5-31-2023

# Creating Usable Research for the Design and Evaluation of Flight Deck Systems and Human Interfaces

Divya C. Chandra

# CREATING USABLE RESEARCH FOR THE DESIGN AND EVALUATION OF FLIGHT DECK SYSTEMS AND HUMAN INTERFACES

Divya C. Chandra
United States Department of Transportation (USDOT) Volpe Center
Cambridge, MA

This paper offers advice to researchers who want their research to be used by regulatory and industry practitioners to design and evaluate flight deck systems and their human interfaces. First, I present a few examples of success and review existing guidance. Next, I explain a design-thinking paradigm that views the design of a research study as a product. In this context, the users of the research (i.e., the product) are the practitioners. Researchers can smooth the path from research to practice by using this paradigm.

Many researchers in aviation human factors are motivated to make an impact in the real world. We do our best work with good intentions, thoughtful studies, and thorough analyses, managing hurdles along the way. Then we are surprised and disappointed to find that people who are in a position to apply our data to real systems do not do that, as though the results are not useful. How can we anticipate and prevent this scenario and, instead, smooth the transfer of research to practice? The benefits would be significant because we could improve safety, justify the investment in research, and perhaps even amplify the scope and longevity of its impact.

Here I share insights on how to create research for the flight deck that is usable and useful, not just to a pilot, but to potential users of the research results who are practitioners in industry (e.g., avionics manufacturers) and regulatory organizations (e.g., the Federal Aviation Administration, FAA, and other international aviation authorities). The principles are generalizable, but here I focus on the design of research aimed at improving flight deck systems, including their human-system interfaces. This is just one of several research needs for aviation practitioners. The same reasoning can be applied to other research needs (e.g., how to make flight operations more efficient while maintaining safety, or how to ensure that pilots have the right training and procedures to use existing flight deck systems effectively).

Human factors researchers working on flight deck systems typically see the pilot as the end user, but that is true only at one level. At another level, researchers can treat the *design* of a research study as a usability problem itself, where practitioners are the users of that product. When the design of the research study is seen as a *product*, it clarifies what steps researchers can take to help ensure a path to practice. I begin by presenting examples of research that have successfully impacted real flight deck systems. I also review advice on this subject from the FAA. Then I show how to apply a design-thinking paradigm to this problem.

## Examples of Success and Existing FAA Research Guidance

The FAA Aviation Safety Office (AVS) is one main sponsor and user of flight deck human factors research. This office posts regulations and policy related to flight deck human factors issues within the FAA AVS website. Some research studies that have transferred to flight deck systems are listed on the FAA NextGen Human Factors Division website. For example, multiple studies done by the Civil Aerospace Medical Institute supported the development of

FAA regulations for low-visibility operations and use of Enhanced Flight Vision Systems (EFVS). Be aware, however, that regulatory and industry documents do not necessarily cite research reports; citations are not a requirement for successful transfer of research to practice.

Yeh et al. (2016) elaborates on what the FAA needs from research products (Appendix B, pp. 283-292). The report describes the roles, responsibilities, and background experience of FAA personnel who may use research products (Table B-1, p. 283). It also describes various uses of research, such as direct support of the design/evaluation of systems. An important point is that, to be legally enforceable, FAA *minimum* requirements for equipment must be tied to a specific regulation or to a Technical Standard Order (TSO) for specific components (e.g., avionics equipment). This may be unintuitive to researchers who try to improve systems, not just meet minimum standards.

Yeh et al. (2016) lists four categories of research needed by the FAA, with examples of exemplary products. Three of the four categories cover the synthesis and consolidation of existing knowledge: developing checklists for system evaluations, developing recommended requirements and guidelines (which may form a basis for the checklists), and industry reviews that help the FAA to understand how the market might drive approval needs. The fourth category focuses on experiments related to flight deck systems. Yeh et al. references Zuschlag, Chandra, & Grayhem (2013) as a positive example of such research. This study compared different symbol-fill options for flight deck displays of traffic information to examine the effect of symbol design on rating of threat by pilots. The report has clear research objectives and makes a direct connection to FAA guidance documents. The work was documented in both a full government report (Zuschlag et al., 2013) and a short paper presented at an engineering conference (Zuschlag, Chandra, & Grayhem, 2011). Both versions are publicly accessible online.

## Design-Thinking Paradigm

Norman (2013) points out that design thinking is central to all innovation regardless of the domain. There are four main activities in human-centered design: observation, idea generation, prototyping, and testing. Design thinking also emphasizes iterative progress; the activities reoccur in a cycle. Table 1 below translates each design activity into activities for the design of human factors research to support practitioners. While the analogy does not transfer perfectly, the paradigm does remind researchers of helpful steps. I add one step to this list: communication of the results. Each section below takes a closer look at these activities.

### Observation: Understanding the Practitioner

Understanding the needs of the user is a standard human-centered design activity. Here, the user is a practitioner who can apply human factors research. Regulatory practitioners regulate or evaluate flight deck systems. Industry practitioners might represent an operator or be involved in developing software or hardware to support flight deck tasks. Industry systems may be subject to review by regulators. All practitioners care about safety and minimum standards.

Table 1.
*Applying a design-thinking paradigm to the creation of research to support practitioners.*

| Design Activity | Application to Design of Research |
|---|---|
| Observation | Understand and identify what practitioners know, what they do, and what they need. |
| Idea Generation | Generate multiple possible technical approaches and research methods. |
| Prototyping | Evaluate different possible technical approaches and methods through thought experiments. Imagine what the results might be and discuss the potential patterns of results with practitioners. Assess whether the different patterns would affect decisions or not. |
| Testing | After running the study, explain the results to a variety of stakeholders. Listen to the feedback and questions. Refine the takeaway messages by addressing criticisms and limitations of the work. |
| Communication | Present results clearly and succinctly, focusing on the practitioners' needs. |

Many practitioners are not trained researchers and instead have a different knowledge base. Practitioners might be trained as engineers or pilots. Most are not trained formally on human factors subjects. Some field evaluators might have trained in the military or may not be college graduates. Despite the knowledge differences, both regulators and industry practitioners are familiar with FAA policies and guidance documents, and industry guidance. They understand the capabilities of the technology, how it may evolve, and how it would be used. Yeh et al. (2016) provides more detail on the different types of FAA users of human factors research.

Regulators may use human factors research to update regulatory documents including policies and guidance to help evaluate systems. Different types of evaluations happen at different levels of system maturity (e.g., bench tests in a laboratory setting, flight-simulator tests, in-flight tests, and operational implementation). FAA and industry practitioners may use human factors research to develop standards and recommendations (e.g., voluntary industry standards, or standards that are cited by as a possible means of compliance by FAA).

Flight deck systems are typically evaluated by different parts of the FAA's Office of Aviation Safety. In brief, the Aircraft Certification Service approves that the equipment works "as intended" and does not interfere with other equipment. (See Yeh et al., 2016, for an introduction to the concept of "intended function.") The Flight Standards Service reviews the equipment from the perspective of the human operator; it approves the use of the equipment by the flightcrew. The operational approval establishes the crew training necessary to operate the equipment under various scenarios, including full or partial failures.

The goals of practitioners become clearer in the context of their job responsibilities and demands. For example, they may have to ensure that the system meets minimum standards for the intended function in (sometimes messy) real-world situations. They consider all conditions under which that system may be used in the context of the full flight task. In contrast, researchers often prefer to study a limited set of conditions to isolate the issues of interest. For example, pilots constantly switch between different flight deck systems. How are human interfaces that support different functions integrated within the limited physical (and screen) real estate of the flight deck? In turbulence, or with smoke in the flight deck? These are real conditions that practitioners consider, which researchers might not.

Practitioners are less able to take the time to understand details that may be important to researchers (e.g., about the study design or statistical analysis). They appreciate short documents (or just Executive Summaries) that highlight key points, especially if these writeups make direct connections to regulatory and guidance documents that they use. Human factors researchers can help practitioners absorb key messages from research studies quickly and effectively so that they know how, when, and whether to apply the results of the research.

**Idea Generation: Designing a Technical Approach**

Researchers have different goals for applied research. They might test a system to determine if it (a) could work *at all*, (b) could work *better than* an existing design, either in terms of efficiency or functionality, or (c) meets guidelines or minimum standards. Option (c) is important to regulators but unrelated to optimization or efficiency. Industry researchers may also be interested in conducting the types of studies suggested by options (a) and (b).

The study needs to address a problem statement with a clear scope, clear motivation, and clear purpose for the results. Researchers benefit from the engagement of practitioners in developing the problem statement. Together they can refine and craft an initial proposal. For example, how will the data inform specific government policies or industry products? Sometimes problem statements need to be broken down into reasonable steps towards addressing a bigger issue. Sometimes initial problem statements are overly specific and could promote studies that are not generalizable. As with other design exercises, developing a clear problem statement is an iterative process.

When designing a study for practitioners, researchers should be creative and flexible, considering a range of options. Often, researchers have specialized knowledge of some methods and less expertise in others, but it is important to customize the research method for the problem the practitioner needs to address. The researcher should be willing to apply unfamiliar or even novel approaches. They should consider the full range of data that could be gathered, including subjective data, objective data, quantitative data, and qualitative data. They should even consider whether an existing data set could produce useful insights with further analysis. Collecting data on human performance is important (because we know that preference does not equal performance), but we need to collect data that are most relevant to the practitioners' goals.

For research to transfer to practice, data should have face validity and operational significance. One means of increasing face validity is for study participants to be realistically representative of the real users (e.g., airline pilots) as well as possible within cost constraints. Operational significance can be affected by the level of task and equipment fidelity; researchers should consider various levels carefully to maintain operational relevance while staying within budget. Sometimes less expensive, lower fidelity approaches are better because they focus on the problem statement at hand. For example, using a flight simulator may introduce distractions that draw attention away from the task of primary interest (e.g., using aeronautical charts).

**Prototyping: Assessing Potential Research Methods**

Once researchers have sketched out multiple study approaches, they should "test drive" these approaches with thought experiments before finalizing the specific methodology. In a thought experiment, the researcher hypothesizes all potential patterns of results and determines, in consultation with the practitioners, how each of the patterns might affect practitioner decisions. If none of the possible result patterns would change what the practitioner does, then that approach is not useful. Sometimes, practitioners unknowingly advocate for research approaches that, in the end, would not affect their policy/evaluation decisions (e.g., requesting use of a flight simulator when it may not be necessary). Researchers may need to sketch out all the possible results and their lack of impact on decisions to demonstrate that those studies are not worth doing. One good strategy is to ensure that there are lessons to be learned regardless of the pattern of results. Perhaps different result patterns would support different recommendations and considerations, if not specific policies and guidance.

Sometimes thought experiments reveal weaknesses in the problem statement or the research method. If it is not easy to hypothesize potential results and interpretations, then the problem statement needs to be clarified. Keep revising it until you can state what practitioner need(s) will be addressed and how the data collected will be used. It is also useful to document and critique methods that were considered but discarded. For example, why did that method not meet the needs at the time? What would have to change for that method to be useful?

**Testing: Refining the Research Takeaways**

After the research results are in, but before takeaways are finalized, socialize the study and its interpretation. Present the study even as the analysis is in progress. Present to a variety of stakeholder audiences. Gather feedback on preliminary takeaways, i.e., conclusions, highlights, and recommendations. This is especially important if the results connect to guidance documents because guidance needs to be clear and acceptable to many audiences. The draft takeaways, and even the analyses if needed, should be enhanced based on stakeholder questions and feedback.

Recommendations are a particularly important type of takeaway. Well-constructed recommendations are easy to read, actionable, they make sense, and are believable (i.e., have face validity). They are clear about their scope and limitations. Maximize the detail in the recommendation without going beyond what the data support. Flight deck human-interface recommendations should be traceable to their impact on pilot tasks. Ideally, recommendations should converge with findings from other sources (e.g., data from other studies or industry working group discussions).

However, keep in mind that recommendations from research do not always give specific *solutions* to flight deck problems because they do not take into account all the constraints for the problem. Let the appropriate stakeholders (not researchers) determine who implements a solution and how. Industry practitioners are partners; they want to understand what the problems are and the principles and rationale that move towards a solution, without being overly constrained. Industry is happy to use insightful findings that they can tailor to their situation.

**Communication: Documenting the Study**

Once key takeaway points are settled, researchers must communicate them effectively to practitioners. The communications should be brief, to the point, clear, and public. Researchers typically document studies for other researchers who may want to replicate a study, but practitioners just want to know why the results matter. The most useful elements of a report for practitioners are the Executive Summary and recommendations. An Executive Summary summarizes the study and key takeaways in one page ideally, or a few pages at most. It succinctly explains the purpose of the research and the main results relevant to regulatory policies or industry products. These points also should stand on their own because, realistically, they might be the only aspects of the research that circulate widely among practitioners.

Researchers also need to critically assess and acknowledge the limits of their study. An honest assessment will greatly help practitioners determine when and whether to apply the results in practice, improving trust between researchers and practitioners.

## Summary

The generalized human-centered design process is iterative and combines observation, idea generation, prototyping, and user testing. I add one final step to this list, to communicate the results and fine tune the takeaway messages. To create usable research, treat the research as a product that will be used by a practitioner.

Although flight deck human factors research problems will change, this design paradigm for creating usable research will apply in general. Cooperative and open dialogue between practitioners and researchers is necessary. Researchers need to do their homework to learn about practitioners' needs, and they need to be creative and willing to learn new ways to think about the problem. The overall goal is to realize the safe and effective use of flight deck equipment in operations. Design thinking will lead the way.

## Acknowledgements

The views expressed herein are those of the author and do not necessarily reflect the views of the Volpe National Transportation Systems Center or the United States Department of Transportation (USDOT). I would like to thank Bill Kaliardos, Tracy Lennertz, Sherry Chappell, Janeen Kochan, Don Fisher, Maura Lohrenz, and Stephen Popkin for their support and review of early drafts of this paper. This work was funded by the USDOT Volpe Center.

## References

Norman, D. (2013) The Design of Everyday Things. Basic Books.

Yeh, M. Swider, C., Jo Y.J., and Donovan, C. (2016) Human factors considerations in the design and evaluation of flight deck displays and controls. Version 2.0 DOT/FAA/TC-16/56. https://rosap.ntl.bts.gov/view/dot/12411

Zuschlag, M, Chandra, D.C., and Grayhem, R. (2013). The Usefulness of the Proximate Status Indication as Represented by Symbol Fill on Cockpit Displays of Traffic Information, DOT-VNTSC-FAA-13-03; DOT/FAA/TC-13/24. https://rosap.ntl.bts.gov/view/dot/9982

Zuschlag, M., Chandra, D.C., & Grayhem, R. (2011). The Use and Understanding of the Proximate Status Indication in Traffic Displays. *Proceedings of the 30th Digital Avionics Systems Conference*, 16-20 October 2011, Seattle, Washington. https://rosap.ntl.bts.gov/view/dot/9513