Scenario-Based Task Design for Airline Pilot Anticipatory Behaviors

Asynchronous Assessment of Complex Cognitive Skills

Peterson, M., Baron, B., Cusano, L., Mumaw, R. J., Corry, P. C., Hoffman, D., & Billman, D. O.

anticipatory behaviors
aviation safety
cognitive skills training
complex systems assessment.
pilot monitoring
proactive safety thinking
resilience engineering
scenario-based assessment

In airline operations, pilots are asked to actively produce positive safety outcomes and to learn from their own and others' successes. They do this, in part, by monitoring conditions present on each flight and either reacting to or proactively planning for threats to airplane safety. However, there are few ways to assess these anticipatory and monitoring behaviors, and little is understood about how to train these complex cognitive skills. Pilots have historically learned these skills informally from peers or from personal experience. The current study seeks to both assess monitoring and anticipation and evaluate a short tutorial for advancing awareness of these skills, especially among early-career pilots. Scenario-based tasks were designed to assess these skills at multiple points along a flight path and

were tested in a pre-test – intervention – post-test design. The design and operationalization of these assessment tasks are described here.

Introduction

In 2023, 4.2 billion passengers flew on 35 million commercial flights, with a single fatal accident (ICAO, 2024). While this accident rate is unquestionably low, global aviation traffic is forecast to increase. Even with a low but steady accident rate, the increase in departures will necessarily result in more frequent accidents (Arbuckle et al. 1998). Therefore, the global aviation safety community continues to seek ways to improve safety outcomes.

Historically, approaches to improving flight safety involved investigating accidents and asking, "What went wrong?" This approach, which seeks to identify the human lapses which led to the mishap, has provided important safety improvements. However, focusing on failures has led to an emphasis on highlighting pilots as an impediment to safety, expressed in a statistic that human error is an important contributor in 80% of accidents (Rankin, 2008). More recently, a new perspective in aviation safety research has emerged, recognizing for every accident cited in the 80% statistic above, pilots successfully address 157,000 challenging inflight events (Holbrook, 2019). In other words, instead of focusing on "what went wrong," researchers now also seek to understand "what goes right" in the frequent times that pilots produce safe outcomes. This approach is proactive, defines safety as having as many things as possible go right, and views humans as a resource to providing flexible solutions to problems.

Hollnagel (2011) developed a framework for proactive safety thinking, or resilience, with four capabilities: anticipate, monitor, respond, and learn. The present study seeks to measure individual anticipatory behaviors through monitoring tasks. During flight, pilots monitor a constant stream of variables: the aircraft's attitude, speed, altitude, and position relative to the position it is expected to be in to comply with flight path assignments. These monitored variables help pilots anticipate the future state of the aircraft.

Until recently, much of the knowledge and skills related to resilience have been passed down informally between pilots while on the job (Baron et al., 2023). Unfortunately, these opportunities for learning have become rarer as pilots retire and their expertise leaves the system, a phenomenon that causes concern amongst researchers and regulators (GAO, 2023). While researchers would like to study resilience and anticipatory behaviors in the richly contextual environment of a flight simulator, simulators are extremely expensive and these skills are characteristically hard to assess (Neville et al., 2020; Rogers et al., 2023). This paper reports on the assessment methods developed for a study which trains complex cognitive skills for pilot anticipation and monitoring.

Methods

This research seeks alternative methods for training and assessment which can be used asynchronously and inexpensively. Given the cognitive complexity of the skills to be assessed, the assessment design portion of this project constituted a challenge. By the nature of their work, pilots as a group are hard to connect with in-person; thus, asynchronous methods were required. Web delivery was chosen to provide the necessary accessibility, despite the challenges of designing both training and evaluation for this environment. Assessment of performance in operationally relevant tasks is as important as assessing skills

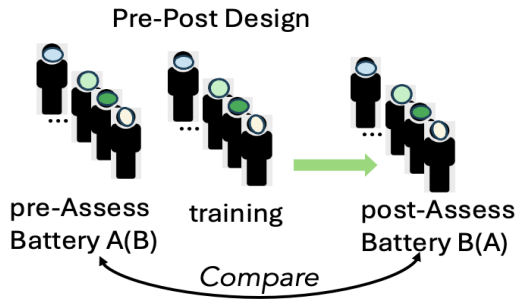
and knowledge; thus, focus groups or interviews were inappropriate. Rather, scenario-based tasks were needed. Each of these tasks is described in detail below.

Study Design

The current study used a pre-test – intervention – post-test design. The participants were airline pilots at US passenger carriers, all of whom flew Boeing 737s. Participants took a pre-test consisting of the measures described below, completed an online tutorial about anticipatory behaviors, and then completed a post-test a few days later (see Figure 1). The pre- and post-test items were designed to match for both difficulty and the knowledge content or skills assessed. Participants completed all activities at a time and place of their choosing.

Figure 1

Study Design



Task Design

There were three major types of scenario-based tasks designed for this study:

- Generative tasks: either a pre-descent briefing item list or a list of potential impacts when confronted with a change to the arrival path.
- Review tasks: pilots reviewed researcher-generated lists of briefing items and action plans.
- Arrival chart analysis tasks: pilots identified potential challenges on simplified charts.

For the generative and review tasks, two real-world airports were selected for the scenarios, Raleigh-Durham and Oklahoma City. These airports were chosen because they contained potentially difficult segments in the flight path, had two or more potential approaches, and had points where it was possible to anticipate a challenge and plan ahead. These scenarios involved a high tailwind that caused ATC to change the runway for landing.

For the analysis task, simplified versions of real-world arrival charts were created and modified slightly to provide specific challenges for the participants to identify. These challenges included short distances to descend a certain amount (i.e. steep sections), high terrain, and potential shortcuts, which could cause high speeds on arrival. For the purposes of this study, a steep section is one which is close to the heuristic of a plane needing at minimum 3 miles to descend 1000 feet.

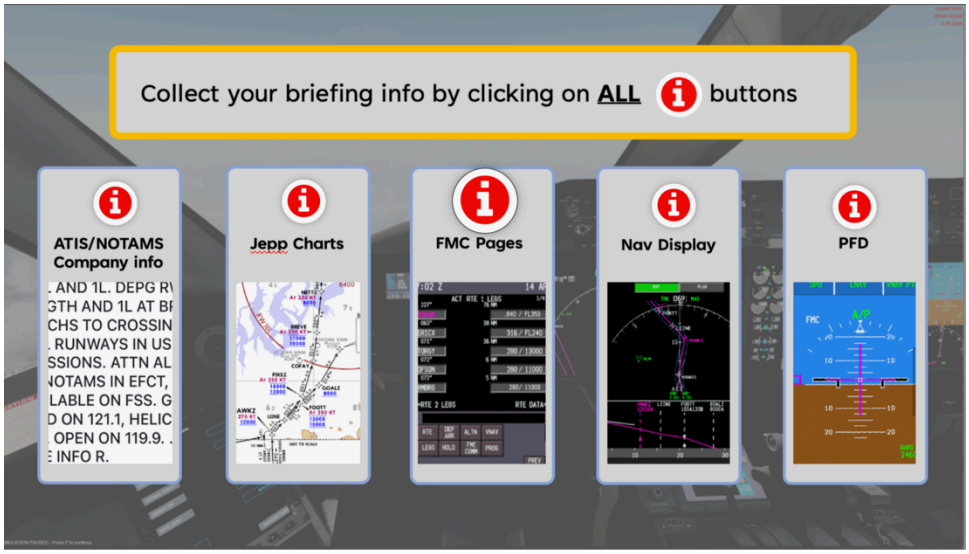
Generative Tasks

The generative tasks consisted of two parts: the generation of a pre-descent briefing, and then a reactive impact identification in response to a communication from air-traffic control (ATC) changing the flight parameters, such as the landing runway. These tasks are routine activities on every flight. These briefings and plans are historically prescribed by standard operating

procedures. Participants were given the information they needed to prepare both the briefing and the plan in an online interactive format (see Figure 2).

Figure 2

Interactive Flight Deck with Scenario Information



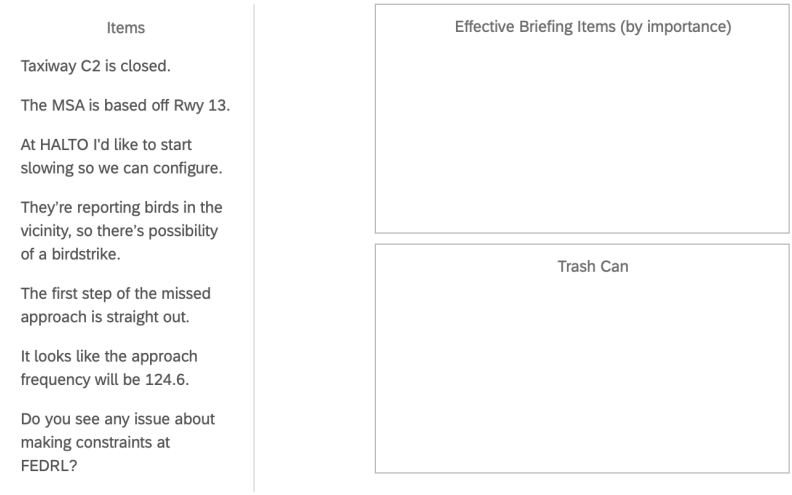
After investigating all this information the first time, participants created a briefing list of the three most important and next two most important items for their fellow pilot to know. They were then told ATC had communicated a change for their flight. Participants were allowed to review updated information and then were asked to list impacts the change would have on their flight.

Review Tasks

Like the generative tasks, the review tasks involved two parts: a briefing and then a reactive action plan. The same scenario was used for these review tasks as was used for the generative task within each version of the assessment. Participants then sorted briefing or action plan items into “effective items” or a trash can and ordered the effective items by importance (see Figure 3).

Figure 3

Review Task Example

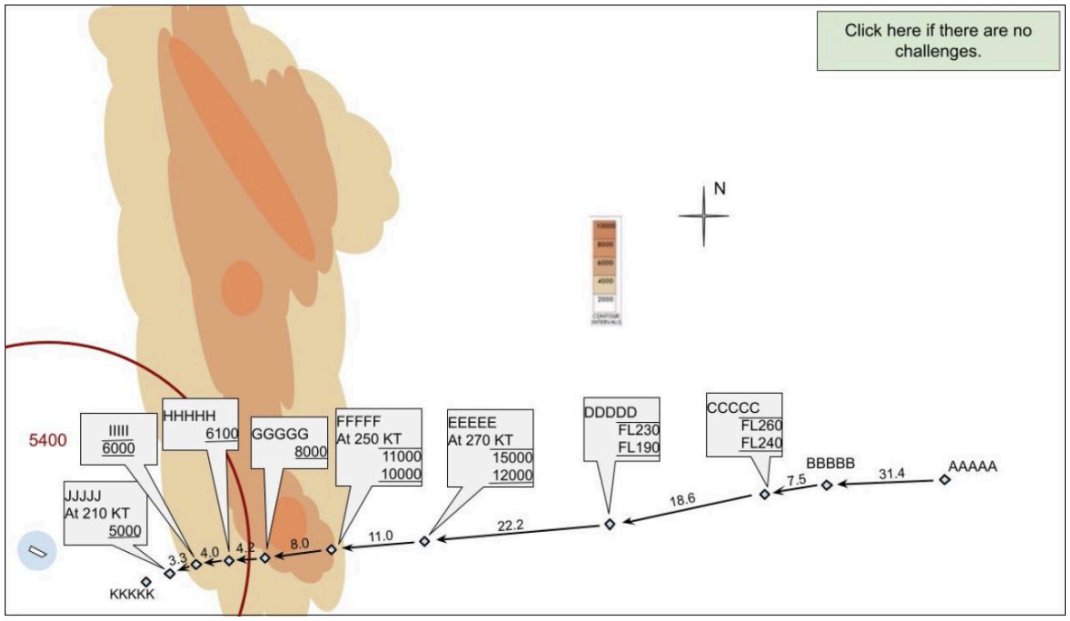


Arrival Chart Analysis Tasks

The arrival chart analysis tasks included three simplified charts, one with steep sections and a shortcut, one with steep sections and terrain, and one with no challenge, to use as a control for guessing. In Figure 4, the high terrain can be seen in orange, and there is a steep section from GGGGG to JJJJJ. It should be noted participants were also given the option to say there were no challenges present in the chart.

Figure 4

Example Arrival Chart with Challenges



Discussion

At the time of this writing, data analysis is ongoing. However, the operationalization of these methods for use within pilot training is a key objective of this study. Further, there are potential applications to other complex and distributed fields.

Application and Practice

In practice, pilots are a distributed audience who require continuous learning and assessment of their skills to maintain their certificates. While much on-the-job training has been moved to asynchronous methods, the assessment of the complex skills necessary to perform well as a pilot are often limited to in-person simulator sessions. Crucially, none of the methods described here require in-person time on the part of the pilot. This allowed the pilots to engage with the assessments from almost anywhere in the world, and on their own time. Most importantly, rather than being just a knowledge assessment, the methods described here required complex contextual analysis of scenarios and decision-making capabilities. They also provided a low-stakes chance to practice these skills.

This form of training and evaluation could be of high value for use by airlines. The approach could provide pilots with engaging learning and valuable assessment during routine travel, and on their own time. Training-assessment activities could be designed in small units and pilots might be given choice of unit to do within a certification period. This could provide a more integrated learning experience for pilots and allow them to practice skills before certification assessments. These web-based activities could align with and reinforce the related simulator scenarios and assessments.

Domains outside of aviation likely share the need for assessing cognitive skills used in complex and high-stakes systems. This work may help inform similar assessments in other contexts, such as multi-agency emergency response or safety logistics for large public events. In practice, matching the user's environment with deeply contextual scenario-based tasks, even asynchronously, may work when synchronous assessment is not feasible, due to widely distributed collaborators, or a lack of resources for in-person simulations.

Future Work

The design of these assessment items was intended to allow researchers to assess complex cognition related to flight safety and anticipation of challenges to safe flight. This work is part of a larger study, exploring the efficacy of the learning intervention and evaluating baseline knowledge and skills related to anticipatory behaviors. Additionally, these methods lead to further research questions:

1. How does performance on these methods correlate with simulator performance?
2. How does this method potentially impact pilot learning and assessment?

A simulator study is in the planning stages to begin to answer the first of these questions. Further, an analysis of the data collected so far is in progress, which will be discussed in future papers.

References

Arbuckle, P. D., Abbott, K. H., Abbott, T. S., & Schutte, P. C. (1998). Future flight decks. In *Proceedings from the 21st Congress of the International Council of the Aeronautical Sciences*, Melbourne, Australia.

Baron, B., Billman, D., & Mumaw, R. J. (2023). Exploring informal learning structures in airline operations. *22nd International Symposium on Aviation Psychology*, 82. https://corescholar.libraries.wright.edu/isap_2023/57

Holbrook, J. B., Stewart, M. J., Smith, B. E., Prinzel, L. J., Matthews, B. L., Avrekh, I., Cardoza, C. T., Ammann, O. C., Adduru, V., & Null, C. H. (2019). *Human performance contributions to safety in commercial aviation*. National Aeronautics and Space Administration.

Hollnagel, E. (2011). RAG – The resilience analysis grid. In E. Hollnagel, J. Pariès, D. D. Woods, & J. Wreathall (Eds.), *Resilience Engineering in Practice. A Guidebook*. Farnham, UK: Ashgate.

International Civil Aviation Organization (ICAO). (2024). *Safety Report*. Montreal, Canada.

Neville, K., Chen, D.-W., Massey, L., Cowell, T., Burbelo, G. A., Blankenbeckler, P. N., Normand, S. L., & Uhl, E. R. (2020). *A cognitive skills research framework for complex operational environments (Technical Report 1379)*. United States Army Research Institute for the Behavioral and Social Sciences.

Rankin, W. (2008). *MEDA: Investigation process*. The Boeing Company.

Rogers, C. J., Tolmie, A., Massonnié, J., & Thomas, M. S. C. (2023). Complex cognition and individual variability: a mixed methods study of the relationship between creativity and executive control. *Frontiers in Psychology, 14*, 1191893. https://doi.org/10.3389/fpsyg.2023.1191893

US Government Accountability Office (GAO). (2023). *Aviation workforce: Current and future availability of airline pilots and aircraft mechanics*. United States Government Accountability Office. Retrieved from https://www.gao.gov/products/gao-23-105571

Acknowledgments

Originating group and funding provided by NASA System Wide Safety (SWS), Human Contributions to Safety (HC2S)