2015

Techniques for the Human Centered Evaluation of Designs for the Future Aviation System

Philip J. Smith

Kathy Abbott

Lawrence J. Prinzel

Amy Pritchett

Tanya Yuditsky

TECHNIQUES FOR THE HUMAN CENTERED EVALUATION OF DESIGNS FOR THE FUTURE AVIATION SYSTEM

Philip J. Smith
The Ohio State University
Columbus OH

Kathy Abbott
Federal Aviation Administration
Washington D.C.

Lawrence J. Prinzel
NASA Langley Research Center
Hampton VA

Amy Pritchett
School of Industrial and Systems Engineering, Georgia Institute of Technology
Atlanta GA

Tanya Yuditsky
FAA William J. Hughes Technical Center
Atlantic City NJ

In order to evaluate new operational concepts, system designs, procedures and technologies for the future aviation system, we need to develop and validate a range of techniques to ensure the safe and effective performance of human-machine systems. This becomes increasingly important as such systems incorporate increasing levels of automation and autonomy for technologies, and as they attempt to integrate increasingly complex subsystems. It is challenging to evaluate the individual components of such systems relative to meeting their design requirements. It is orders of magnitude more challenging to evaluate performance when they are embedded in the larger system context. While there is no perfect method for such an assessment, a number of complementary techniques have been developed, applied and evaluated and will be discussed. Some can be applied early in the design process, while others focus on assessment as a system has been released for field trials or actual operations.

One theme of this panel is the need to apply a range of techniques over the development life cycle for a new system or subsystem in order to increase comprehensiveness and provide converging evidence. Methods across this range are outlined below, using examples from concrete aviation systems to help communicate the nature of the assessment methods and their actual use.

A second theme is the need to understand the strengths and weaknesses of such methods, individually and together, addressing questions such as:

- What is the state of the art?
- How good is it?
- What are the weaknesses of each individual method?
- When are they practical?
- What are the barriers to their use?

A third theme focuses on how to get better: What are the most promising directions for further developing our repertoire of techniques for verification and validation of human-machine systems, not just at the individual level but at the level of the complex, distributed work systems in aviation with a wide range of embedded technologies and forms of "automation"? Below is a summary of the topics to be addressed in this panel discussion.

Evaluating Design-Related Pilot Error

In 2013, the Federal Aviation Administration published a new regulation that requires evaluation of new aircraft flight deck systems/equipment for design-related pilot error. Good design standards must be applied as described in the regulation. In addition, the regulation recognizes that even well qualified pilots using well designed systems will make errors, so the systems/equipment designs must incorporate means to enable the pilots to manage those errors. The extent to which the system design needs to be evaluated depends on the novelty, complexity, and level of integration of the systems/equipment. Thus, discussion of future designs need to be framed in terms of regulations, methods used for complying with them, and challenges in applying them.

NASA Research Techniques for Future Aviation Systems: The Case of Synthetic and Enhanced Vision Systems

The NASA Synthetic and Enhanced Vision System (S/EFVS) is one of the enabling technologies that can provide additional margins of safety and aircrew performance in low-visibility surface, arrival, and departure operations. This work provides a case study of research techniques often employed in NASA human factors research.

Synthetic Vision Systems (SVS) use terrain/obstruction databases to present a computer rendered view of the outside world, often on a Head-Down Display (HDD). Enhanced Flight Vision Systems (EFVS) use real-time sensor input to present an enhanced visual image of the outside view on a Heads-Up-Display (HUD) or "equivalent" display, such as a Head-Worn Display (HWD). Synthetic Vision Systems (SVS) use terrain/obstruction databases to present a computer rendered view of the outside world, often on a Head-Down Display (HDD). Enhanced Flight Vision Systems (EFVS) use real-time sensor input to present an enhanced visual image of the outside view on a Heads-Up-Display (HUD) or "equivalent" display, such as a Head-Worn Display (HWD).

Research on such systems has provided an opportunity to study the use of a number of techniques, flight and simulator assets and resources, and newly developed and/or non-traditional aviation human factors approaches for evaluating new aviation technologies and systems based on research techniques often employed in NASA human factors research. The panel presentation shall outline the various methodological approaches taken to evaluate NASA SVS and EFVS technologies.

Preventing Human Factors Problems Early in Design

Too often, human factors concerns are latent within a design because of some aspect of the underlying concept of operations. Thus, human in the loop testing late in the design cycle may find that a decision made early in the design cycle will lead to, for example, a workload spike where the pilot must quickly execute a large number of key presses to respond to unexpected air traffic controller instructions, or a situation where the pilot performing interval management will need to continuously monitor a task during already-high-workload phases of an arrival and approach. At these late stages in the design and implementation cycle, such human factors issues are often labeled as problems in the interface or with training, even when their genesis is more fundamental in the design.

Thus, it is important to consider how we can examine, early in design, what the fundamental impacts on workload, teamwork and information requirements will be in response to a new concept of operation, to new function allocations between humans and automation and/or between air and ground, and the implementation of new technologies. In particular, at the early stages of design, our models should not seek to predict what a human operator *will* do, but instead should first be checking for what the new design *will ask* the human operator to do. Particularly in the dynamic contexts inherent to aviation, this analysis needs to include computational fast-time simulation to predict when tasks will be demanded of the human operator. Such analyses can then highlight to all the designers involved where the concept of operation or underlying technological functions need to be changed.

Human Factors in the Wild

Traffic Managers continually evaluate the future status of the National Airspace System and make decisions that greatly impact its efficiency. Future systems for Traffic Flow Management will provide increased support to drive those decisions to be more precise in where they affect traffic flows and by how much. Understanding the decisions that are made today and what drives them is critical to the design of future systems, but this has been quite challenging. We typically learn about today's processes by conducting "Human Factors in the Wild:" we go to operational facilities and observe the experts in their natural habitat. This works well when we are counting steps or key presses, but not as well for deconstructing decisions. The environment is so dynamic, the options so varied, and so many factors are in play, that deconstructing the decision process becomes messy. So much of the decision making occurs in the Traffic Manager's head that observation alone is not sufficient. We propose using a modified observation approach where a Subject Matter Expert is part of the observation team and provides an interpretation of what drove the subject's decision. A version of this technique has been used successfully for identifying the drivers of operational errors by having subject matter experts review replays of traffic scenarios.

Structured Knowledge Elicitation to Envision the Impact of Future Designs

There are a variety of complementary approaches to identify potential issues human-automation design concerns associated with integration of some new component with the broader aviation system. Some involve computer modeling and some involve empirical testing or observation.

Another approach is to take advantage of the knowledge of a team of human experts to envision potential safety critical scenarios. This approach focuses on knowledge elicitation from Subject Matter Experts (SMEs) to predict potential incidents or accidents by developing scenarios where the automation embedded within some new technology could contribute to incidents or accidents.

There are several important features defining this method: First, a sequence of stages is used to progressively expose the SMEs to different types of prompts to help stimulate scenario generation. These stages use probes that are increasingly more detailed and suggestive. In the first stage, only nominal scenarios (success stories) are presented in order to avoid any biasing of the SMEs as they generate scenarios. In addition, the SMEs work individually in order to avoid having one SME influence the scenario generation by another.

In the next two stages, increasingly specific probes are presented to stimulate additional ideas for scenarios. The first set of probes uses fairly general categories from the Threat and Error Management literature; the second set provides very specific prompts for the SMEs to consider in generating scenarios, based on system design features and cognitive processes such as the potential impacts of:

- alarm prioritization
- autonomous mode changes
- inadequate knowledge of intent
- slips (errors of omission and commission).

The fourth stage finally brings the SMEs together in a focus group (individuals with relevant operational experience, human factors experts and experts in the underlying technology for the human-automation system of interest) and asks them to work together to identify additional critical scenarios. This focus group uses a variety of structured probes as well, including the presentation of historical accidents and abstract characterizations of these accidents in terms of contributing factors.

The end result is a very concrete set of scenarios predicting potential incidents or accidents for consideration by system designers.