Modeling aviation crew interaction using a cognitive architecture

Robert W. Holt (bholt@gmu.edu)
Jeffrey T. Hansberger (jhansber@gmu.edu)
Ronald S. Chong (rchong@gmu.edu)
Deborah A. Boehm-Davis (dbdavis@gmu.edu)
Department of Psychology, George Mason University
4400 University Dr., Fairfax, VA 22030 USA

Holt (2001) proposed developing Scientific Information Systems to construct and validate theory concerning complex multi-person systems. Holt described a process of successive cycles of theory refinement using information in databases. Holt, Boehm-Davis, and Beaubien (2001) discussed the development of theory for describing crew performance in the aviation domain by statistically analyzing performance measures. These inductive, theory-building approaches require good data and analyses. Unfortunately, obtaining good quality measures may be difficult in domains such as aviation which are complex, dynamic, and multi-person (Holt, Johnson, & Goldsmith, 1997; Holt, Hansberger, & Boehm-Davis, in press).

An alternative approach is to carefully extend theory from a field closely related to the focus of research and subsequently validate it. This study was focused on aviation crew performance using flight deck automation during the descent phase of flight. The theory that was extended to this domain was the ACT-R 4.0 cognitive architecture (Anderson & Lebiere, 1998). The ACT-R architecture was extended to describe the highly procedural nature of crew performance in this context (e.g. checklists, Standard Operating Procedure, etc.). The initial development of this model focused on an ACT-R model of the Pilot Flying (PF) who had to receive directives from Air Traffic Control (ATC), decide on how to use the automation to achieve flight goals, and monitor the success or failure of actions.

Based on lessons learned from this initial effort, the approach was extended to constructing a crew model with a simulated PF and Pilot Not Flying (PNF). These crew members were simulated by separate ACT-R models based on a cognitive task analysis of the duties for each person. The simulated task scenario was the time period just before and after Top of Descent (TOD) in the descent phase of flight. The PNF tasks included verification and programming of the Flight Management System (FMS) computer as well as gathering appropriate information for completion of the flight. The PF monitors and flies the aircraft except for required briefings and responses.

Required aspects of crew interaction such as crew communication (e.g. briefings, acknowledgments) were implemented by a communication link between the PF and PNF simulations using a multi-model extension of ACT-R. Simulated communications involved goals, specific actions, or situational facts and features.

The linked PF and PNF models were evaluated by manipulating the simulated expertise of the crew. Expertise was simulated by changing ACT-R parameters and structures. Specifically, higher expertise was simulated by combinations of high strength of associative links for procedural behavior, higher working memory capacity, and cognitive strategies such as the systematic reactivation of goals cued by external stimuli such as a checklist.

One advantage of using the cognitive architecture was that a complete profile of cognition and performance could be measured for each simulation run. Model performance measures include total time for all tasks, average time for each task, checklist steps skipped, repeated, or performed out-of-order, automation programming delayed, skipped, or incorrect, and the omission of required communications.

Qualitative results such as step skipping, repetition, and intrusion of incorrect steps were observed at lower levels of simulated expertise. Emergent results included crew miscommunication, differential situation awareness, and forgetting relevant goals under certain conditions of delays and interruptions. The precise profile of performance differences for different levels of crew expertise can be used to develop assessment items, strategies, and guidelines for assessing performance of commercial crews.

Acknowledgments

This research was supported by grant NAG 2-1289 from the NASA, and 99-G-010 from the FAA.

References

Anderson, J.R. & Lebiere, C. (1998). *The Atomic components of thought*. Mahwah, NJ: Erlbaum.

Holt, R. W. (2001). *Scientific information systems.* Aldershot, UK: Ashgate Publishing.

Holt, R. W., Boehm-Davis, D. A., & Beaubien, J. M. (2001). Evaluating resource management training. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations: Applications of resource management training.* Mahwah, NJ: Lawrence Erlbaum Associates.

Holt, R. W., Hansberger, J. T., & Boehm-Davis, D. A. (in press). Improving rater calibration and performance in aviation. *International Journal of Aviation Psychology*.

Holt, R. W., Johnson, P. J., & Goldsmith, T. E. (1997). Application of psychometrics to the calibration of air carrier evaluators. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting,* 916-920.