intentionally not word-for-word

² Before we began the transcription process, we used the Audio Mining Development System voice recognition software (ScanSoft, 2002) on one recording to determine whether it could provide a useful “first pass” transcript. We found the resulting transcript required such extensive revisions that it was not useful to the human transcribers. Unlike recordings made during simulations, voice communications from the field contain static and noise, which can affect voice recognition accuracy. In addition, because we intended to review many different facilities and sectors, it was too resource intensive to build site-specific vocabularies (e.g., local fixes, airports, common airlines) for each site. Without site-specific vocabularies, recognition accuracy is reduced substantially. If future research on safety alerts focuses on a smaller number of facilities and sectors, building site-specific vocabularies can make the voice recognition process accurate enough to be useful.

in most cases. Instead, the transcriptions took the form of a summary of the substance of the communication with extra wording and phraseology excluded. Table 7 shows a sample of a summary transcript and an equivalent word-for-word transcript.

Table 7. Sample Summary Transcript Compared to a Word-for-Word Transcript for the Same Recording

Summary Transcript	Word-for-Word Transcript
<p>*UPS109* C: following heavy Airbus traffic, 10:00 6 nm northbound 4000; P: acknowledges</p> <p>*UPS109* P: did he pass us now? C: yes, he's paired with another aircraft 3000 for parallel; P: got him; C: follow heavy airbus caution wake turbulence clear visual app 35L additional traffic 11:00 10 miles for parallel 3000 B767</p>	<p><u>Controller</u>: U-P-S one zero nine heavy, you'll be following heavy Airbus traffic just coming in behind you now at ten o'clock six miles northbound descent out of four thousand.</p> <p><u>Pilot</u>: Ok, we're looking. I got a couple of 'em over there.</p> <p>[pause]</p> <p><u>Pilot</u>: Did he pass us now?</p> <p><u>Controller</u>: Uh, yes sir, he just passed you. He's paired up with another aircraft, three thousand, going to the parallel.</p> <p><u>Pilot</u>: Ok, we got him.</p> <p><u>Controller</u>: U-P-S one zero nine heavy, follow the heavy Airbus, use caution wake turbulence. Cleared visual approach runway three five [left]. Additional traffic eleven o'clock, one zero miles for the parallel runway three thousand feet a Boeing seven six seven.</p>

2.5 Time Synchronization

Much of our analysis is based on determining when a controller action occurred relative to the activation of the corresponding alert. Doing this with precision requires accurate timekeeping. ATC voice recordings are coded using the Inter-Range Instrumentation Group-B (IRIG-B) time code signal, which records the time in Coordinated Universal Time (commonly known as UTC or “Zulu” time). The HCS, ARTS, and STARS also mark events in UTC, which allows synchronization between the automation data and voice recording data.

Special equipment or software is required to read the time code signal. We used an ESE LX-270U/9 time code reader to interpret the IRIG-B signal. The voice communications were recorded on the left audio channel and the time code on the right. Output from the right channel was fed into the time code reader and the left channel into the computer speakers or headset.

However, some sites did not provide the IRIG-B time codes on the voice recordings. In other cases, the time code signal was recorded incorrectly. We believe that these errors were due to unfamiliarity with the process of making audio recordings for use outside the facility or to a misunderstanding in what we requested. For recordings without time codes, we assumed that the first second of the recording corresponded to the first second of the time period we requested and computed synchronization times accordingly. In some cases, sites provided us with explicit information that they had recorded different time periods than we requested (e.g., they started recording 5 minutes early). In these cases, we used the times the QA personnel provided to

compute the synchronization times. In two cases, S56 and LOU ATCT, we were unable to determine time synchronization with confidence, so we did not include these facilities in analyses where time synchronization is a factor.

While analyzing the audio data, we became aware of a phenomenon that added further complexity to the analysis of some recordings. Though voice recordings in the field begin as digital audio, some facilities lack a capability to record the data directly to a WAV format. In those cases, the facilities created an analog cassette first, and then they or we converted the analog cassette to a WAV format. However, small differences in speed between tape decks (e.g., the tape deck in the field ran slightly slower than the tape deck at our laboratory) can lead to small differences in timing that compound over time. For example, we found that a recording from M98 gained approximately 10 seconds over the course of an hour. That is, if 21:00:00 UTC occurred at 00:00:00 on the tape, 22:00:00 occurred at 01:00:10 on the tape. This corresponds to a discrepancy between their tape deck and ours of 0.003 seconds per second. Because we required precise time synchronization across the entire recording, we used the time code reader to manually locate each alert activation in the audio recording rather than compute its location based on the start time of the recording and the alert's time stamp. This process introduces a small amount of error based on human reaction time while manually marking times using the time code reader. We estimate this error to be less than 0.5 seconds.

2.6 Response Analysis Sample

In cases where there were no communications with the affected aircraft in the 10-minute window surrounding the alert, the analysts had to determine whether the affected aircraft were actually on the frequency near the time of the alert. If the affected aircraft are not on the frequency near the time of the alert, it is nearly impossible to say with confidence whether the alert received a response. It would be invalid to consider such alerts in the response analysis, because they would count as “no responses” when they actually may have received a response from a sector we did not examine. However, it is not easy to determine whether an aircraft is on a frequency at a particular moment. No system records which aircraft are on which frequency. The automation systems record when handoffs are made, not when communications are transferred. It is common for an aircraft to be listed in the automation system as being controlled by one sector while it is communicating with another, especially in the minutes surrounding handoffs. This is also true when aircraft are handed off from Final to Local. In addition, because of the limitations of SMON data in ARTS, many of the alerts in the sector-period sample occurred in nearby sectors and did not appear on the recordings we obtained.

To create the sample for the response analysis, analysts listened to the voice recordings for *any* communication between the controller and the affected aircraft. They decided whether to include the alert in the response analysis sample using the following criteria. In the examples, “Sector A” refers to the sector associated with the sector period we were examining. “Sector B” refers to adjacent sectors or facilities that hand off to or accept handoffs from Sector A.

1. If the recording contained no communications whatsoever between Sector A and the affected aircraft, we **excluded** the aircraft from the response analysis.
2. If the recording contained communications between Sector A and the affected aircraft that occurred in the 10-minute window (5 minutes before, 5 minutes after) surrounding the alert, we **included** the aircraft in the response analysis.

3. If the recording contained communications between Sector A and the affected aircraft, and those communications occurred both before and after but not during the 10-minute window, we **included** the aircraft in the response analysis. This is a standard “no response” situation.
4. If the recording contained communications between Sector A and the affected aircraft, but those communications occurred *only* before or *only* after the 10-minute window, we used the following additional criteria.
 - a. If the alert occurred more than 90 seconds before the very first communication on the recording for that aircraft, we **excluded** the aircraft from the response analysis. Usually the first communication on the recording was the initial check-in between the aircraft and Sector A (e.g., pilot says, “N123AB with you at one-zero thousand” or Sector A says, “N123AB are you on frequency?”). If the alert occurred more than 90 seconds before the initial check-in with Sector A, the pilot was probably still on the Sector B frequency at the time of the alert. Because we did not obtain voice recordings for every adjacent sector, we cannot know whether Sector B responded to the alert.
 - b. If the alert occurred more than 90 seconds after the very last recorded communication between the aircraft and Sector A, we **excluded** the aircraft from the response analysis. In almost all cases, this final communication was the Sector A controller transferring the aircraft to the Sector B frequency and the pilot acknowledging the transfer. After 90 seconds, the aircraft would likely have checked in with Sector B who would then be responsible for responding to the alert. Because we did not obtain voice recordings for every adjacent sector, we cannot know whether Sector B responded to the alert.
 - c. If the alert occurred more than 90 seconds but less than 15 minutes after a Final sector transferred the aircraft to the frequency of a Local sector, we **included** the aircraft from the response analysis. We used this exception because Final controllers typically do not make handoff entries in ARTS or STARS when handing off aircraft to Local for arrival. Without the handoff entry, the automation system considers the aircraft to still be in the Final sector, even though Local has responsibility for the aircraft. If an alert occurs during this period, Final will receive the alert as well as Local. After 15 minutes, however, the aircraft has probably landed and any new alerts that occur for that aircraft probably result from departure of a new flight with the same call sign or from a *go-around* that is being resequenced and will be treated as a new flight when it reaches Final again. In those rare cases, because we do not have voice recordings for all sectors, we cannot know who is communicating with the aircraft at the time of the alert and we cannot determine whether the alert was responded to appropriately.

By applying these criteria to the alert sample and the voice recordings, we created the sample for the analysis of controller responses. The response analysis sample consists of all the alerts where one or more of the aircraft involved appear in the voice recordings near the time of the alert. The totals for each facility are listed in Table 8 and Table 9.

Table 8. Alerts Included in the Response Analysis Sample by ARTCC

ARTCC	CA	MSAW
Atlanta	84	19
Houston	28	4
Indianapolis	3	0
Los Angeles	78	2
Salt Lake City	38	21
Total	231	46

Table 9. Alerts Included in the Response Analysis Sample by TRACON

TRACON	CA	MSAW
Atlanta	42	29
Boston	67	4
Burlington	7	12
Dallas-Fort Worth	25	7
Denver	18	22
Detroit	13	5
Indianapolis	3	4
Las Vegas	46	0
Louisville	35	16
Minneapolis	7	7
Northern California	1	8
Oklahoma City	0 ^a	0 ^a
Orlando	48	4
Philadelphia	10	8
Potomac	34	4
Salt Lake City	0 ^b	0 ^b
Southern California	20	2
Total	376	132

^aNo audio recordings received for this facility.

^bTechnical difficulties with the audio recordings prevented us from analyzing responses from this facility.

2.7 Response Categorization

For each alert in the response analysis sample, we used the transcripts to determine whether the controller made a response to at least one of the aircraft involved. A response is a controller-pilot or controller-controller communication that attempts to address the conflict or low altitude situation. To be counted as a response, the communication must be relevant to the situation in terms of the aircraft involved and the action taken. An alert can receive multiple responses. For example, a controller might issue a traffic advisory and issue a turn to the same aircraft in response to the conflict situation. Because CAs involve multiple aircraft, a single alert may receive many responses.

The following sections describe the categories of controller responses and provide examples of each category. The examples are taken from transcripts in the sample, but the call signs have been changed.

2.7.1 Control Instructions

In the model alert situation, the alert prompts the controller to issue a control instruction to the pilot. The pilot acknowledges the instruction and implements it in the cockpit. Control instructions commonly issued in response to safety alerts include the following:

- Altitude (e.g., climb, descend, cross fix at altitude)
 - Controller: US Airways 123, Houston center, good morning, climb and maintain flight level 230.
 - Pilot: Good morning. Up to 230, US Airways 123.
- Speed (e.g., increase speed, decrease speed, maintain speed)
 - Controller: N123DL, say your indicated speed.
 - Pilot: Now indicating two-eight-zero.
 - Controller: N123DL, reduce back to two-five-zero knots for sequencing.
 - Pilot: Slow to two-five-zero for spacing, N123DL.
- Heading (e.g., turn left, turn right, fly heading)
 - Controller: FedEx 123, turn 30 degrees left, vector for traffic.
 - Pilot: 30 degrees left, FedEx 123.
- Route (e.g., clear direct to fix, clear for approach)
 - Controller: ExecJet 123, cleared direct Daggett.
 - Pilot: Direct Daggett, ExecJet 123.

In some cases, instructions that are part of the routine procedures and routes in a sector will resolve an alert situation. We further categorized control instructions to distinguish instructions that seemed to be in direct response to the alert situation from instructions that we could not determine whether the instruction was a direct response to the alert. We categorized instructions as *directly applicable* when most or all of the following characteristics applied:

- Controllers issued a traffic advisory or alert in the same communication as the control instruction.
- Controllers did not issue the same control instruction to other aircraft in the sector that were not involved in alerts.
- Instructions included phraseology such as “vector is for traffic.”
- Instructions included phraseology indicating urgency such as “expedite,” “immediately,” or “maximum rate.”

2.7.2 Advisories and Alerts

In general, controllers issue traffic advisories to aircraft under their control when, in the controller’s judgment, the aircraft “proximity may diminish to less than the applicable separation” or to other aircraft on the frequency when, in the controller’s judgment, “their proximity warrants it” (FAA, 2006a, section 2-1-21). Traffic advisories are critical tools controllers may use to ensure that pilots are aware of the traffic situation around them. Until systems such as Automatic Dependent Surveillance-Broadcast (ADS-B) and Cockpit Displays of Traffic Information are widely available, traffic advisories are the primary mechanism by which controllers provide information to pilots about what is happening with other aircraft.

Issuance of a traffic advisory indicates that the controller has recognized the potentially hazardous situation, assessed its urgency and severity, and concluded that the aircraft may soon lose separation. A traffic advisory without accompanying control instructions indicates that the controller has determined that the situation does not require immediate action in terms of a turn or altitude change. Of course, a controller can misinterpret a traffic situation or assign it an improper priority. However, we know of no clearer indication that a controller has identified a potentially hazardous situation than issuing an advisory. The following is an example of a traffic advisory.

- Controller: United 345, traffic 2 o’clock, 5 miles, northeast bound, one-one thousand, and again, delete the speed restriction. I’ll have higher once you’re past Zalle.
- Pilot: 345 has traffic in sight.
- Controller: Excellent.

A safety alert is a more formal and serious version of an advisory. A controller is required to issue a safety alert if the aircraft is in an imminent situation that places it in “unsafe proximity to terrain, obstructions, or other aircraft” (FAA, 2006a, section 2-1-6). The following is an example of a safety alert for traffic.

- Pilot: 5EF is with you.
- Controller: 5EF, Atlanta approach, maintain VFR below Atlanta class bravo. Traffic Alert, 5EF, 11 o’clock, a mile eastbound, same altitude, unverified.
- Pilot: 5EF, we have the traffic.
- Controller: Roger.

2.7.3 Visual Separation

When the controller has already assured proper separation, the pilot can see another nearby aircraft, and other traffic, weather, and aircraft performance conditions are met, a controller can issue a visual separation clearance. In this case, the pilot assumes responsibility for maintaining proper separation from the other aircraft, for example:

- Controller: United 123, traffic 12 o'clock, 12 miles opposite direction, one-five thousand, Galaxy.
- Pilot: United 123, looking. [pause] Traffic in sight, United 123.
- Controller: United 123, can you maintain visual separation from that traffic, sir?
- Pilot: Yes, sir.
- Controller: United 123, roger, maintain visual separation with that traffic, descend, and maintain one-four thousand.
- Pilot: Down to one-four thousand, maintain visual, United 123.

In rare low altitude situations, controllers instruct pilots to maintain a safe distance from obstructions at the pilot's discretion. Issuing a visual separation clearance indicates that the controller is aware of the traffic situation and is taking action.

2.7.4 Repeats and Due Diligence

Even after an advisory or alert has been issued, controllers may make additional communications to verify that a pilot is aware of the situation or determine whether a pilot has identified the appropriate traffic. These "due diligence" communications may be prompted by the activation of a safety alert and take the form of repeating the advisory information given to the pilot ("traffic now at 10 o'clock and 3 miles") or asking the pilot to verify whether he has the target aircraft or obstruction in sight ("verify traffic in sight"). This repetition helps to ensure that the pilot has received the correct information from the controller. In some cases, the controller may check back two, three, or even four times, especially in cases where multiple aircraft are involved in the alert.

The following is an example of multiple due diligence communications in which the pilot and the controller talk about the traffic three times. The first advisory is counted as an advisory, the second and third are counted as due diligence.

- Controller: Bluestreak 123, [...] You do have an RJ, traffic 12 o'clock, 5 miles, turning southbound at 2000.
- Pilot: All right, I'm looking, 123.
- Controller: Bluestreak 123, maintain 2000.
- Pilot: 2000, 123.
- Controller: Bluestreak 123, turn right heading 180.
- Pilot: Turn 180, Bluestreak 123.

- Controller: Bluestreak 123, traffic's up at 3 o'clock and 3 miles, there's a company RJ at 2000.
- Pilot: I'm looking, 123.
- Controller: Bluestreak 123, turn right heading 240. You've got company in sight, 1 o'clock, 4 miles.
- Pilot: 240, negative on the traffic, Bluestreak 123.
- Controller: Bluestreak 123, roger, turn to localizer runway 26, report the airport.
- Pilot: We have him in sight, Bluestreak 123.

2.7.5 Coordination with Other Controllers

When the traffic situation or procedures require, controllers communicate with other controllers about the alert situation. If the controllers are seated face-to-face, these communications sometimes take place directly, but they normally occur over the voice communication system. Coordination such as this is especially likely when an aircraft enters alert status shortly after being handed off to the next sector. Due to recent incidents where automation was configured so that tower controllers did not receive audible alerts for aircraft under their control, a memo was issued reminding TRACON controllers of their responsibility to notify tower controllers of MSAW alerts (NTSB, 2006). The following is an example of communication between two controllers regarding an MSAW alert.

- TRACON Controller: Wilbur satellite, low altitude alert, November 678DL.
- Tower Local Controller: Thank you, AA.

3. RESULTS

The following sections discuss the results of the data analysis. The first section examines data from the full alert sample, regardless of whether audio data was obtained for the aircraft involved. The second section examines the response analysis sample, which includes only those alerts for which we obtained voice recording data.

3.1 Alert Sample Data

The result contained in the following two sections are based on the alert sample described in section 2.2.4. The alert sample includes the alerts occurring in or near the selected sector during the selected time period.

3.1.1 Operational Errors

We examined the Operational Error Database for the facilities and time periods included in the alert sample. Out of 1573 CAs and 480 MSAWs examined, none resulted in an OE or deviation. Therefore, in every case in our sample, controllers either resolved the situation by taking action or the situation resolved itself without action. That is, we found no instances like those reported in the ATO memo and the NTSB report where an alert activated, controllers took no action, and a safety incident resulted.

3.1.2 Duration

The duration of an alert is suggestive of its usefulness. Many alerts that deactivate quickly after activating are borderline alert conditions or could result from surveillance or tracking error. Such alerts provide no useful information and can create distraction and increase workload for controllers. Other short-duration alerts occur when the system detects a maneuver shortly after the alert activation. For example, a controller might recognize a potential conflict situation and issue a turn to the aircraft; however, time is needed for the pilot to implement the instruction, for the aircraft to respond, and for the tracker³ to detect the turn. During this time, the aircraft might go into alert status until the tracker recognizes the turn. This alert can be considered a nuisance because the controller had already resolved the situation and the alert provided no additional information. At best, these alerts serve as validation that the controller did the right thing by issuing the turn.

How long does an alert need to last before we can consider it not a nuisance? Research has shown that controller and pilot communications about maneuvers last 10.8 seconds on average and traffic advisories 10.9 seconds (Cardosi & Boole, 1991). In addition, research has shown that the time required for an aircraft in the landing configuration to level off and start a climb is 10-12 seconds (Rogers, 1999). Therefore, at least 20 seconds, and probably longer, are necessary for a controller to recognize an alert, formulate a response, communicate it to the pilot, and the pilot and aircraft to respond. Alerts lasting less than 20 seconds that do not result in OEs can be considered nuisances because they deactivated before a response by the controller could have taken effect.

In en route, aircraft positions update every 12 seconds based on the update rate of long range radars. In terminal, aircraft positions update every 4.8 seconds based on the update rate of short range radars. Using the 20-second criterion and being conservative, we consider alerts lasting only one update in en route or one, two, or three updates in terminal to be nuisance alerts. To simplify the analysis, we divided durations into five categories based on how long the alert lasted. Table 10 shows the number of alerts in each category. The definitions of the categories are:

- 1 update only (0-12 seconds in en route; 0-5 seconds in terminal);
- 2 or 3 updates (13-36 seconds; 6-15 seconds);
- 4 or 5 updates (37-60 seconds; 16-25 seconds);
- 6 or 7 updates (61-84 seconds; 26-35 seconds); and
- > 7 updates (> 84 seconds; > 35 seconds).

³ In ATC automation systems, the tracker is a set of algorithms and functions that process surveillance and flight plan data to determine aircraft positions and to associate targets with flight data.

Table 10. Frequency Distribution of Alert Durations

Alert Duration	CA		MSAW	
	ARTCC	TRACON	ARTCC	TRACON
1 update only	31% (94)	1% (2)	7% (6)	0% (0)
2 or 3 updates	34% (101)	35% (80)	39% (32)	53% (24)
4 or 5 updates	14% (41)	19% (43)	10% (8)	13% (6)
6 or 7 updates	6% (19)	12% (26)	7% (6)	9% (4)
> 7 updates	15% (44)	33% (75)	37% (31)	24% (11)

Note. The number in parentheses is the number of responses in that category. Due to how ARTS records alert activations, it is not possible to calculate the duration of the alert precisely. As a result, data from only HCS and STARS facilities are included in this analysis.

There were many short-duration CAs in both en route and terminal. In en route, 31% of the alerts lasted only one update (12 seconds or less). In terminal, 36% of the CAs lasted one, two, or three updates (15 seconds or less). In both cases, there was not enough time between the alert activation and deactivation for a controller response to have had an effect. These CAs ended because the situations resolved themselves without action by the controller or the controller had already taken action before the alert activated.

MSAWs showed a different pattern. In en route, 7% of MSAWs lasted just one update and, therefore, only a small percentage can be considered nuisances because of alert duration. In terminal, however, 53% of MSAWs lasted one, two, or three updates. Again, these MSAWs ended because the situations resolved themselves without action by the controller or the controller had already taken action before the MSAW occurred.

There were also numerous alerts with durations longer than seven updates. In these cases, the alert lasted longer than would have been necessary to resolve the situation if controllers and pilots had responded after the alert activated. We suspect that controllers delayed issuing control instructions because they concluded that the situation did not require immediate action and they wished to let the situation play out before committing to changing the flight paths of the aircraft. In addition, military flights frequently fly so close to other military flights or to the ground that they remain in alert status for a large portion of the flight. These flights contribute to the relatively large number of long-duration alerts.

3.2 Controller Responses to Alerts

The analyses contained in this section are based on the response analysis sample described in section 2.6. The response analysis sample includes all the alerts involved in CAs or MSAWs for which we obtained communication data at the time of the alert activation. Table 11 presents the number of controller responses to CAs and MSAWs.

Table 11. Controller Responses to Alerts

	CA		MSAW	
	ARTCC	TRACON	ARTCC	TRACON
Number of Alerts	231	376	46	132
Number of Alerts Receiving One or More Responses	87 (38%)	209 (56%)	4 (9%)	52 (39%)

In en route, controllers responded to CAs 38% of the time; in terminal, 56% of the time. En route controllers responded to 9% of the MSAWs whereas terminal controllers responded to 39%. This leaves substantial portions of both alert types that received no response. Reasons why a controller may not have responded to an alert include the following.

- The alert may have occurred while one or both aircraft was in a special status, such as military aircraft in Special Use Airspace, training flights that have pre-coordinated actions like missed approaches, or helicopter flights that have previously arranged their flight paths with the controllers. Of the aircraft involved with CAs in en route, 52% were in one of these special categories. In terminal, 43% were in one of these special categories. Of the aircraft involved with MSAWs in en route, 85% were in one of these special categories. In terminal, 13% were in one of these special categories. This suggests that a sizeable portion of the alerts that received no response were in one of these special categories. In these cases, no response *was* the appropriate controller action because the flights had been coordinated beforehand. However, the alerts associated with these flights still activated and the controller still had to attend to and make decisions about the alerts. These are true nuisance alerts.
- The controller may have assessed the situation and determined that no response was necessary. This is likely to be common for MCIs, where one aircraft is VFR, because only one aircraft is under control and can receive a response. This is also likely to be common for MSAWs, because the controller probably knows where the obstruction is and can quickly determine whether the aircraft is too close.
- The alert may have occurred while the aircraft was still on the previous sector’s frequency or after the pilot had changed to the next sector’s frequency.
- The controller may not have seen or heard the alert.

For purposes of counting the number of responses in different categories, alerts where no response occurred were excluded from further analysis. Table 12 and Table 13 present the percentage of controller responses in each category. Note that these tables count responses to aircraft involved with alerts rather than alerts. For MSAWs, CAs involving one uncontrolled aircraft, and CAs involving one aircraft in an adjacent sector, these numbers are the same. However, when a CA involves two controlled aircraft in the same sector, the controller can make multiple responses to the same alert (e.g., the controller may issue two traffic advisories – one to each aircraft). To account for this, we treat each aircraft involved in an alert as a “response opportunity.” The denominators in these tables reflect the number of aircraft receiving a response rather than the number of alerts receiving a response.

Table 12. Controller Responses to Aircraft Involved with CAs

Controller Action	Count	Denominator	Percentage
First Control Instruction (Directly Applicable)	18	394 ^a	4.6%
First Control Instruction	107	394	27.2%
Second Control Instruction	9	125 ^b	7.2%
Third Control Instruction	3	9	33.3%
Advisory/Alert	357	394	90.6%
First Repeat/Due Diligence	133	357 ^c	37.3%
Second Repeat/Due Diligence	48	133 ^d	36.1%
Third Repeat/Due Diligence	12	48 ^d	25.0%
Fourth Repeat/Due Diligence	4	12 ^d	33.3%
Visual Separation	16	394	4.0%
Controller-to-Controller Coordination	2	394	0.5%

^aNumber of aircraft involved in a CA that received a response. This number is not equal to the number of alerts receiving responses (296). This is because some CAs involved two aircraft in the sector simultaneously, and *both* aircraft may have received responses.

^bAircraft that received a first control instruction.

^cAircraft that received an initial advisory or alert.

^dAircraft that received a previous due diligence response.

Table 13. Controller Responses to Aircraft Involved with MSAWs

Response Category	Count	Denominator	Percentage
First Control Instruction (Directly Applicable)	2	56 ^a	3.6%
First Control Instruction	34	56	60.7%
Second Control Instruction	8	34 ^b	22.2%
Third Control Instruction	4	8	50.0%
Advisory/Alert	19	56	33.9%
Obstruction in Sight	3	56	5.4%
Controller-to-Controller Coordination	16	56	28.6%

^aNumber of aircraft involved in an MSAW that received a response. Because MSAWs involve only one aircraft, this number is equal to the number of MSAWs receiving a response (56).

^bAircraft that received a first control instruction.

Among the most important results of this analysis is that when controllers decide to respond to CAs, they provide a traffic advisory or alert to aircraft involved in the situation more than 90% of the time. The results also show that controllers follow up on traffic advisories with additional due diligence communications 37% of the time. These results show that when controllers do decide to respond, they are careful to inform the pilots about what is happening and increase their situation awareness.

Because only about one third of aircraft involved with CAs received a control instruction, it seems that controllers judged many situations to be serious enough to warrant an advisory but not serious enough to change the flight paths of the aircraft. This makes sense because an advisory is a relatively low cost response. A control instruction, however, is more costly in terms of its impact on the aircraft and controller workload: routes are modified, the controller must look for conflicts created by the new vector, the controller may have to make data entries, and so on. Because none of these aircraft was involved in an OE or deviation, these decisions seem to have been reasonable. In the cases where the controllers did issue control instructions, however, it is unknown whether the situation would have resulted in an OE if the control instruction had not been issued. Making this determination would require extensive review of the aircraft trajectories and flight plans and was beyond the scope of this project.

The low amount of coordination in response to CAs is somewhat surprising. In cases where CAs occurred between aircraft in adjacent sectors or near a sector boundary, we would expect controllers to coordinate with adjacent sectors, especially when they issue control instructions. The low number suggests that CAs between two sectors or near sector boundaries are less common than we thought. Alternately, controllers may simply be resolving the situations without informing adjacent sectors.

Responses seem to be somewhat different for MSAWs. In particular, controllers appear to issue advisories for MSAWs (34%) less frequently than for CAs and issue control instructions (61%) and coordinate (29%) more frequently. These results suggest that controllers regard MSAWs and CAs differently in terms of what response is most appropriate. Controllers seem to regard MSAWs as more likely to require a change in course, whereas CAs more likely require the transmission of information. MSAWs also seem to require controllers to provide information to other controllers more often than with CAs.

3.2.1 Time Between Alert and Controller Action

Table 14 and Table 15 show the median times before and after an alert activation that controllers took action to address the situation. In some cases, controllers took action on aircraft involved with the alert both before and after the alert activation. For example, a controller might issue a traffic advisory before a CA and then issue a control instruction after. The final columns in the tables show the median time after the alert only for those responses that were the first response of any kind taken on that aircraft involved with the alert. These “first” responses are the ones most likely to have been prompted directly by the alert.

Table 14. Time Between CAs and Responses

Controller Action	Median Seconds Before Alert	Median Seconds After Alert	Median Seconds After Alert
First Control Instruction (Directly Applicable)	86.1 (13)	102.6 (5)	4.1 (3)
First Control Instruction	85.5 (66)	80.0 (41)	87.7 (23)
Second Control Instruction	120.6 (7)	95.0 (2)	N/A
Third Control Instruction	129.8 (2)	200.7 (1)	N/A
Advisory/Alert	95.4 (245)	78.3 (112)	78.3 (104)
First Repeat/Due Diligence	71.2 (77)	85.1 (56)	N/A
Second Repeat/Due Diligence	84.8 (25)	142.1 (23)	N/A
Third Repeat/Due Diligence	24.9 (1)	45.1 (11)	N/A
Fourth Repeat/Due Diligence	None	261.0 (4)	N/A
Visual Separation	42.0 (13)	24.7 (3)	None
Controller-to-Controller Coordination	221.8 (1)	89.3 (1)	None

Note. The number in parentheses is the number of responses in that category.

Table 15. Time Between MSAWs and Responses

Controller Action	Median Seconds Before Alert	Median Seconds After Alert	Median Seconds After Alert (First Action)
First Control Instruction (Directly Applicable)	None	56.4 (2)	3.4 (1)
First Control Instruction	125.8 (23)	42.6 (11)	37.7 (9)
Second Control Instruction	218.4 (7)	42.9 (1)	N/A
Third Control Instruction	153.3 (4)	None	N/A
Advisory/Alert	7.8 (15)	4.0 (4)	3.0 (2)
Obstruction in Sight	136.1 (3)	None	None
Controller-to-Controller Coordination	6.2 (9)	4.5 (7)	4.5 (5)

Note. The number in parentheses is the number of responses in that category.

The data show that controllers most often take action to address potential safety situations *before* the CA or MSAW activates. For the 394 aircraft involved in a CA that received a response, 67% received the response before the alert activated. For the 56 aircraft involved in an MSAW where controllers made a response, 68% received the response before the alert activated. On the encouraging side, this result shows that controllers are continually searching for hazardous situations and proactively taking action rather than waiting for automation to tell them when action is necessary. On the negative side, taking action prior to the alert is strong evidence that most CAs and MSAWs notify controllers about situations of which they are already aware. At best, these alerts are redundant and create additional workload for controllers who must ignore or suppress them. At worst, these alerts distract controllers at critical moments, increase desensitization toward CAs and MSAWs, and increase distrust in the automation systems.

In addition, the median time from the activation of a CA to the controller response is quite long. The median time from the activation of a CA until the controller begins to issue a control instruction is approximately 88 s. The median time for an advisory is approximately 78 s. These long gaps suggest that controllers typically wait to see how a situation develops before taking action. This is not an indication that such alerts are unnecessary, because the alerts eventually lead to controller action. However, such long delays suggest that controllers do not consider most CAs to be situations requiring immediate attention. Further analysis should be conducted to see what finally prompts action from the controllers, such as the aircraft reaching some other internal threshold.

For MSAWs, the pattern of response is different. Controllers still take a long time, approximately 38 s, before they begin issuing a control instruction in response to an MSAW. Like CAs, this shows that controllers usually wait for low altitude situations to develop before instructing the pilot to climb or turn. However, they wait less than half as long before issuing control instructions when responding to MSAWs as when responding to CAs (38 s versus 88 s). The directly applicable control instructions do not show any difference (3.4 s versus 4.1 s). Although issuing control instructions in response to an MSAW is usually not required immediately, these results suggest that controllers do regard most MSAW situations as more urgent than most CAs. However, unlike CAs, controllers seem to begin issuing low altitude advisories almost immediately (3 s or 4 s) after the MSAW activates, though this is based on only two observations. This suggests that when a low altitude alert is needed, controllers consider this an urgent event that requires immediate attention.

4. DISCUSSION AND RECOMMENDATIONS

The following sections interpret the results and provide recommendations for improving the human factors attributes of ATC safety alerts. We group the recommendations into two categories (a) recommendations that attempt to reduce the number or impact of nuisance alerts and (b) recommendations that attempt to improve the presentation of the alerts in the automation systems. In simplest terms, alerts should have the characteristics described in section 1.3.

- Effective alerts accurately detect potential hazards
- Effective alerts do not activate when potential hazards do not exist
- Effective alerts provide new information
- Effective alerts are perceptible and salient
- Effective alerts are distinguishable from one another
- Effective alerts are easy to locate
- Effective alerts are legible
- Effective alerts communicate urgency
- Effective alerts communicate appropriate actions
- Effective alerts draw attention but do not disrupt operations

4.1 Nuisance Alerts

If the controller makes no response to an alert and the situation does not develop into an OE, we can conclude that the situation resolved itself without action by the controller and therefore the alert did not provide useful information. We count such alerts as nuisance alerts.

The results indicate that controllers experience a large number of nuisance CAs and MSAWs. The number of nuisance alerts constitutes the most serious human factors issue facing the alerting systems. As discussed in previous sections, nuisance alerts create workload and distractions and can lead to desensitization toward the alerts and poorer performance overall. Nuisance alerts also decrease trust and reliance on automation.

Determining whether an alert was a nuisance is a difficult task, after the fact, with limited data. In the moments after the alert activated, we do not know whether the controller considered the alert to be a nuisance. We cannot know what a controller knew about the traffic at the time the alert activated. An alert that might be a nuisance to one controller might be a necessary warning to another. We also cannot know what actions the controller took that were not recorded by the automation or the voice tapes. At best, we can estimate the nuisance alert rate based on what the controllers did or did not do when the alert activated.

We make the following conclusions about alerts in en route. Of the alerts examined in the project, 62% of the CAs and 91% of the MSAWs were nuisance alerts. They were nuisances in that no additional action from the controller was necessary after the alert to prevent the situation from developing into an OE. In addition, of the aircraft involved with CAs that received a controller response, 67% received the response *before* the alert activated. Of the aircraft involved with MSAWs that received a controller response, 68% received the response before the alert activated. Although not entirely nuisances, these alerts can be considered redundant or unnecessary. Furthermore, of all the CAs we examined, 31% lasted such a short time that controllers took action to address the situation prior to the alert activation or the alert situation resolved itself without action. Taken together, we estimate that as many as 87% of CAs and 97% of MSAWs did not provide useful information beyond what the controllers already knew and were not necessary to maintain safety.

We make similar conclusions about alerts in terminal. Of the alerts examined in the project, 44% of the CAs and 61% of the MSAWs were nuisance alerts. In addition, of the aircraft involved with CAs that received a controller response, 67% received the response *before* the alert activated. Of the aircraft involved with MSAWs that received a controller response, 68% received the response before the alert activated. Although not entirely nuisances, these alerts can be considered redundant or unnecessary. Furthermore, 36% of all CAs and 53% of all MSAWs lasted such a short time that controllers took action to address the situation prior to the alert activation or the alert situation resolved itself without action. Taken together, we estimate that as many as 81% of CAs and 87% of MSAWs did not provide useful information beyond what the controllers already knew and were not necessary to maintain safety.

Though this does not tell us what effect this large number of nuisance alerts had on controller performance, the human factors literature from other domains is clear: Professionals respond poorly to genuine alerts when those alerts occur in environments where many nuisance or low-priority alerts occur (see section 1.4.3). The high number of nuisance alerts decreases controller trust in the automation systems and desensitizes controllers toward CAs and MSAWs. When controllers become desensitized, they are more likely to overlook genuinely hazardous situations because they are accustomed to treating most as nuisances. In the accidents reported by the ATO and the NTSB, desensitization may have contributed to the controllers not responding to the alerts.

Research in other domains has shown that if the reliability rate of an alert is lower than 70%, professionals would perform better, overall, if they ignored the alert entirely and relied on their own detection and decision-making abilities (Wickens & Dixon, 2005). With more than 80% of the alerts examined here being nuisances or unnecessary, the reliability of ATC safety alerts

would need to be improved dramatically before we can conclude that they significantly help controllers maintain safety.

4.2 Recommendations to Reduce Nuisance Alerts

We recommend that the FAA develop alert algorithms and presentations that reduce the number of nuisance alerts and their potentially negative impact on controller performance.

In the following sections, we offer several recommendations for ways the negative effects of nuisance alerts could be reduced. We emphasize that salience and nuisance alerts are linked. Controllers may attribute the disruption and annoyance produced by nuisance alerts to the alert being too loud or obnoxious. Without also addressing the number of nuisance alerts, changes to the presentation that increase salience probably will be ineffective and may make problems worse. For example, if controllers already distrust an alert because of its high nuisance rate, making the alert louder without reducing the number of nuisance alerts will only serve to irritate the controllers and reduce trust further.

4.2.1 Provide Better Suppression Functionality

Controllers can manually suppress alerts using keyboard or trackball entries. Alerts can be suppressed in geographic regions using adaptation parameters. For example, the facility might create a suppression zone corresponding to a military warning area so that alerts are suppressed when the warning area is active. For MSAWs, the automation systems also incorporate geographic grids that specify safe altitudes in each cell. The grid allows facilities to determine where and at what altitude MSAWs activate. However, the MSAW grid is not necessarily designed to minimize nuisance alerts.

We realize that there is a natural reluctance and some official policies against suppressing safety alerts, despite the human factors problems nuisance alerts create. Supervisors and controllers remain concerned that if an OE occurs and the alert has been suppressed that the decision to suppress will be questioned and the controller will be blamed for the error. The NTSB has formally recommended that the FAA reduce the size of MSAW suppression regions (NTSB, 2000). On this view, reducing the amount of alert suppression is seen as a way to reduce risk (i.e., *better safe than sorry*). This is a reasonable position given the extreme costs of a miss compared to the minor costs of a false alarm. However, in an environment where many nuisance alerts occur, the greater risk may be in not suppressing nuisance alerts because they can cause such serious human factors problems. We recommend that the FAA improve the suppression functionality available in the automation systems and encourage controllers and facilities to use it.

From the controller's perspective, manually suppressing alerts increases workload with little perceived benefit. First, as our data show, many CAs and MSAWs ended quickly after the alert activated. In these cases, it is easier to simply wait for the alert to go away instead of making a manual entry. Second, if a controller suppresses an alert and the situation becomes more hazardous afterward, the alert may not reactivate when it really is needed. We do not have a strong human factors recommendation for how to address the workload issue, except that the suppression entry should be simple and require as few keystrokes or trackball movements as possible.

To address the second issue, however, suppression could be made to work more like a *snooze* function rather than removing the alert entirely. The controller could manually suppress the alert after it activates, but the suppression would deactivate if additional criteria are met. The criteria could be based on elapsed time (e.g., suppression deactivates if the suppressed aircraft are still in CA status 20 seconds after the suppression) or urgency (e.g., suppression deactivates if the suppressed aircraft are projected to lose separation in 45 seconds instead of 75 seconds). Alerts that already meet the stricter criteria at the time of their first activation would not be suppressible.

We recommend that automation systems deactivate alert suppression (i.e., allow the alert to reactivate) when a CA or MSAW lasts longer than or becomes more urgent than a second set of criteria.

From the local adaptation perspective, controllers and supervisors usually can identify general areas in their facilities where nuisance alerts are common. In some cases, facility personnel are aware of the exact locations and circumstances that cause nuisance alerts. For example, at one facility, we transcribed the following exchange between a controller and a pilot about nuisance MSAWs and their cause.

- Controller: [...] Turn left when able, you can taxi to the ramp, remain on this frequency, have a good one.
- Pilot: [...] Stay on this frequency. I was wondering, how come when you shoot the approach as published you get low altitude alerts?
- Controller: Uh, they've got a thing called an MSAW grid programmed on the radar and—that's minimum safe altitude warning—and anybody within 10 miles, I think, they've got the grid set up at the wrong altitudes, but we're required to give it and I don't know if or when they'll ever fix it.
- Pilot: All right, I was just wondering. It happened a couple times in that approach, so I was just wondering.
- Controller: Yeah, if you're shooting it on the profile, like if you're out there and once you pass the fix you descend to 1300 feet, the thing starts going bananas because it thinks you're making too rapid a descent.
- Pilot: Yes sir, I appreciate it, I was just wondering.

If these locations are known by ATC, why are suppression zones not available or not used there? Why are the MSAW grids not constructed to minimize nuisance alerts? Creating and maintaining good suppression zones and MSAW grids can be complex and expensive. Traffic patterns change throughout the day and at different times of year. Facilities would need to create numerous suppression zones tailored for very specific circumstances. If such zones were available, the number of nuisance alerts displayed to controllers could be reduced. However, a large number of suppression zones would increase the complexity for controllers and supervisors who would need to determine which zones should be active and which are appropriate for a given traffic pattern. This could be a serious problem if a controller or supervisor selected the

wrong suppression zone and suppressed alerts that they intended to see. In addition, increasing the number of suppression zones increases the complexity associated with creating and maintaining local adaptation parameters by technical operations personnel.

Creating good MSAW grids is also complex. New obstructions, such as communication towers, are constructed regularly. In addition, the cells of the MSAW grid are larger than necessary. We support efforts currently underway by the service units to reduce the size of the cells of the MSAW grid. Smaller cells mean more precise altitude parameters, which should reduce the number of nuisance alerts.

We recommend that field facilities establish processes that periodically determine where and under what circumstances nuisance alerts commonly occur at the facility for the purpose of building and improving suppression zones and MSAW grids.

This determination can be made through observations and discussions between automation and ATC or by examining logs recorded by the automation systems listing manual suppression entries made by controllers. If the logs reveal a location in the facility where controllers frequently suppress alerts manually, that location may be a good candidate for a suppression zone. When analysis reveals a consistent pattern of nuisance alerts or manual suppressions, a suppression zone should be built and tested to determine whether it successfully reduces the number of nuisance alerts or manual suppression actions. If so, the zone should be made permanently available and controllers should be appropriately trained and encouraged to use it. Furthermore, the suppression functions available in automation systems should incorporate additional factors such as the following:

- Beacon codes (e.g., prevent suppression of alerts for aircraft squawking special condition codes like 7700)
- Aircraft type or category (e.g., suppress alerts for all helicopters)
- Special designations (e.g., suppress alerts for military or training flights)
- Arrival airport or runway (e.g., suppress alerts if aircraft A is arriving at runway 35R and aircraft B is arriving at runway 35L)

Determining the most effective combinations of factors for a given situation will involve some trial and error. The complex interactions between multiple factors can be difficult to predict and should be tested carefully. We believe, however, that some of the negative effects of nuisance alerts can be mitigated with the use of more sophisticated suppression features.

Like manual suppression, zone suppression does not need to be all-or-nothing. Instead, zone suppression could delay the alert rather than suppress it entirely. That is, if an alert is suppressed because it falls in the suppression zone, it could be assigned another set of criteria rather than turned off completely. For example, the system might automatically suppress CA for two aircraft landing at closely spaced parallel runways and within 3 miles of the outer marker and that are projected to lose separation in 75 seconds. However, if the aircraft are projected to lose separation in 25 seconds, regardless of which runways they are landing on and their geographic location, the automatic suppression would be overridden and the alert would activate.

4.2.2 Develop Graded Alert Presentation

In environments where the base rate of a genuine negative event like an OE is low, even an accurate alerting system will have a low posterior probability that an alert indicates that an event is about to occur (Parasuraman & Hancock, 1999; Swets, Dawes, & Monahan, 2000). It is this low posterior probability that corresponds directly to a high false alarm rate. Increasing this posterior probability and reducing the false alarm rate while keeping the hit rate high is a critical and difficult problem (Swets et al.).

Because this is so difficult to accomplish, one way to improve the system is to reduce the impact of the false alarms on the human operators by providing alerts that have several levels or gradations rather than being all-or-nothing.

We recommend that the FAA develop prototype CA and MSAW presentations that incorporate gradations of urgency or likelihood and then evaluate the prototypes through human factors testing.

The alert gradations should be selected based on the urgency and/or the severity of the situation. A severe problem that is projected to occur very soon would receive the highest alert level, whereas a minor problem that is projected to occur in a long time would receive the lowest alert level. Because the most common situations will be coded with the least salient presentation, some of the negative effects of nuisance alerts will be reduced. In addition, because the most salient presentations will be reserved only for very urgent situations, users will be less likely to turn the brightness or volume of the alerts down because they will not be exposed to them very often. Because safety alerts are intended to prevent OEs, it makes sense for graded ATC alerts to focus mainly on urgency. Table 16 shows an example of a graded alert presentation in which the presentation becomes more salient as the urgency increases. In the table, we define urgency as the time until the projected loss of separation occurs, though urgency could be calculated in other ways. The main difficulty in designing graded alerts according to urgency is determining where to make the breakpoints. The breakpoints should be based on operational requirements such as the time necessary to complete critical actions like recognizing the alert and communicating the action to the pilot.

An alternative to basing the gradations on urgency is to base them on likelihood. In a likelihood alert, the system makes a separate determination of the likelihood of the hazardous situation actually occurring. For example, the algorithm might project that loss of separation will occur in 75 seconds, but it knows that an actual loss of separation occurred only 5% of the time in previous similar situations. The algorithm would activate an alert only when the likelihood crossed certain thresholds. The difficulty in developing likelihood alerts is that the algorithm must incorporate a model that calculates likelihoods of different outcomes based on available information. For example, given the positions, altitudes, and aircraft types involved in the current situation, what is the likelihood that a loss of separation will occur? Creating such a model requires examining previous alerts and comparing them to outcomes. Developing such a model would be complex and would require an in-depth data collection and analysis, but might be based on the data we collected for this project.

Table 16. Example of a Graded CA Based on Urgency

Urgency Level	Alert Presentation
<p>Level 1: Notification</p> <ul style="list-style-type: none"> • Long time until projected loss of separation • Controller should monitor the situation • Traffic advisory recommended • Immediate control instruction not necessary 	<ul style="list-style-type: none"> • “CA” appears in line 0 (above the call sign) for each affected data block • “CA” appears in red text • “CA” flashes at a slow rate
<p>Level 2: Alert</p> <ul style="list-style-type: none"> • Short time until projected loss of separation • Controller must place high priority on monitoring the situation • Traffic advisory required; traffic alert recommended • Immediate control instruction recommended 	<ul style="list-style-type: none"> • “CA” appears in line 0 (above the call sign) for each affected data block • Entire data block displayed in red text • Entire data block flashes at a slow rate • Audible alert with slow, low pitched tone
<p>Level 3: Alarm</p> <ul style="list-style-type: none"> • Loss of separation imminent • Situation requires controller’s highest priority • Traffic alert required • Immediate control instruction required • Urgent phraseology recommended 	<ul style="list-style-type: none"> • “CA” appears in line 0 (above the call sign) for each affected data block • Entire data block displayed in red text • Entire data block flashes at a fast rate • Audible alert activates with fast, high pitched tone

There is experimental evidence that the detrimental effects of nuisance alerts and false alerts can be mitigated by the use of graded and likelihood alerts (Gupta, Bisantz, & Singh, 2001; Sorkin, Kantowitz, & Kantowitz, 1988; Wickens & Xu, 2002). Allowing the look-ahead time to remain long but providing the user with additional information about the urgency of the alert improves performance, increases safety and trust, and decreases distraction (Lee, Hoffman, & Hayes, 2004; Thomas, Wickens, & Rantanen, 2003). These benefits are especially apparent under high workload conditions (Sorkin et al.). By providing information about the urgency of alerts, controllers would be encouraged to act quickly to resolve urgent situations and to better prioritize responses to less urgent ones.

4.2.3 Base Alert Times on Human Factors Research

Safety alerts are designed to provide controllers with enough time to prevent a loss of separation. It is critical that the determination of how much time is *enough* account for human abilities and limitations and that any estimate of human performance included in the alert algorithms be based on the latest human factors research.

In this project, we collected data that can be used in making these determinations. Ideally, we would like to be able to break down these values into their components, but that would require additional research, most likely in a human-in-the-loop simulation. In Table 17, we present reaction time parameters for which we can calculate values based on the data we have collected.

Table 17. Reaction Time Parameters

Reaction Time Parameter	Source
Time for controllers to react to the alert + Time for controllers to obtain necessary information + Time for controllers to decide what action to take =	Duration from alert activation until beginning of relevant communication from controller to pilot.
Time for controller to communicate instructions to the pilot + Time for pilots to hear and acknowledge the instructions =	Duration of communication between controller and pilot in which controller takes relevant action.

In this report, we have not calculated the values for these reaction time parameters. The data require further refinement and filtering before we can be confident that only applicable alert situations are included in the calculations. Furthermore, additional information must be obtained regarding pilot and airframe reaction times for algorithms to incorporate a complete $T_{critical}$.

We recommend that the FAA conduct further human factors studies or analyses of controller and pilot response times to alerts to develop more precise parameters for use in safety alert algorithms.

4.3 Recommendations to Improve Alert Presentations

In addition to reducing the number and impact of nuisance alerts, we offer several recommendations for improving alert presentations. We emphasize, however, that reducing nuisance alerts should be the first priority. Increasing the salience of the alert presentation may have negative effects if nuisance alerts are not also reduced.

4.3.1 Collocate Speakers with Displays

The intent of an audible alert is to draw attention to the display or information needed to address the situation. The audible alert should say, in essence, “Look here!” Audible alerts are more effective when the audible alert sounds from a location very close to the display (Brown, O'Hara, & Higgins, 2000, sections 4.5.6.3-8, 4.10-3, 4.10-6, and 4.10-7). At TRACONs with STARS or ARTS with ACDs, the audible alert speakers are mounted at each controller position. For ARTS sites without ACDs, however, the speakers may be located elsewhere in the row of consoles or may be located at positions unaffected by the situation.

In addition, in all ATCTs, controllers move around frequently and visual attention can be focused in a variety of places. The radar displays may be mounted from the ceiling or in the console close to the cab windows. The alert speaker may or may not be located close to the display, especially when the display is mounted from the ceiling. The programs should investigate how the speakers could be mounted closer to the radar displays. This may have technical, installation, or maintenance implications that should be examined.

We recommend that all audible alerts be located as close as possible to the displays affected by the alert.

4.3.2 Restrict Audible Alert Volume Settings

Human factors guidelines commonly advocate providing flexibility for users to adjust equipment to suit their individual needs and preferences. However, human factors guidelines also advocate that systems be designed to reduce the likelihood for human error. To be effective, audible alerts must be loud enough that they can be reliably heard over other sounds in the environment. Ahlstrom and Longo (2003) provide the following standards with regard to the volume of audible alerts:

- **Exceeding ambient noise** (section 7.2.8.3). Auditory signals shall exceed the prevailing ambient noise level by at least 10 dB(A) or any maximum sound level with a duration of 30 seconds by at least 5 dB(A), whichever is louder, without exceeding 115 dB(A) for emergency signals or 90 dB(A) for other signals.
- **Volume limits** (section 7.2.8.6). Volume control movement shall be restricted to prevent reducing the volume to an inaudible level or increasing it to an unacceptably high level.

ARTS and STARS allow users to adjust the audible alert volume at each working position. At ARTS sites with ACDs, users can adjust the alert volume from 1 to 100, nominally corresponding to 100 steps of 1% of full volume each. At STARS sites, users can adjust alert volume from 1 to 10, nominally corresponding to 10 steps of 10% of full volume each. However, the perceived increase in volume between the steps has not been validated. Both systems provide a mechanism for the user to test the volume of the alert before applying the setting. However, in both systems, the available adjustments are not tied to the sound level in the facility or individual operational positions. Facilities can restrict the volume settings through adaptation (e.g., allow settings 3 through 10 only), but it is not clear how these settings are selected or if such restrictions are used. At some TRACON and tower positions, some or all available volume settings may not be sufficient to meet the requirements of the FAA HF-STD-001.

It is not possible to be more specific about which settings are inadequate because no comprehensive survey of ambient sound levels at operational ATC positions is available. A survey of the locations in nine facilities where Technical Operations personnel work found ambient levels between 52 and 63 dB(A) (Ahlstrom & Panjwani, 2003). However, equivalent data are not available for the locations in facilities where controllers work. Informal sound level data collected for other purposes found ambient sound levels at two small TRACONs between 56 and 84 dB(A) (FAA, 1998; 2007b) and at one small tower between 49 and 68 dB(A) (FAA, 2007b). The most extensive survey available of sound levels in ATC facilities was conducted by

researchers from the Volpe National Transportation Systems Center (VNTSC) who collected sound data at nine ATCTs. They found that average ambient sound levels ranged from 48 dB(A) to 62 dB(A), with maximum levels ranging from 75 dB(A) to 90 dB(A) (VNTSC, 2007).

Informal measurements of the STARS CA audible alerts show them to be in the high 70s dB(A) and MSAW in the low 80s dB(A) when adjusted to the highest available setting (10). These levels appear to be adequate for most of the facilities surveyed, but a comprehensive dataset is necessary before such conclusions can be drawn for the NAS.

We recommend that the FAA conduct a comprehensive survey of ambient sound levels at operational working positions in ATC facilities.

A survey of this type has been conducted for ambient illumination levels (Wilson, Wilson, & Jha, 2007) but not for sound. We believe such a survey is necessary to ensure that audible alerts are effective throughout the NAS. This survey could be accomplished by human factors personnel or by facility safety and health officials. The survey should examine multiple operational positions in each facility and should consider the complex interactions between factors such as

- room size, shape, and materials (e.g., a large carpeted room with high ceilings will have different acoustic profile than a small room with low ceilings and no carpet);
- equipment and layout (e.g., sound level should be measured in locations with different configurations of equipment);
- environment (e.g., sound level should be measured at different locations relative to air conditioners, maintenance areas, break rooms); and
- time of day and staffing (e.g., ambient sound level should be measured at different times of day).

We recommend that automation systems use the ambient sound level data to restrict the minimum and maximum audible alert volume settings so they follow the FAA HF-STD-001.

That is, the system should prevent a controller from adjusting the volume quieter than the standard even when it is the controller's preference to do so. The human factors literature is clear: audible alerts must be louder than the ambient noise level if they are to draw attention effectively. The systems also should prevent controllers from adjusting the volume louder than the standard allows. We believe that ATC safety alerts do not qualify as "emergency signals" and therefore should not be adjustable louder than 90 dB(A), except as needed to accommodate very loud facilities.

Furthermore, the minimum and maximum volume settings should be established through local adaptation parameters on a position-by-position basis. Positions located in louder parts of a control room may require louder minimum volumes than other positions. However, the louder minimum setting may be unnecessarily loud for other positions. We discourage using volumes that are much louder than necessary because they can create disruption and distraction, especially

for positions that are not affected by the alert. In particular, this situation can arise when a minimum setting designed for the loudest position in the facility is used as the minimum setting for all positions.

Finally, facility personnel may need to occasionally conduct maintenance tests or other activities that cause safety alerts to activate. To reduce the potential impact on operations, systems could make volume settings available that are quieter than the standard but only for maintenance, training, and other non-operational purposes and restrict those settings to designated access levels or positions.

4.3.3 Improve Salience and Readability of Alert Data blocks

We recommend that the data blocks for active CAs and MSAWs follow human factors standards and best practices for salience and legibility.

When a controller's attention is drawn to the display, the critical information in the data block should be legible and interpretable. Because of clutter and data block overlap, however, this may not always be the case. To make the alert presentation optimal from a human factors point of view, alert data blocks should have the following parameters.

- Alert data blocks should use flash coding to draw attention. When text is flashing, the flash should have good attention-getting properties while maintaining good legibility, 1/3 to 1 flash per second with an on/off ratio of 70/30 (Ahlstrom & Longo, 2003, section 8.6.11.9). If only a symbol is flashing, a faster flash rate can be used, 2 to 5 flashes per second with an on/off ratio of 50/50 (Ahlstrom & Longo, section, 8.6.11.3).
- Alert data blocks should allow controllers to stop the flashing by making a keyboard or trackball entry, while the data block remains in the alert state (Ahlstrom & Longo, section, 8.6.11.5).
- Alert data blocks should be full data blocks only, not limited or partial.
- Alert data blocks should be drawn last (i.e., on top) of other data blocks and other displayed information.
- Alert data blocks should be displayed at the size that provides optimum legibility, 20 to 22 minutes of arc (Ahlstrom & Longo, section, 8.2.5.6.5).
- Alert data blocks should be displayed at a brightness that provides optimum contrast, between 6:1 and 10:1 (Ahlstrom & Longo, section, 8.2.5.6.12).
- Alert data blocks should be assigned highest priority in automatic data block offset algorithms (i.e., the offset algorithms should move other nearby data blocks before any adjustments are made to the alert data blocks).

The current automation systems appear to meet many of these standards, although none has been formally tested against the standards.

4.3.3.1 Color

Color coding is using differences in color to communicate different states or categories. Research has shown that color coding can substantially decrease visual search time for information (Treisman & Gelade, 1980). Controller opinions also show a strong preference for using color coding (Cardosi & Hannon, 1999).

We recommend that automation systems employ color coding to distinguish CA and MSAW data blocks from others. Because of numerous human factors issues associated with color coding, we further recommend that color coding be used only when it is redundant with other information, such as text.

The colors yellow and red are culturally associated with warning and alert conditions (Ahlstrom & Longo, 2003, section 8.6.2.4.3). Orange is also a traditional warning color commonly used in road signs and other safety applications. In ARTS, STARS, and DSR, yellow data blocks are already used for non-alert information; therefore, we recommend against using yellow as the alert color because of possible confusion and additional training this would require. We recommend that color coding for alerts use red or orange. The following RGB values for red and orange have been tested on monitors currently used in ARTCCs and TRACONs, and the results show that these RGB values provide adequate contrast and accurate color identification (Xing, 2007).

- Red: RGB [255, 0, 0]
- Orange: RGB [255, 165, 0]

ARTS and STARS follow these recommendations by providing red text for alert data blocks that provide redundant coding of “CA” or “LA” in the first line. Note that ARTS currently uses a color called orange to code precipitation intensity, but this orange does not provide suitable contrast to use for safety alerts. In addition, this color orange is difficult to identify by people with color vision deficiencies (Xing, 2006b). In cases where red and orange are used by the same system, red should always code the more urgent situation.

We recommend that automation programs test all safety alert colors in every environment where they will be used to ensure they provide adequate contrast.

Pure red, such as RGB (255, 0, 0), provides only marginal contrast on a black background in bright environments (Crown Communications, 1998). In cases where pure red does not provide adequate contrast, other RGB values, such as RGB (255, 50, 50), can be considered that increase the brightness of the color, as long as the resulting color can be accurately identified as red and cannot be confused with pink, brown, orange, or magenta. Accurate color identification should be verified through human factors testing (Cardosi & Hannon, 1999). For other guidance on the effective use of color in ATC displays, consult reports by Cardosi and Hannon (1999) or Xing (2006a; 2007).

4.4 Summary of Recommendations

The following list is a summary of the most important recommendations discussed in this report.

- We recommend that the FAA develop alert algorithms and presentations that reduce the number of nuisance alerts and their potentially negative impact on controller performance.
- We recommend that automation systems deactivate alert suppression (i.e., allow the alert to reactivate) when a CA or MSAW lasts longer than or becomes more urgent than a second set of criteria.
- We recommend that field facilities establish processes that periodically determine where and under what circumstances nuisance alerts commonly occur at the facility for the purpose of building and improving suppression zones and MSAW grids.
- We recommend that the FAA develop prototype CA and MSAW presentations that incorporate gradations of urgency or likelihood and then evaluate the prototypes through human factors testing.
- We recommend that the FAA conduct further human factors studies or analyses of controller and pilot response times to alerts to develop more precise parameters for use in safety alert algorithms.
- We recommend that all audible alerts be located as close as possible to the displays affected by the alert.
- We recommend that the FAA conduct a comprehensive survey of ambient sound levels at operational working positions in ATC facilities.
- We recommend that automation systems use the ambient sound level data to restrict the minimum and maximum audible alert volume settings to follow the FAA HF-STD-001.
- We recommend that the data blocks for active CAs and MSAWs follow human factors standards and best practices for salience and legibility.
- We recommend that automation systems employ color coding to distinguish CA and MSAW data blocks from others. Because of numerous human factors issues associated with color coding, we further recommend that color coding be used only when it is redundant with other information, such as text.
- We recommend that automation programs test all safety alert colors in every environment where they will be used to ensure they provide adequate contrast.

5. FUTURE ACTIVITIES AND RESEARCH

The following sections discuss further work that is necessary to improve the effectiveness of ATC safety alerts. In some cases, this work is already underway. In other cases, a longer-term research and development program should be established.

5.1 Causes of Nuisance Alerts

Identifying nuisance alerts by listening to voice recordings and analyzing controller responses does not tell us why the nuisance alerts occurred. What factors led the algorithms to conclude

that an alert was needed to maintain safety when, in fact, it was not? What information could the algorithms have considered that would have led to a different conclusion? Answers to these questions are necessary to assist algorithm developers in selecting new information to include in the algorithms and developing better alert parameters.

Determining why a nuisance alert occurred requires a detailed analysis of the trajectories and flight plans of the aircraft involved. For example, the CA algorithm may detect two aircraft at the same altitude about to lose separation. However, unknown to the algorithm, one aircraft is planning to begin a descent when it reaches its next fix in 20 seconds. In this case, if the algorithm had incorporated information about aircraft short-term intent, the nuisance CA might have been avoided. The HFRG has provided a grant to human factors professionals from Alion Sciences and New Mexico State University to examine CAs from this perspective, using data obtained for this project.

5.2 ARTS Algorithm Development

In 2006, engineers from the ARTS program and its contractors began conducting tests and engineering analysis of alternate parameter sets for the ARTS safety alerts. In particular, they tested various look-ahead times and examined the effects on nuisance alerts. During the conduct of our project, the engineers from ARTS contacted us and requested assistance in determining research-based parameters for testing. In particular, they were interested in obtaining specific values for controller reaction times and communication times. The automation and voice communication data collected here lend themselves to determining these values.

In the coming months, we intend to work with the ARTS program to further analyze our data to determine the necessary parameters and assist them with testing and evaluating potential changes to the algorithms.

5.3 Incorporating Intent into Alert Algorithms

One major source of nuisance alerts is that the automation systems are not aware of every action that controllers take. For example, if the system knew that an aircraft had been cleared for an instrument approach, and the controller had instructed the aircraft to turn and intercept the localizer, the algorithms could better project routes of flight and better determine whether to activate an alert for that aircraft. En route systems incorporate some of this functionality already, such as routes and interim altitudes, but current terminal systems either lack the functionality entirely or controllers do not commonly use it. In addition to improving the safety alerts, in theory, other automation tools such as URET and the Traffic Management Advisor could also benefit if they had access to this type of information.

In theory, this problem could be easily solved by modifying systems and procedures to require controllers to make a corresponding entry for each action and communication. In en route, controllers already make entries when they change routes and altitudes. They also have the option of entering changes in speed and heading. In terminal, however, controllers currently do not normally enter these actions into ARTS or STARS. They also do not indicate to the system when they clear an aircraft to take off or land, intercept the localizer, or other important terminal actions. Providing information like this to the system would help the system make better route projections and could reduce the number of nuisance alerts. In practice, however, this solution

would create human factors problems. For routine traffic patterns and at slower times of day, making such entries might be feasible. However, when traffic is heavy or when the situation is time critical, controllers devote most of their attention to assuring separation and communicating rather than keeping the automation system updated. As a result, during situations when the safety alerts are needed most, the algorithms would not have the up-to-date information about controller actions that they need. We do not believe that asking controllers to provide the automation system with more information about their actions is the best way to reduce nuisance alerts.

Instead, we believe that technologies now entering the NAS or under development may help improve the safety alert algorithms. For example, when controller-pilot Data Communications (DataComm) becomes widely available in en route and terminal, controllers will be able to uplink instructions to pilots rather than communicate them over the radio. In theory, the safety alert algorithms could use the information transmitted over DataComm to determine when certain control actions have been issued and acknowledged. The algorithms could use this information to make better projections of routes of flight and perhaps eliminate or suppress some number of nuisance alerts.

Furthermore, ADS-B allows aircraft to broadcast a variety of information about the aircraft that could be used by safety alert algorithms. For example, algorithms currently determine whether an aircraft is making a maneuver and may delay alert activation until the maneuver is complete. If the aircraft can broadcast its turn status through ADS-B, the maneuver determination could be made more accurately and faster. Giving the safety alert algorithms precise information about what an aircraft is doing should help to reduce nuisance alerts.

We believe that the ATC research community should examine how DataComm and ADS-B could improve the effectiveness of safety alerts. In particular, researchers should determine what information available from these systems would be beneficial to the alert algorithms and then develop prototype algorithms that incorporate this information. Then, as part of test and evaluation activities for DataComm and ADS-B, researchers should examine whether the potential benefits to safety alerts are realized. In particular, does implementing these technologies and the corresponding changes to the algorithms reduce the rate of nuisance alerts? Are there new types of nuisance alerts introduced?

5.4 Examination of Relationship Between Handoffs and Alerts

Accidents LAX04FA205, LAX06FA106A, and DFW05FA202 described in the NTSB report occurred within a few minutes of a handoff and involved some degree of ambiguous responsibility (NTSB, 2006). In ATC, a handoff occurs when an aircraft approaches a sector boundary and the controller who currently holds responsibility for it transfers that responsibility to the next controller. A handoff consists of the following steps described here with an example in which Sector A hands off AAL1234 to Sector B.

1. The transferring controller initiates a handoff in the automation system, typically through a keyboard or trackball entry. Handoff status is indicated in the data block, typically via text and flashing. For example, the Sector A controller makes a keyboard and trackball entry to initiate the handoff of AAL1234 to Sector B. This causes the data block at Sector B to change from a limited to full data block and begin flashing.

2. The receiving controller indicates acceptance of the handoff by making an entry. For example, the Sector B controller clicks on the AAL1234 target with the cursor. The data block stops flashing and shows Sector B as the owner.
3. The transferring controller instructs the pilot to change frequencies to the receiving sector frequency. For example, the Sector A controller says, “American 1234, contact Genera Approach on 123.14.”
4. The pilot acknowledges the frequency change. For example, the pilot says, “123.14, American 1234. Good day.”
5. The transferring controller may take some action to indicate that the frequency change has occurred. Such actions are typically not mandatory and are a matter of controller technique and preference. Examples of actions include marks on flight progress strips or a *slant zero* in which a controller uses a data block offset entry to position the data block as a reminder.
6. The pilot changes the frequency on the cockpit radio.
7. The pilot checks in with receiving controller. For example, the pilot of AAL1234 says, “Genera Approach, American 1234 with you at flight level 210.”
8. The receiving controller acknowledges the pilot. For example, the Sector B controller says, “American 1234, Genera Approach, roger.”
9. Some time later, the transferring controller takes an action to drop the full data block to a limited data block.

Ideally, all of these actions would occur without error or delay. In reality, there are numerous opportunities for mistakes during handoffs. For example, the transferring controller may make a data entry error and hand off the wrong aircraft or to the wrong sector. The controllers or pilots may make errors in speaking or understanding the radio transmissions.

Problems can be magnified when responsibility for an aircraft becomes ambiguous from the perspective of the overall human-machine system.⁴ Responsibility becomes ambiguous when one part of the human-machine system believes one thing about the situation and other parts of the system believe something else. For example, controllers often initiate handoffs in the automation system several minutes prior to directing pilots to change frequencies (see Figure 5). This creates a situation in which the automation system believes that the receiving controller is communicating with the aircraft when, in fact, the transferring controller still is. In other cases, controllers may transfer communications but not make handoffs in the automation system at all, as is commonly done when a controller working Final in a TRACON hands off to a Local controller in an ATCT. The automation knows nothing about the transfer of communications; it knows only about the entries the controllers make.

⁴ We are using the term “responsibility” in a general sense. From a legal standpoint, responsibility for an aircraft at a given time is clearly defined in ATC procedures. However, it is a separate question whether the automation system, the controllers, and the pilots involved in a given situation know, understand, and agree on every rule that applies.

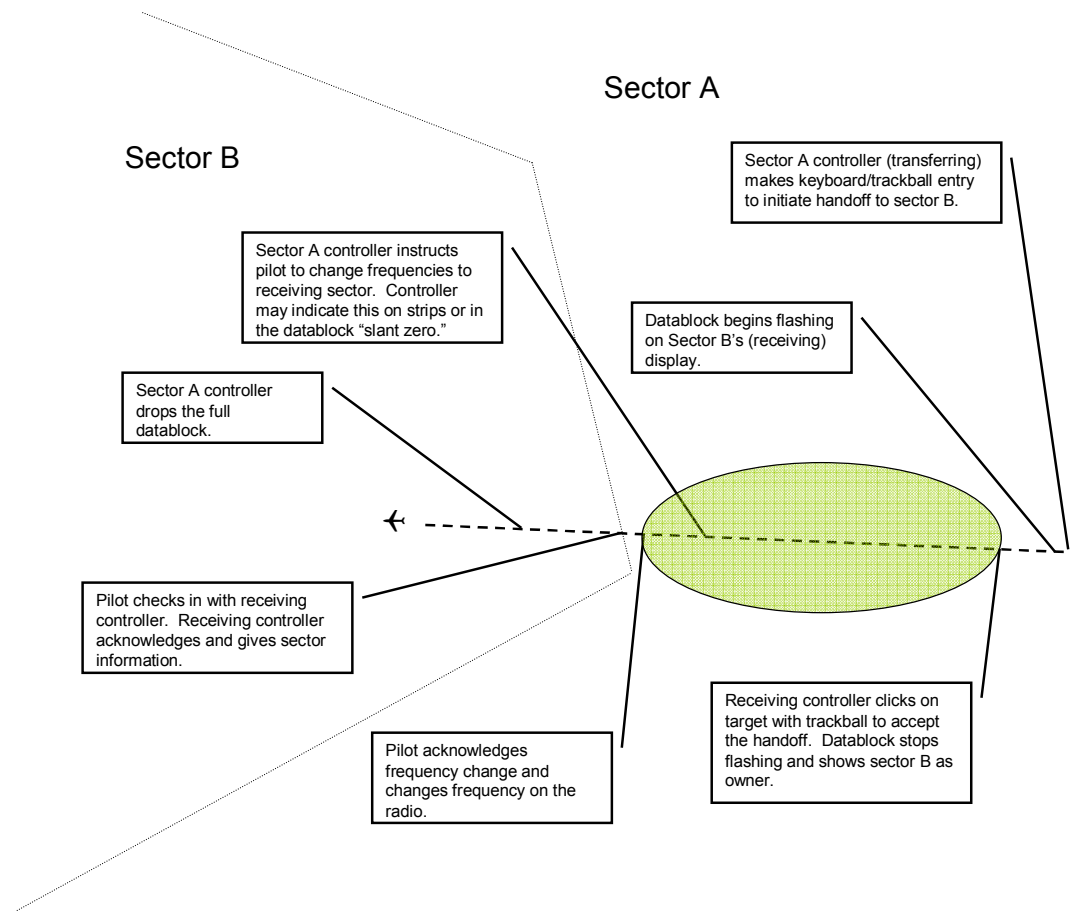


Figure 5. A typical handoff. The shaded oval indicates a time period in which the automation system believes that Sector B is responsible for the aircraft, but Sector A is actually in communication with the aircraft.

These situations create ambiguity for the system and may increase the likelihood of mistakes by controllers. A controller normally can tell who is communicating with the aircraft involved in a CA or MSAW by looking at information in the data block. In these situations, the automation does not represent the current truth about who is communicating with the aircraft (and hence who is able to take action on that aircraft if needed). The controller must rely on memory or consult other sources for the information. For example, a controller working *Final* who sees an MSAW needs to determine whether that aircraft has already been transferred to the tower frequency. The controllers cannot determine this by looking at the ARTS or STARS display but instead must remember the transfer of communications, examine flight progress strips, or call the Local controller. Relying on memory in these situations can lead to assumptions and mistakes. In addition, if the aircraft has already been transferred and the Local controller is already addressing the situation, the coordination call would create additional workload for both controllers. If the aircraft has not been transferred already, mistakenly trying to coordinate could delay the first controller from taking action.

Situations of ambiguous responsibility do not just affect the automation system. Even when controllers know who is responsible “by the book,” they may be inadvertently relying on other controllers to manage situations. For example, as shown in Figure 5, if an alert were to occur in the ambiguous period, Sector A might assume that because Sector B has the data block, Sector B will take care of the problem; whereas Sector B might assume that because Sector A has communications, Sector A will take care of the problem. If this occurs, neither controller might respond until it is too late to prevent an error. Ambiguity of this type can also arise when aircraft in two different sectors come into conflict. During these situations, each controller may assume that the other controller will take action to address the situation. This may delay communication of safety-critical information to pilots.

Due to the potential for problems in these situations, alerts occurring near handoffs may be key places to look for human factors issues associated with safety alerts. If a safety alert occurs during a period when responsibility is ambiguous, especially if the automation system believes something different than the controllers believe, there may be an increased likelihood for a controller to misattribute the alert as a nuisance or mistakenly defer responsibility to another controller. If this is true, changes to the automation or to procedures may be necessary to reduce the ambiguity.

5.5 Analysis of Controller Awareness and Responses

The data collected for this project tell us how and whether controllers responded to alerts. However, they cannot tell us with certainty if a controller was aware of a situation before the alert activated and cannot explain why a controller chose whether or not to respond. These data also cannot tell us why the controller chose that response. That type of detailed analysis could be conducted using techniques similar to those used for OE and accident investigations. For example, researchers could monitor automation data to determine when an MSAW with chosen characteristics occurs. When it occurred, the researchers could then obtain the automation and communication data for that MSAW and conduct interviews with the controllers and supervisors involved. Researchers could also review replays of the situation with QA personnel. The resources required to conduct such analyses would dictate examining a smaller number of alerts than the current study but in much more detail.

Like incident investigations, there is some chance that controllers could forget or misinterpret why they made certain decisions about a safety alert. These misinterpretations also are possible when reviewing replays because reviewers may have their own ideas about what “should” have happened. These problems could lead to biases in the data and incorrect conclusions. To reduce these problems, reviews of safety alerts should occur as soon as possible after the alert. In addition, it should be made completely clear to all involved that human factors research on safety alerts of this sort would be completely separate from any formal investigation or QA audit.

In another approach, human-in-the-loop simulations of alert situations could be conducted in which controller situation awareness and response times are measured. Interviews and replays also could be conducted as part of the simulations. The advantage of using simulations instead of field data is that the number and characteristics of alert situations can be manipulated experimentally. Researchers would not need to wait for situations to occur that match the characteristics they want. Interviews could be conducted immediately after the alert without affecting operations.

In addition, data collection tools, such as eye tracking, can be used in simulations that cannot be used in the field. For example, how often and how long a controller looks at a data block can be considered a measure of how concerned the controller is about that aircraft. If a controller looks at an aircraft very often just prior to a CA involving that aircraft, we might conclude that the controller was at least partially aware of the situation before the alert activated. The eye tracking data would, in theory, correlate with other indicators of concern about the aircraft such as communications and coordination. However, if the controller looks at the aircraft only as often as every other aircraft before the CA but then looks at the aircraft much more often afterwards, we might conclude that the controller was unaware of the situation before the alert.

5.6 User-Centered Design and Evaluation

Finally, we strongly believe that no new safety alert presentation or algorithm should be deployed to the NAS without user-centered design and evaluation. For new alert presentations (e.g., new data block colors or alert sounds), designs should be developed following the iterative rapid prototyping process used on past programs (Allendoerfer & Yuditsky, 2001). The new alert presentation should be developed by a multidisciplinary team composed of controllers and supervisors, ATC procedure and operational specialists, system engineers, and human factors professionals. Visual and auditory prototypes of new alert presentations should be created using high-fidelity tools and tested through multiple small scale part-task assessments. The team's final designs should be tested in human-in-the-loop simulations where controllers work the same traffic scenarios but with different alert algorithms or parameters. Operational metrics such as the number of nuisance alerts and losses of separation can be measured. In addition, human performance metrics such as reaction time, visual attention, and level of situation awareness also can be measured, allowing human factors researchers to compare the effectiveness of different algorithms.

References

- Ahlstrom, V. (2003). *An initial survey of National Airspace System auditory alarm issues in terminal Air Traffic Control* (DOT/FAA/CT-TN-03/10). Atlantic City International Airport, NJ: FAA, William J. Hughes Technical Center.
- Ahlstrom, V., & Longo, K. (Eds.). (2003). *Human factors design standard for acquisition of commercial-off-the-shelf subsystems, non-developmental items, and developmental systems* (DOT/FAA/CT-03/05/HF-STD-001). Atlantic City International Airport, NJ: FAA, William J. Hughes Technical Center.
- Ahlstrom, V., & Panjwani, G. (2003). *Auditory alarms in the airway facilities environment* (DOT/FAA/CT-TN-04/04). Atlantic City International Airport, NJ: FAA, William J. Hughes Technical Center.
- Allendoerfer, K., & Yuditsky, T. (2001). Further lessons learned in the design of computer-human interfaces for terminal air traffic control systems. *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, 45, 195-199.
- Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science*, 26, 39-83.
- Barras, C. (2005). Transcriber (Version 1.5.1) [Computer software]. Retrieved October 29, 2007, from <http://trans.sourceforge.net>
- Brown, W. S., O'Hara, J. M., & Higgins, J. C. (2000). *Advanced alarm systems: Revision of guidance and its technical basis* (NUREG/CR-6684). Washington, DC: U.S. Nuclear Regulatory Commission.
- Cardosi, K., & Boole, P. W. (1991). *Analysis of pilot response time to time-critical Air Traffic Control calls* (NTIS No. DOT/FAA/RD-91/20). Cambridge, MA: National Transportation Systems Center.
- Cardosi, K., & Hannon, D. (1999). *Guidelines for the use of color in ATC displays* (DOT/FAA/AR-99/52). Washington, DC: Office of Aviation Research.
- Christ, R. E. (1975). Review and analysis of color coding research for visual displays. *Human Factors*, 7, 542-70.
- Crown Communications, Inc. (1998). *STARS tower display workstation visual characteristics final report* (Crown Document No. C012-003-0010). Washington, DC: Author.
- Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors*, 33, 205-231.
- Federal Aviation Administration. (1995). *Human factors policy* (DOT/FAA/Order 9550.8A). Retrieved June 5, 2007, from <http://www.hf.faa.gov/docs/508/docs/HForder.pdf>
- Federal Aviation Administration. (1998). [Data collected at Atlantic City International Airport (ACY) TRACON]. Unpublished raw data.

- Federal Aviation Administration. (2006a). *Adjustable volume for the TRACON, Case File Number ATOOT-STARs-1021*. Internal system configuration management communications between a field site and the FAA Standard Terminal Automation Replacement System (STARs) program, initiated February 1, 2006.
- Federal Aviation Administration. (2006b). *Air Traffic Control* (DOT/FAA/Order 7110.65R). Retrieved June 29, 2007, from <http://www.faa.gov/atpubs>
- Federal Aviation Administration. (2006c). *Human factors study for safety alerts*. Internal FAA memorandum sent by Vice Presidents of Terminal Services, En Route and Oceanic Services, and Safety Services to Acting Vice President of Operations Planning, dated June 9, 2006.
- Federal Aviation Administration. (2007a). *Administrator's fact book*. Retrieved June 15, 2007, from <http://www.faa.gov/abouttoofficeorg/headquarters/offices/aba/adminfactbook/>
- Federal Aviation Administration. (2007b). [Aural alarm data from Greer, SC]. Unpublished raw data.
- Getty, D. J., Swets, J. A., Pickett, R. M., & Gonthier, D. (1995). System operator response to warnings of danger: A laboratory investigation of the effects of the predictive value of a warning on human response time. *Journal of Experimental Psychology: Applied*, 1, 19-33.
- Guillaume, A., Drake, C., Rivenez, M., Pellieux, L., & Chastres, V. (2002). Perception of urgency and alarm design. *Proceedings of the 8th International Conference on Auditory Displays*, 357-361.
- Gupta, N., Bisantz, A. M., & Singh, T. (2001). Investigation of factors affecting driver performance using adverse condition warning systems. *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, 1699-1703.
- Jones, D. M. (1999). The 1997 BPS Broadbent Lecture: The cognitive psychology of auditory distraction. *British Journal of Psychology*, 90, 167-187.
- Lee, J. D., Hoffman, J. D., & Hayes, E. (2004). Collision warning design to mitigate driver distraction. *Proceedings of the 2004 SIGCHI Conference on Human Factors in Computing Systems*, 65-72.
- Maltz, M., & Meyer, J. (2001). Use of warnings in an attentionally demanding detection task. *Human Factors*, 43, 217-226.
- Mazzoni, D. (2006). Audacity (Version 1.2.6) [Computer software]. Retrieved from <http://audacity.sourceforge.net/>
- Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors*, 43, 563-572.

- National Transportation Safety Board. (2000). *Aircraft accident report: Controlled flight into terrain-Korean Air flight 801* (Report Number NTSB/AAR-00/01). Washington, DC: Author.
- National Transportation Safety Board. (2006). *Safety recommendation A-06-44 through 47*. Internal communication sent by NTSB to FAA Administrator dated July 12, 2006. Retrieved June 29, 2007, from http://www.nts.gov/recs/letters/2006/A06_44_47.pdf
- Parasuraman, R., & Hancock, P. A. (1999). Designing automated warning systems: A signal detection and Bayesian analysis. In M. W. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 63-67). Mahwah, NJ: Erlbaum.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, and abuse. *Human Factors*, 39, 230-253.
- Rogers, R. (1999). *Flight test results of the controlled flight into terrain avoidance maneuver in fly-by-wire transports*. Report of the Air Line Pilots Association. Airworthiness Performance Evaluation and Certification Committee. Retrieved June 29, 2007, from http://www.rvs.unibielefeld.de/publications/Incidents/DOCS/Research/Other/Article/Rog Reports/Rogers_99_CFIT_FBW.pdf
- Sanquist, T. F., Thurman, D. A., & Mahy, H. A. (2005). The human element in homeland security systems: Application issues and research needs. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 913-917.
- Sarter, N. B. (2005). Graded and multimodal interruption cueing in support of preattentive reference and attention management. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 478-479.
- ScanSoft. (2002). Audio Mining Development System (Version 1.2) [Computer software]. Burlington, MA: Author.
- Seagull, F. J., & Sanderson, P. M. (2001). Anesthesia alarms in surgical context: An observational study. *Human Factors*, 43, 66-77.
- Sorkin, R. D., Kantowitz, B. H., & Kantowitz, S. C. (1988). Likelihood alarm displays. *Human Factors*, 30, 445-459.
- Sorkin, R. D., & Woods, D. D. (1985). Systems with human monitors: A signal detection analysis. *Human-Computer Interaction*, 1, 49-75.
- Stokes, A., & Wickens, C. (1988). Aviation displays. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 387-431). San Diego: Academic.
- Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Sciences in the Public Interest (a supplement to Psychological Science)*, 1, 1-25.

- Thomas, L. C., Wickens, C. D., & Rantanen, E. M. (2003). Imperfect automation in aviation traffic alerts: A review of conflict detection algorithms and their implications for human factors research. *Proceedings of the Human Factors & Ergonomics Society 47th Annual Meeting*, 344-348.
- Topf, M., & Dillon, E. (1988). Noise-induced stress as a predictor of burnout in critical care nurses. *Heart & Lung*, 17, 567-74.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Volpe National Transportation Systems Center. (2007). [Sound levels from nine Airport Traffic Control Towers]. Unpublished raw data.
- Watson, M., Russell, W. J., & Sanderson, P. (2000). Anaesthesia monitoring, alarm proliferation, and ecological interface design. *Australian Journal of Information Systems*, 7(2), 109-114.
- Watson, M., Sanderson, P., & Anderson, J. (2000). Designing auditory displays for team environments. *Proceedings of the 5th Australian Aviation Psychology Symposium (AAvPA)*, Manly, Sydney, New South Wales, Australia. November 20-24. Retrieved June 29, 2007, from <http://www.itee.uq.edu.au/~cerg/publications/AAvPsy2000-WatsonSandersonAnderson.pdf>
- Wickens, C. D., & Dixon, S. R. (2005). *The benefits of imperfect diagnostic automation: A synthesis of the literature* (Technical Report AHFD-05-01/MAAD-05-1). Savoy: University of Illinois, Aviation Human Factors Division. Retrieved June 29, 2007, from <http://www.humanfactors.uiuc.edu/Reports&PapersPDFs/TechReport/05-01.pdf>
- Wickens, C. D., & Xu, X. (2002). *Automation trust, reliability and attention* (Technical Report AHFD-02-14/MAAD-02-2). Savoy: University of Illinois, Aviation Human Factors Division. Retrieved June 29, 2007, from <http://www.humanfactors.uiuc.edu/Reports&PapersPDFs/TechReport/02-14.pdf>
- Wilson, E., Wilson, D., & Jha, P. D. (2007). *Airport traffic control tower lighting and viewing measurements* (DOT/FAA/TC-07/09). Atlantic City International Airport, NJ: FAA, William J. Hughes Technical Center.
- Xiao, Y., & Seagull, F. J. (1999). An analysis of problems with auditory alarms: Defining the roles of alarms in process monitoring tasks. *Proceedings of the Human Factors & Ergonomics Society 43rd Annual Meeting*, 256-260.
- Xiao, Y., Seagull, F. J., Nieves-Khouw, F., Barczak, N., & Perkins, S. (2004). Organizational-historical analysis of the failure to respond to alarm problem. *IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and Humans*, 34, 772-778.
- Xing, J. (2006a). *Color and visual factors in ATC displays* (DOT/FAA/AM-06/15U). Oklahoma City, OK: FAA, Civil Aerospace Medical Institute.

- Xing, J. (2006b). *Reexamination of color vision standards. Part III: Analysis of the effects of color vision deficiencies in using ATC displays* (DOT/FAA/AM-06/11). Oklahoma City, OK: FAA, Civil Aerospace Medical Institute.
- Xing, J. (2007). *Developing the Federal Aviation Administration's requirements for color use in ATC displays* (DOT/FAA/AM-07/10). Oklahoma City, OK: FAA, Civil Aerospace Medical Institute.
- Zabyshny, A. A., & Ragland, D. R. (2003). *False alarms and human-machine warning systems* (UCB-TSC-RR-2003-07). Berkeley: University of California, Traffic Safety Center. Retrieved August 15, 2007, from <http://repositories.cdlib.org/its/tsc/UCB-TSC-RR-2003-07>

Acronyms

ACD	ARTS Color Display
ADS-B	Automatic Dependent Surveillance - Broadcast
ARTCC	Air Route Traffic Control Center
ARTS	Automated Radar Terminal System
ATC	Air Traffic Control
ATCT	Airport Traffic Control Tower
ATO	Air Traffic Organization
ATO-P	Air Traffic Organization – Operations Planning
CA	Conflict Alert
DataComm	Data Communication
ERP	Engineering Research Psychologist
FAA	Federal Aviation Administration
FABE	False Alert But Early
HBL	Hit But Late
HCS	Host Computer System
HFREG	Human Factors Research and Engineering Group
IOT&E	Independent Operational Test and Evaluation
IRIG-B	Inter-Range Instrumentation Group-B
MCI	Mode-C Intruder
MSAW	Minimum Safe Altitude Warning
NAS	National Airspace System
NTSB	National Transportation Safety Board
OE	Operational Error
QA	Quality Assurance
SDT	Signal Detection Theory
SME	Subject Matter Expert
SMON	ARTS Monitor
STARS	Standard Terminal Automation Replacement System

TCAS	Traffic Alert/Collision Avoidance System
TRACON	Terminal Radar Approach Control
URET	User Request Evaluation Tool
UTC	Coordinated Universal Time
VFR	Visual Flight Rules
VNTSC	Volpe National Transportation Systems Center
WAV	Digital audio file in WAV format