

THE INSERTION OF HUMAN FACTORS CONCERNS INTO NEXTGEN PROGRAMMATIC DECISIONS

Bettina L. Beard, NASA Ames Research Center
Jon Holbrook, San Jose State University
Rachel Seely, FAA Human Factors Division

Since the costs of proposed improvements in air traffic management exceed available funding, FAA decision makers must select and prioritize what actually gets implemented. We discuss a set of methods to help forecast operational *and* human performance issues and benefits before new automation is introduced. This strategy could minimize the impact of politics, assist decision makers in selecting and prioritizing potential improvements, make the process more transparent and strengthen the link between the engineering and human factors domains.

INTRODUCTION

The Federal Aviation Administration (FAA) is leading an effort to streamline U.S. air traffic operations within a campaign referred to as NextGen. NextGen emphasizes the introduction of automated decision support for air traffic controllers and traffic managers. To realize NextGen, FAA stakeholders, such as program planning and implementation teams, must consider and make important trade-offs between multiple factors.

From the viewpoint of human factors practitioners, the introduction of automation into a cognitively demanding profession raises several red flags, including:

1. Unintended or unanticipated consequences that may result if decision support is not appropriately designed and introduced (Norman, 1990),
2. Limitations of human-in-the-loop (HITL) simulations are often misunderstood by decision makers and therefore underestimated,
3. User acceptance of technology observed in simulation and shadowing studies can be a poor predictor of user acceptance in actual operations (e.g., pFAST), and
4. At times, technologies under development are solutions in search of a problem.

For tower operations alone, the FAA has proposed 85 operational improvements (OI) and increments (<https://nasea.faa.gov>). In addition, with the progression of time, reduced budgets and pragmatic limitations in hardware and software maturity have produced an inconstant trade-space. This creates a multidimensional trade-space that is extremely difficult to manage. From the viewpoint of program management, just getting the system built becomes the number one priority. Human factors concerns are considered less pressing, particularly when the concerns are presented in an obscure fashion.

The question is, how can we, the human factors community, first, get our arms around the complex HF issues so that we may then provide defensible input to NextGen prioritization and program decisions and, second, provide input that is consistent with their cost/benefit analyses. Here we introduce the feasibility of an approach for clarifying and quantifying the operational and human performance merits or risks of potential NextGen improvements to help stakeholders make these difficult decisions.

PRACTICE INNOVATION

In this section, we describe four methods for obtaining data that address the four human factors issues raised in the introduction. In the Findings section we will describe strengths and weaknesses of our approach.

1. Unintended or unanticipated consequences

Experts in the field of aviation human factors were recruited to provide input into an on-line survey. These experts rated each NextGen OI (or related-OI grouping) on its predicted impact across 23 human performance metrics (refer to Table 1).

1. Individual SA, Level 1: Perceiving
2. Individual SA, Level 2: Comprehending
3. Individual SA, Level 3: Projecting
4. Sensory Information Acquisition
5. Training
6. Opportunity for Error
7. Error Detection
8. Recovery from error
9. Retrospective memory
10. Prospective memory
11. Skill retention
12. Monitoring
13. Mental effort
14. Physical Effort
15. Trust in automation
16. Decision-making
17. Communication
18. Coordination/Collaboration (human-human)
19. Coordination/Collaboration (human-automation)
20. Work Flow
21. Time to Perform Tasks
22. Assignment of Roles and Responsibilities
23. Clarity of Roles and Responsibilities

Table 1. A group of human factors experts derived a list of potentially important human performance metrics. The experts then rated the potential impact of NextGen improvements on each of the 23 metrics.

Ratings were made on a bipolar scale ranging from strong negative impact to strong positive impact. Although the human factors experts were knowledgeable about NextGen solutions, we provided training material about the proposed solutions to ensure a baseline level of knowledge. We also

provided the opportunity for the experts to qualify or expand on their ratings, to discuss lessons learned implications, to identify affiliated research questions and to discuss as-yet unpublished research.

2. Limitations of human-in-the-loop simulations can be underestimated

Numerous Human-In-The-Loop (HITL) concept and technology evaluations have been performed to assess NextGen proposed solutions. We critiqued the methods, results and conclusions of articles published by the U.S. government, industry, academia and Europe. For each publication, at least two researchers coded the sample size, simulation fidelity, dependent and independent variables, demonstrated efficiencies and risks of the new tool, study weaknesses and future research needs. Our main interest was to determine whether the research design and data analysis were experimentally rigorous and the conclusions valid (Beard, 2012).

3. User acceptance

Visits to air traffic control facilities reveal cases where certain pieces of automation are not being used as intended. These new roles range from only one function out of a suite of functions being utilized to banishment under the console. Because NextGen decision support tools will be highly co-dependent, the miss-use or dis-use of any sub-system can make the entire system fail. For this reason, it is critical to understand the sub-system dependencies and to obtain buy-in and advice from the user community. All mid-high fidelity HITL research includes air traffic controllers in the subject pool. What is missing is input from the larger ATC community.

We developed a brief, web-accessible survey (see Figure 1), that asked individuals in the tower controller community to rate the potential impact of NextGen solutions across five metrics: impact to their job, to airport safety, airport efficiency, airport capacity, flexible operations and predictable operations (Holbrook, Parke, Oyung, Collins, Gonter & Beard, 2013). These ratings were made for the seven proposed NextGen capabilities and three broad enablers listed below.

Seven Proposed NextGen Capabilities

- Departure Metering at the Ramp
- Taxi Routing and Scheduling
- Departure Runway Assignment
- Runway Scheduling
- Departure Flow Management
- Integrated Arrival/Departure Scheduling
- Runway Configuration Management

Three Enablers

- Enhanced Surveillance
- Electronic Flight Data Displays
- Data Communications

In addition, we compared the broad community results to in-depth interviews with two highly experienced controllers who have an extremely deep understanding of proposed NextGen OIs.

4. Identifying if Proposed Solutions fix today's problems

Early identification of OI applicability to existing problems supports investment decisions by providing criteria for prioritizing OIs based on the specific operational challenges they address. Safety issues were identified through an analysis of 200 Aviation Safety Reporting System (ASRS) incident reports over a five-year period (Holbrook, Stasio, McDonnell, Puentes, Jobe & Beard, 2011).

FINDINGS

In this section, we discuss the pros and cons of the methods used to address the four human factors issues and list lessons learned about our implementation of each method.

1. Unintended or unanticipated consequences

We found a very high attrition rate. The survey of human factors experts was much too long. Only the most patient survey participants completed the entire survey. The ratings themselves provided quantitative insights into potential human factors issues. However, the most informative results came from the descriptive discussions and clarifications provided by the survey participants. The identification of critical linkages between the 23 human performance metrics underscores the importance of research addressing some metrics in concert to understand how one metric trades with another. A few participants commented that it was difficult to rate the capabilities without a firm concept in mind.

Lessons Learned. (1) Prior to survey distribution, hold a workshop devoted to discussions about potential concepts. (2) Provide an incentive to survey participants to increase the likelihood of survey completion.

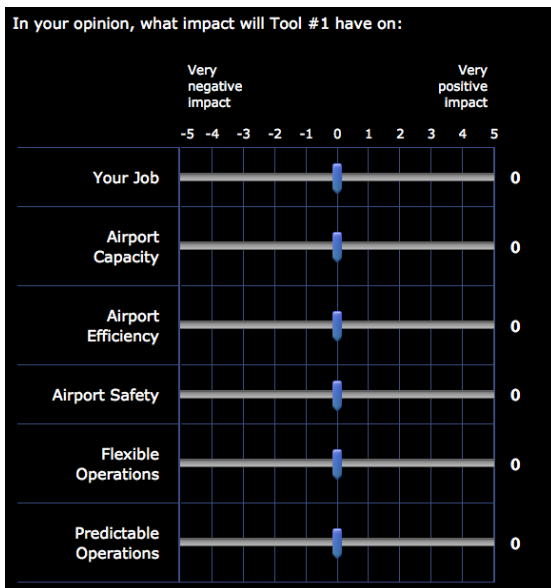


Figure 1. Screen shot of the web-based survey. Controllers rated each of seven proposed NextGen capabilities (that were combinations of related OIs) across five metrics (e.g., airport capacity). Controllers used the slider to make impact ratings for each capability.

2. Evaluation of published literature

We found that the amount of time required for less experienced researchers to review each article was untenable. On the other hand, seasoned researchers could evaluate a given publication within a few hours. A meta-analysis can be used to identify the critical factors impeding the development of implementable tools and to identify further issues that may be fragmenting the tool development process.

Lessons Learned. (1) It takes a discerning and well-trained eye to appropriately evaluate the results of published research.

Less experienced researchers can be used to catalogue the variables of the experiment, but the critique should be performed by a very experienced researcher.

(2) Use only reviewers with no stake in the results of the analysis.

3. User acceptance

With considerable help from the controller union, we were able to obtain ratings from air traffic controllers across the nation and from large to smaller airports. Based on queries of controllers who chose not to take the survey, the controllers who chose to participate tended to be those who were more amenable to change.

Lessons Learned. (1) Provide an incentive to survey participants to increase the likelihood of survey completion.

(2) Be sure to include space for comments. Responses were dependent upon the controllers understanding of what automation can actually provide.

4. Identifying if Proposed Solutions fix today's problems

It is informative to identify how proposed changes address current issues. The analysis of ASRS reports underscored how NextGen solutions emphasize enhanced situation awareness. In addition, classes of errors were identified that are not addressed by NextGen. The FAA operational improvements focus heavily on the advancement of technology and procedures. They do not address concerns such as organizational culture that have been known to plague large institutions. A limitation of this assessment was that only safety-related issues could be identified.

DISCUSSION

Forecasts of the value of the NextGen automation solutions have predominantly focused on the main objectives and metrics proposed in the modernization plan (i.e., efficiency, capacity and safety). Here we provide a feasible set of methods that may be used by any organization developing automation for human use to forecast human performance and organizational issues and benefits before implementation. First, we gathered anticipated human factors issues from the broad human factors community couched within the framework used by the FAA (i.e., the operational improvements). Second, we found that a HITL meta-analysis can help program decision makers to better understand the degree of uncertainty in scientifically based conclusions and provides a basis for scoping future research. It is challenging to proactively evaluate systems before implementation. Our analysis involved a review of human-in-the-loop simulations of prototype systems; therefore it included an early, and

formal, instantiation of the proposed system. Because the instantiation is known, any modifications to the design over time can be incorporated into the analysis. Third, involving the broad user community can aid in the identification of issues that may not be revealed in isolated simulations or shadowing experiments. Finally, identification of how proposed changes address current issues can provide further knowledge about gaps in NextGen solutions.

Turochy (2001) summarizes methods for prioritizing potential improvements in the highway transportation domain. Similar to our approach, each method follows a rational procedure and includes both objective and subjective evaluations. They do not, however, incorporate human factors issues.

There have been several efforts to identify likely NextGen human factors issues (Sheridan, Corker & Nadler, 2006; Funk, Mauro & Barshi, 2009). The novel aspects of our approach are that we ground the issues within the language used by the NextGen engineers (i.e., operational improvements) and we include both the human and operational tradeoffs.

Our goals were to:

- Strengthen the overall FAA performance budget to include appropriate human performance metrics in the cost/benefit calculations,
- Provide immediate access to human performance indicators for better decision-making,
- Improve communication and collaboration between FAA modernization programs and human factors experts, and
- Rapidly produce budget publications personalized to the decision-makers priorities.

In line with what we are attempting here, Funk (2009) proposed a method to identify potential human factors issues in the NextGen flight deck. His method involves functional modeling, task analysis, human fallibilities analysis, and Failure Modes and Effects Analysis. Although of promising value, Funk's method does not incorporate operational cost/benefit estimates.

MITRE is working closely with the FAA to prioritize the operational improvements. Their methodology, while addressing both operational and some human performance concerns, provides a concrete prioritization. It does not permit the decision maker to weight the disparate inputs going into the analysis. We are currently building an interactive capability for FAA decision makers to weight the importance they want to instill on each analysis. For example, the decision maker may want to assign a higher weight to user acceptance than to whether the automation addresses contemporary problems in the national airspace system. The Performance Budget Tool we are now developing will help FAA decision makers prioritize the proposed improvements based on their own viewpoints.

Acknowledgements. The FAA Human Factors Division coordinated the research requirement and its principal representative acquired, funded, and technically coordinated execution of this research.

REFERENCES

Beard, B. L. (2012). Human-in-the-loop research supporting mid-term NextGen air traffic control tower operational

improvements. Acquired from:

<https://www.hf.faa.gov/hfportalnew/Search/SearchReport.aspx>

- Funk, K. (2009). A methodology and tools for the prospective identification of NextGen human factors issues. *Proceedings of the 15th International Symposium on Aviation Psychology*, Dayton, OH, April 27 – 30, 2009, pp. 106 – 111.
- Funk, K., R. Mauro, R. and Barshi, I. (2009). NextGen flight deck human factors issues. Published in Conference Proceedings of the *15th International Symposium on Aviation Psychology*, Dayton, OH, 208 – 213.
- Holbrook, J., Parke, B., Oyung, R., Collins, R., Gonter, K. and Beard, B. L. (2013). Perceived usefulness of planned NextGen tools by air traffic control tower controllers. Published in Conference Proceedings of the *16th International Symposium on Aviation Psychology*, Wright State University, Dayton, Ohio.
- Sheridan, T. B., Corker, K. M. and Nadler, E. D. (2006). Final report and recommendations on human-automation interaction in the Next Generation Air Transportation System. Cambridge, MA: US Department of Transportation, John A. Volpe National Transportation Systems Center.
- Turochy, R. E. (2001). Prioritizing proposed transportation improvements: Methods, evaluation, and research needs. *Journal of the Transportation Research Board: Transportation Research Record* 1777, 123-128.