INITIAL MEASURES SHOW WEB-DELIVERED LEARNING MODULE IMPROVES PILOTS' MONITORING & ANTICIPATION

Dorrit Billman
NASA Ames Research Center
Moffett Field, CA

Barth Baron, Jr.
University of Hawai‘i at Mānoa,
San Jose State University, CA

P. Christine Corry
San Jose State University
San Jose, CA

Lucas Cusano
NASA Ames Research Center
Moffett Field, CA

Melissa Peterson
University of Hawai‘i at Mānoa,
San Jose State University, CA

Randall J. Mumaw
San Jose State University
San Jose, CA

Anticipation and monitoring are key pilot activities that build safety margin in flight operations. Conversely, inadequate monitoring is frequently identified as a factor contributing to aviation accidents. We propose that: effective, anticipatory monitoring is a proactive cognitive activity; development of tutorials for web-based learning may be an effective method for learning anticipation and monitoring skills; and anticipatory monitoring skills can be measured on the web. Our exploratory study with airline pilots evaluated whether pilots' anticipatory monitoring improved after a tutorial as measured in a web-based assessment. We found large, significant gains on a multiple- choice test that assessed the understanding and application of concepts and strategies taught in the tutorial. We identify challenges and potential value of the research approach and findings.

Turkish Airline's 737-800 flight 1951 crashed attempting to land at Amsterdam-Schiphol airport, 2009. Airline accidents are extremely rare and usually result from alignment of multiple problems or failures. This case was no exception. Here, these included a failed radio altimeter, an initially high airspeed; an unexpected, automated mode-transition providing less protection than expected; hasty execution of the landing check list items; confusing alerts; and focusing on the check list rather than landing the plane. The crew did not communicate about modes nor the descent management plan, nor prioritize their tasks well. Further they were jointly fixated on the checklist and on the confusing messages, rather than flying the plane. In sum, they did not understand the situation, did not coordinate tasks nor communicate information, nor did they prioritize flying the plane. Their anticipation and monitoring was inadequate to notice how the actual and expected states did not match, or to recognize the needed actions.

Inadequate monitoring and anticipation have been noted as a contributing factor in a high proportion of aviation accidents, and improving pilots' monitoring and anticipation is identified as an important path to improving aviation safety (CAST, 2014; FAA, 2013; ICAO, 2016). Further, monitoring and anticipation likely grow in importance with increasing complexity of operations and automation. A major accident often triggers efforts to prevent its recurrence. When efforts to improve safety is retrospective, resulting changes may be narrowly focused on the specific conditions of the accident, and the resulting changes of narrow scope. Ideally,

changes should improve safety margin and prevent accidents across a wide range of conditions. In part because major accidents are extremely infrequent, a reactive approach gives limited guidance for pervasive, effective change. The focus on "what goes wrong" has been labeled Safety I, to contrast to a Safety II focus on "what goes right" (Hoermann &Soll, 2004). This widened focus means looking at effective performance, not just undesired outcomes; at routine as well as exceptional performance; at all data, not just that from problems; and viewing safety as resilience-producing activity as well as absence of accidents.

Motivation and Overview of Our Research

Little research has investigated characterization, training, or assessment of monitoring and anticipation, particularly by pilots flying highly automated aircraft. (Related work has focused on situation awareness in the context of nontechnical skills, e.g., Hoermann & Soll, 2004). While there is much knowledge about acquisition of expertise and about assessment methods, these have not been systematically applied to the domain of flight path monitoring and anticipation.

Our study asked whether a web tutorial on anticipatory monitoring produced measurable improvements in pilot performance. Results from one assessment task are reported here. Our research draws on our characterization of effective anticipatory monitoring and its skills & knowledge (S&K); on our tutorial designed and implemented for remote pilots; and on diverse measures developed for remote assessment. We narrow our scope by focusing on monitoring and anticipation for flight path management, particularly energy management during descent. Our work was further narrowed by striving for high feasibility and easy uptake of any methods developed, specifically an interest in web methods for learning and measurement.

Monitoring and anticipation are tightly linked, as monitoring is centrally an active process of comparing current state with expectation. Effective monitoring and anticipation generate a cycle of posing a question or noticing something puzzling, comparing the actual to expected situation, and assessing the implications for action. This integrates expectations based on a mental model of 'how things work' with information about the current situation to form a situation model (Billman, Mumaw, & Feary, 2020). When pilots communicate information about a situation model, this enables a shared situation model. Effective monitoring also depends on context-specific strategies about what questions to ask, information to seek, and assessments to carry out (Mumaw, Billman, & Feary, 2021).

An effective tutorial provides relevant content and effective delivery. Our tutorial provides a conceptual view of anticipatory monitoring as questioning and testing expectations. It demonstrates a set of strategies for assessing flight path compliance and possible energy management challenges. It identifies specific communications between Pilot Flying (PF) and Pilot Monitoring (PM) likely to be valuable in a Top of Descent briefing. Our tutorial delivered content over the web to remote users. Laptop-delivery with neither personal interaction nor dynamic control of events places limits on both learning and assessment activities. We addressed this challenge because there is a widespread operator preference for web training and matching that preference could facilitate uptake. The interactive tutorial used natural voice narration by an engaging pilot; demonstrated simple anticipatory strategies of what and how to monitor;

prompted the pilot to think of situations where strategies might have helped them; and provided interactive questions. Billman et al (2023) provides additional information about the tutorial.

The purpose of our assessment tasks was to measure the anticipation and monitoring S&K taught in the tutorial and whether performance changed as a result. Many items gave pilots a task based on a simple or extensive scenario. Item design balanced operational relevance with feasible web implementation. Preparation of a Top of Descent briefing from rich information sources about the situation was most directly relevant to operational tasks and also demanding to implement and score. The multiple-choice task was most efficient to design, answer, and score Questions included choosing an action in a situation or the application of a concept introduced in the tutorial. Peterson et al (2025) describes assessment challenges and design.

Exploratory Study

Method

Design & Participants. The within- and between-subject, mixed measure design compared pre- vs post-training performance. Two versions of the assessment were counterbalanced for order (Version A → Version B or B → A). The study was approved by the University of Hawaii IRB. 38 pilots from US airlines volunteered their time, 20 in one order and 18 in the other completing the study, half were Captains. Average total flight hours were 8504.

Procedure. Three activities— a pre-training assessment, the tutorial, and a post-training assessment —were presented via the web to remote learners. A 2-day retention interval separated the tutorial from the post-test. Each pilot accessed all materials, through a web link, on their own device (laptop recommended). Activities were self-paced, and the full study required about 3 to 3.5 hours of the participant's time.

Table 1. Sequence of Events. Assessment Tasks

	time -->				
		Pre-Train Assess	Training.	2-Day Retention Interval	Post-Train
Assessment Tasks	*Demographics*	✓			
	Briefing	✓			✓
	Action Choice	✓			✓
	STAR Assessment	✓			✓
	MULTIPLE-CHOICE	✓			✓
	Monitoring attitudes	✓			✓
	Tutorial & Assessment Feedback				✓

Materials. Broadly, the tutorial goal was to increase pilots' skills, knowledge, and attitudes supporting anticipatory monitoring for flight path management on descent. The tutorial presented both general principles underlying proactive, anticipatory monitoring and situation-specific strategies. Strategies focused on technical methods for assessing and monitoring flight path on descent (particularly on Standard Terminal Arrivals, STARs). The tutorial presented methods for identifying monitoring challenges, particularly across multiple segments. It provided specific communication methods between PF and PM. The design was guided by established learning principles (Mayer, 2022; Merrill 2013). We used short interactive activities based on a task-centered design and having the participants apply the presented concepts, reflect on the concepts presented, and connect the information to experiences they may have had. We used natural

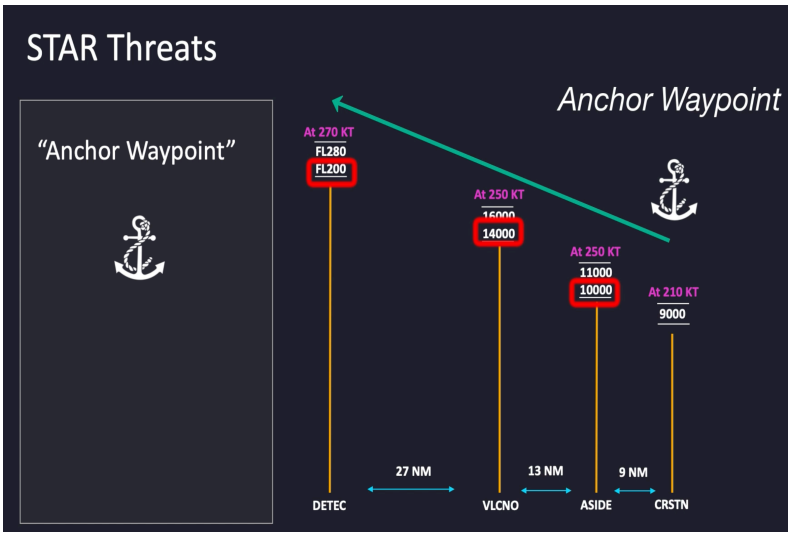

Figure 1. *Tutorial Screen: Anticipating Energy Management*

instead of synthetic voice, conversational instead of formal speaking tone, and voice and images instead of voice and text. Figure 1 shows a screen shot.

The assessment tasks are listed in order in Table 1. They ranged from analogs of operational tasks (e.g., formulating a descent briefing) to familiar test formats of multiple choice. Multiple choice items assessed retention of concepts and terms (matching a concept to an example or description) and skills applying the strategies and concepts to brief scenarios. The 32 four-alternative forced-choice items were divided into Versions A and B. Fifteen pairs (30 items) assessed S&K presented in the tutorial. (One automation pair was an exploratory control item.)

Multiple Choice Results

Overall, scores on the tutorial-relevant items increased 26.1%, from 53.9% to 80.0% correct from before to after the tutorial (χ^2= 100.8, p <2e-16; R mixed model regression with pre-post as fixed effect and participant and item as random effects). Figure 2 shows consistent improvement, for 37 pilots (one pilot unchanged), and were likely limited by ceiling effects.

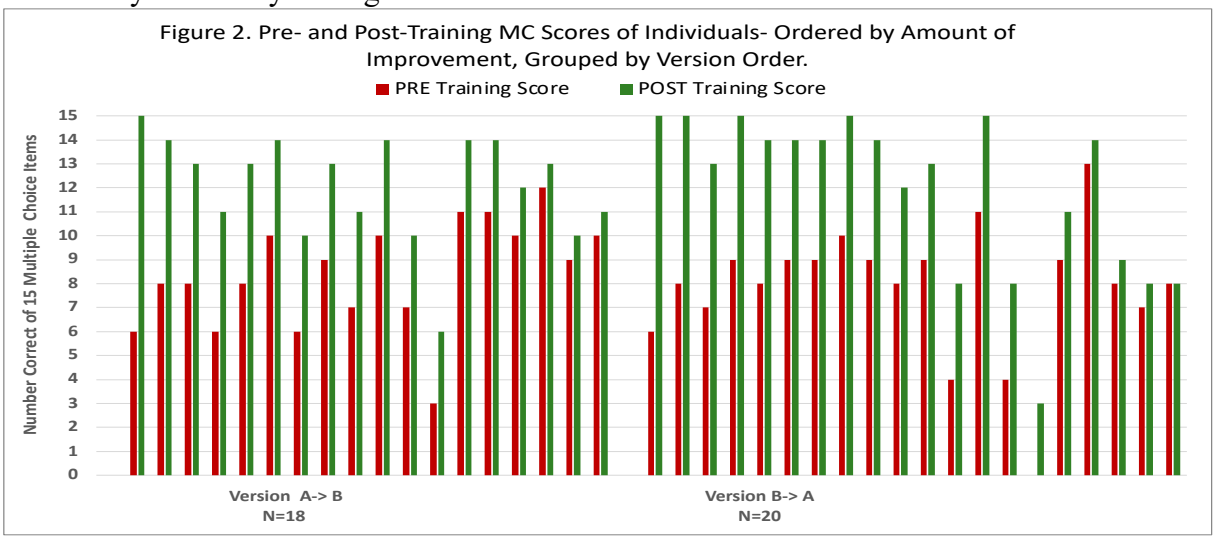

Figure 2. *MC Scores by Individual.*

The multiple-choice items were grouped into three types. Because the items could integrate multiple concepts and strategies, these three types conceptually overlap, though each item was assigned to just one type. The Core Concepts type (14 items) assessed concepts central to anticipatory monitoring that were presented in the tutorial

and unlikely to be familiar, for example, monitoring cues, plans, and objectives; and new strategies for monitoring STARs. The Energy Management type (5 items) assessed use of technical flight path management strategies that likely had familiar components. The Communication type (11 items) included any focus on communication; we expected pilots would have had exposure to related ideas; however, four items addressed elevating the role of PM, which we expected to be previously unfamiliar. Core Concept questions showed most improvement with 33.7% increase in score. Energy Management improved 21.2% and Communication items improved 17.1%. Also, the subset of four Communication items that addressed ideas likely to be unfamiliar about the PM role averaged 31.5% improvement.

We began analysis with the multiple-choice task because it was easy to score and analyze. In parallel, we developed a coding system for the free-text Generate-Briefing task, as the best analog of an operational activity. Descriptive analysis shows improved scores, such as more frequent inclusion in briefing items of an emerging threat at the airport.

Conclusions and Implications

Our results suggest that a sample of anticipatory monitoring skills can be learned through a carefully designed tutorial. Multiple-choice performance after the tutorial showed application of S&K to new content, retention over a few days, and large performance improvement. Our intent was for pilots to learn both effective context-specific strategies and a more generalizable foundation for transfer to novel situations sharing recognized structural elements.

Multiple important methodological challenges remain. Within a web-delivered format, it would be valuable to assess whether and how post-test performance improves even without the tutorial. The assessment activities provide no feedback but are quite rich and might provide an additional, useful learning opportunity. It is valuable to separate any repetition effect from that of the tutorial. We hope to conduct a web study including a control condition with two assessments but no tutorial. The most critical methodological issue, we believe, is whether results from the web intervention transfer to performance on the simulator. Web delivery provides little ability to control or measure times. We are seeking an opportunity for a simulator study.

The research presented here focuses on learning at the individual level. Pilots learned strategies for monitoring and anticipation that required flexible thinking about the case at hand. The goal is learning both effective context-specific strategies, and a foundation for transfer to novel situations sharing recognized structural elements. We suspect that a more proactive approach to monitoring may also help pilots stay engaged, aid in adaptive task prioritization, and add resilience. These suggestions clearly require considerable additional investigation.

The value of proactive monitoring increases as complexity increases, particularly, complexity of automation both in airliners and in the many new types of aircraft entering the airspace. Across these operations, people serve as traditional in-air pilots, remote pilots, and real-time managers and controllers of fleets and airspace. Across these roles, humans are adaptive. They provide skill to understand a dynamic situation, ability to adapt to unexpected combinations of conditions, and capability to learn. Nevertheless, pilots' anticipatory monitoring may not be fully appreciated or supported at an organizational level (aircraft operators, whether

airlines or UAS developers). Opportunities for gathering pilot experience, both to accelerate learning of other pilots and to inform other aspects of operations, may not be well developed. Getting the benefits of increased adaptability and resilience from proactive monitoring may depend on linkage between learning at the individual and organizational level. The work presented here aims to be a small step supporting learning at the individual level, around principles that can link to organizational levels.

Author Notes

This work was supported by NASA System-Wide Safety, HC2S 340428.02.70.01.01

References

Billman, D., Baron, B., & Mumaw, R.J. (2023) Cognitive Engineering in Training: Monitoring and Pilot-Automation Coordination in Complex Environments. In Proceedings of Applied Human Factors and Ergonomics Conference, San Francisco, CA, July 20-24, 2023.

Billman, D, Mumaw, R., & Feary, M. (2020) A Model of Monitoring as Sensemaking: Application to Flight Path Management and Pilot Training. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 64). SAGE Publications.

CAST (2014). Airplane state awareness joint safety analysis team: Final report analysis and results. Located at https://skybrary.aero/sites/default/files/bookshelf/2999.pdf

FAA (2013). Operational use of flight path managements systems; Final report of the performance based operations Aviation Rulemaking Committee / Commercial Aviation Safety Team Flight Deck Automation Working Group. Washington, DC: FAA.

ICAO International Civil Aviation Organization. (2016). Guidance material for improving flightcrew monitoring. Montreal, Quebec, Canada: ICAO.

Hoermann, H-J., & Soll, H. (2004). Training of situation awareness and threat management techniques. In *Practice and Research*. *Aviation Psychology.*

Hollnagel, E. (Ed.). (2013). *Resilience engineering in practice: A guidebook*. Ashgate Publishing, Ltd.

Mayer, R. E. (2022). *Multimedia learning*. Cambridge University Press.

Merrill, M. D. (2013). *First Principles of Instruction: Identifying and Designing Effective, Efficient, and Engaging Instruction*. San Francisco, CA: Pfeiffer.

Mumaw, R. J., Billman, D., & Feary, M. S. (2021). Analysis of pilot monitoring skills and a review of training effectiveness. NASA/TM-20210000047.

Peterson, M., Baron, B.,Cusano, L., Mumaw, R.W., Corry, P.C., Hoffman, D., &Billman, D. (2024). Scenario-Based Task Design for Airline Pilot Anticipatory Behaviors: Asynchronous Assessment of Complex Cognitive Skills. Proceeding of AECT, Kansas City, October 2024.