

Wright State University

**CORE Scholar**

---

International Symposium on Aviation  
Psychology - 2011

International Symposium on Aviation  
Psychology

---

2011

## Situation Awareness, Workload, and Performance in Midterm Nextgen: Effect of Dynamic Variations in Aircraft Equipage Levels

Thomas Z. Strybel

Kim-Phuong L. Vu

L. Paige Bacon

Sabrina Silk Billinghamurst

Robert Conrad Rorie

*See next page for additional authors*

Follow this and additional works at: [https://corescholar.libraries.wright.edu/isap\\_2011](https://corescholar.libraries.wright.edu/isap_2011)



Part of the [Other Psychiatry and Psychology Commons](#)

---

### Repository Citation

Strybel, T. Z., Vu, K. L., Bacon, L., Billinghamurst, S. S., Rorie, R. C., Kraut, J. M., Morgan, C., Battiste, V., & Johnson, W. (2011). Situation Awareness, Workload, and Performance in Midterm Nextgen: Effect of Dynamic Variations in Aircraft Equipage Levels. *16th International Symposium on Aviation Psychology*, 118-123.

[https://corescholar.libraries.wright.edu/isap\\_2011/96](https://corescholar.libraries.wright.edu/isap_2011/96)

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2011 by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

---

**Authors**

Thomas Z. Strybel, Kim-Phuong L. Vu, L. Paige Bacon, Sabrina Silk Billinghamurst, Robert Conrad Rorie, Joshua M. Kraut, Corey Morgan, Vernol Battiste, and Walter Johnson

# SITUATION AWARENESS, WORKLOAD, AND PERFORMANCE IN MIDTERM NEXTGEN: EFFECT OF DYNAMIC VARIATIONS IN AIRCRAFT EQUIPAGE LEVELS

Thomas Z. Strybel, Kim-Phuong L. Vu, L. Paige Bacon, Sabrina Silk Billingham, Robert Conrad Rorie,  
Joshua M. Kraut, Corey Morgan  
Center for Human Factors in Advanced Aeronautics Technologies  
California State University, Long Beach  
Long Beach, CA

Vernol Battiste  
San Jose State University Foundation; NASA Ames Research Center  
San Jose, CA; Moffett Field, CA

Walter Johnson  
NASA Ames Research Center  
Moffett Field, CA

NextGen changes in air traffic management promise to bring many benefits to the current airspace system, but these changes must be evaluated for their impact on mid-term air traffic management in which mixed-equipage is certain. We examined mixed equipage environments in which the equipage levels changed over the course of the scenario to reflect changes in sector characteristic over the course of a day or controller's work shift. Six retired ATCs managed mixed-equipage traffic that either began with low levels of NextGen equipped aircraft and increased midway through the scenario or vice-versa. These were compared to a scenario in which the equipage mix was held constant. ATC performance, workload and situation awareness were affected differently by these scenarios.

The Next Generation Air Transportation System (NextGen), is a program being developed under the guidance of several government agencies, including the Federal Aviation Administration (FAA), Department of Transportation, and NASA, under the umbrella of the Joint Planning and Development Office (JPDO, 2007). The goals of NextGen include expanding the capacity of the National Airspace System (NAS) to handle 2-3 times current day traffic to accommodate projected increase in air travel over the next 10-20 years (see e.g., FAA Implementation Plan, March 2010). NextGen tools, technologies, and concepts will be implemented in phases (JPDO, 2007). Phase 1 (2007-2011) included research and development of avionics technologies for enabling NextGen concepts and procedures being planned for subsequent phases. One of these technologies, Data Comm, allows pilot and controllers to communicate using a text-based messaging system in lieu of voice communications.

The use of Data Comm is anticipated to reduce operator workload under high traffic environments by reducing the serial mode of voice transmissions (Kerns, 1991), reducing ambiguity of the intended message (Flathers, 1987), and reduce working memory demands associated with remembering auditory messages (Hinton & Lohr, 1988). In mid-term NextGen, Data Comm in the cockpit can be integrated with the flight management system. With integrated Data Comm, ATC and pilots can negotiate flight plan requests, and agreed-upon changes loaded into both ground and aircraft systems. ATCs and pilots may also benefit from automated conflict resolution which provides recommended speed or trajectory changes to resolve a conflict. Controllers managing Data Comm equipped aircraft (AC) could benefit from other NextGen technologies such as trial planners and conflict probes.

Although the potential benefits of Data Comm, along with changes to the NAS brought about by NextGen procedures would eventually lead to most aircraft being equipped with Data Comm, the adoption of Data Comm by commercial and private carriers is likely to develop gradually over time due to the cost associated with equipping aircraft. Consequently, air traffic controllers in the near- and mid-term will be challenged to manage AC having very different capabilities. In addition, controllers may not always benefit from NextGen tools for conflict detection and resolution because it requires that both ACs in conflict be appropriately equipped with NextGen technologies. Controllers will need to rely on their "manual" air traffic management skills for detecting and resolving conflicts between AC pairs involving one or more unequipped aircraft. A mixed-equipage fleet may also increase workload because ATCs must maintain awareness of the equipage level of all ACs in the sector.

Preliminary investigations of air traffic management in mixed equipage airspace have shown that the percentage of equipped aircraft affects ATC performance and workload. Corker et al. (1999) examined the effects

of mixed equipage environments on conflict detection and ATC workload. For equipped AC in this investigation, pilots had some responsibility for maintaining separation from other traffic. The authors showed that the equipage levels of 80 and 100% reduced workload, but that a 20% equipage level increased workload over the level achieved with 100% equipage. Prevot et al. (2005) reported that controllers who participated in mixed equipage airspace simulations were somewhat negative about the impact of mixed equipage on situation awareness and safety. Specifically, controllers reported that it would be slightly more difficult to detect non-conforming aircraft and more difficult to cope with unplanned events in mixed-equipage environments. Willems et al. (2008) showed that for 70% equipage levels, ATCs could manage a 33% increase in traffic over current day levels, but could not handle a 66% increase in traffic. Willems et al. also observed that in a 70% equipage environment, controllers attempted to uplink Data Comm messages to aircraft were not equipped to receive it. Hah et al. (2010) compared sector equipage levels of 0%, 10%, 50%, and 100% in a simulation of an en route sector. They noted that significant contributions of Data Comm required at least 50% equipped aircraft in the sector. Moreover, ATCs reported high percentages of equipped traffic changed the way they managed traffic.

In summary, NextGen changes in air traffic management promise to bring many benefits to the current airspace system, but these changes must be evaluated for their impact on mid-term air traffic management in which mixed-equipage is certain. In the present simulation, we investigated another aspect of mixed equipage traffic airspaces. We examined mixed equipage environments in which the equipage levels change over the course of the scenario, because the equipage mixture will most likely be dynamic or changing over the course of a day or controller's work shift. These may bring about changes in ATC workload, which has been shown to reduce operator situation awareness (SA; Hallbert, 1997). ATCs managed traffic in a simulated en route sector having an average of 50% equipped aircraft, but the mixture in some scenarios began with a lower level of equipped aircraft and increased midway through scenario, or vice versa. Workload and situation awareness were measured with an online probe technique that is a variant of the Situation Present Assessment Method (SPAM; Durso & Dattel, 2004) because SA information is not only limited to the contents of working memory, but also can be distributed across the operators' task environment (Chiappe, Strybel & Vu, submitted).

## Method

### Participants

Seven retired ATCs participated in the simulation. All were former radar-certified TRACON ATCs with 9-25 years of experience in either Southern California or Bay Area TRACON facilities. One participant had eight years experience in an en route center, and one had participated in previous simulations at NASA Ames Research Center. None reported having real-world experience with the sector being simulated in the present study.

### Apparatus and Scenarios

The simulation was run using the Multi Aircraft Control System (MACS), Aeronautical Datalink and Radar Simulator (ADRS) developed by the Airspace Operations Laboratory at NASA Ames Research Center (e.g., Prevot et al., 2006). Participants managed traffic in simulated sector ZID-91, using MACS configured as a Digital System Replacement (DSR) display having advanced tools (integrated Data Comm, conflict alerts and probes, and trial planners) that could be used to manage AC designated as equipped. For the remaining non-equipped AC, participants managed traffic with voice commands. All ATC instructions were executed by a pseudopilot located in an adjacent room. ATC participants determined the status of each AC's equipage by checking the data tag for a diamond located to the left of the call sign indicating that the aircraft was equipped. Online situation awareness and workload probes were presented on a touch-screen workstation located adjacent to the DSR display.

Six experimental trials were run, each lasting approximately 50 minutes, with each scenario containing on average 50% equipped and 50% unequipped ACs. Within some scenarios the percentage of equipped AC changed midway through the scenario. Two scenarios started with 25% equipped AC. Midway through the scenario the percentage increased to 75% and remained roughly constant thereafter. Two scenarios began with 75% equipped AC, and the percentage decreased to 25% at the halfway mark. Finally, two scenarios had a constant mix of 50% equipped AC throughout the scenario. For each equipage mixture, one scenario contained 8 conflicts and the other six conflicts, as shown in Table 1, with half the conflicts occurring between equipped aircraft (and subsequently producing an alert on the DSR.) Of the remaining conflicts, the number of conflicts between one equipped and one unequipped AC, and between two unequipped ACs was counterbalanced. In addition, the number of conflicts in the first and second half of each scenario was equated. Note also in Table 1, that the second half of the scenario contained more AC, which reflects potential increases in traffic that occurs at particular times of day.

Table 1

*Traffic Characteristics (Percent Equipped AC, Number Alerted Conflicts) of the Scenarios*

Scenario	% Equipped by Half	Total Num. Conflicts	Num. Equipped Conflicts	Num. Un-equipped Conflicts	Ave. Num.AC: 1 <sup>st</sup> Half	Ave. Num.AC: 2 <sup>nd</sup> Half
1	25-75	6	3	3	6.6	9.1
2	50-50	6	3	3	7.7	10.6
3	75-25	6	3	3	7.6	14.0
4	25-75	8	4	4	7.6	10.9
5	50-50	8	4	4	7.5	12.1
6	75-25	8	4	4	7.0	9.0

### Procedure

Participants were given one week of training consisting of a briefing on traffic flows, ATC procedures, the MACS DSR interface, conflict alerting and use of the trial planner, and online probe questions. Following the briefing, participants practiced managing traffic using current day procedures, with NextGen tools, and with probe questions (see Kiken et al., 2011). The experimental sessions were run in the week after training over two days. Each participant ran one each of the scenarios described in Table 1, while simultaneously responding to probe questions. In each experimental scenario, 16 probe questions were presented, one every three minutes, beginning four minutes into the scenario. The questions were designed to assess ATC's awareness of conflicts between aircraft (e.g., "In what area of your sector will the next conflict occur?"), information relating to sector status (e.g., "Are there more equipped AC in your sector at this moment?") and workload (7-pt scale similar to the ATWIT). Workload prompts were administered four times at regular intervals. The sequence of events for each probe question was as follows. The ATC received a "Ready for Question?" prompt on the touch screen accompanied by an audio alert. The participant responded by touching a button on the screen when he/she had sufficient excess capacity to answer a question. Once the participant responded affirmatively, a probe question and response alternatives were displayed on the panel and the participant answered the question by pushing one of the response buttons. If the participant did not respond to the "Ready" prompt within two minutes the query was withdrawn.

### Results and Discussion

Measures of workload, performance, and situation measures were analyzed with repeated measures ANOVAs. For time-based variables, the analyses were done on log transforms but the results are shown in seconds. Two measures of workload, ATWIT ratings and time to respond to ready prompt, were analyzed with three-factor repeated-measures analyses of variance, with the factors of percent equipped AC, number of conflicts and scenario half. For ATWIT ratings, a significant three-way interaction was obtained ( $p < .05$ ), as shown in Figure 1. For each scenario, the ATWIT ratings increased in the second half, with the greatest increase occurring for the 75%-25% scenario with 6 conflicts, and the 50%-50% scenario with 8 conflicts. Moreover, in the second halves of both these scenarios, the mean rating was very high ( $M=4.9$  for both scenarios), and ATC provided ratings of 6 or higher roughly 50% of the time. From Table 1, these scenarios contained the highest traffic densities. The three-way interaction was marginally significant for ready latency ( $p < .06$ ). As shown in Figure 2, for scenarios containing 6 conflicts, ready latency increased in the second half when the scenario contained 25%-75% equipped AC, but decreased in the second half when the scenario contained 75%-25% equipped AC. For 8-conflict scenarios, the mean ready latency increased in the second half for scenarios containing 50%-50% and 75%-25% equipped AC. Obviously, these ready latencies are inconsistent with ATWIT ratings. We believe this is due to the fact that ATCs did not respond to Ready prompts frequently in the second half of the scenario. For example, in the 50%-50% scenario with 8 conflicts, the ready prompt was ignored (and timed out) on 35% of the workload probe queries. Nevertheless, a marginally significant correlation was obtained between ATWIT rating and Ready latency ( $r = .20$ ;  $p = .07$ ). Only ATWIT ratings were significantly correlated with average number AC being worked ( $r = .21$ ,  $p < .05$ ), suggesting ATWIT subjective workload is related to ATCs' perception of the number of AC being managed.

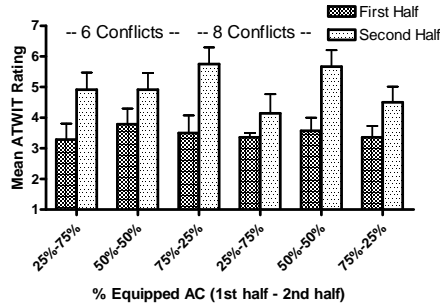


Figure 1. Mean ATWIT Rating as a function of equipage and number of conflicts.

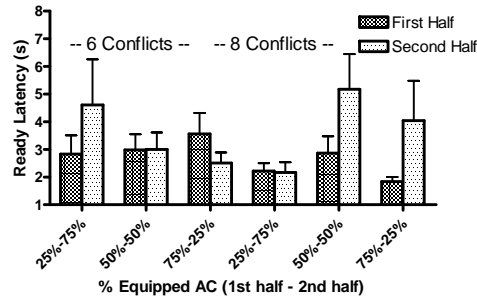


Figure 2. Mean Latency to Ready Prompt as a function of equipage and number of conflicts.

## Performance

Based on the extremely high workload in two scenarios, subsequent analyses was limited to those scenarios having roughly the same traffic densities, 25%-75% with 6 and 8 conflicts, 50%-50% with 8 conflicts, and 75%-25% with 6 conflicts. Performance was measured in terms of safety and efficiency. The number of alerted LOS and non-alerted LOS were analyzed for second halves only because no LOS occurred in the first half of any scenario.

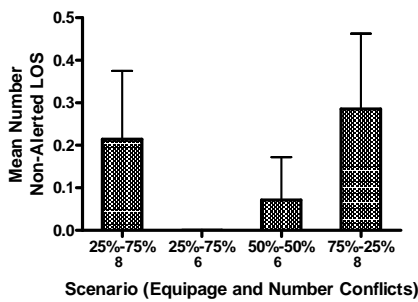


Figure 3. Average non-alerted LOS as a function of scenario for the second half.

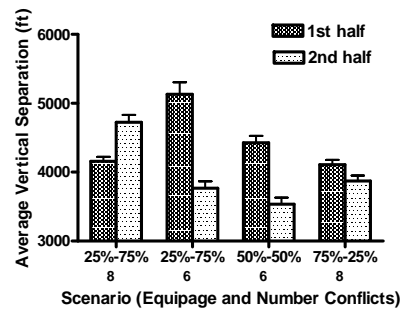


Figure 4. Vertical separation between AC as a function of scenario and scenario half.

For non-alerted LOS, a marginally significant effect of scenario was obtained ( $p < .07$ ). As shown in Figure 3, more LOS occurred in scenarios containing 8 conflicts, and equipage did not seem to affect the number of LOS. For vertical and average lateral separation between AC, repeated measures analyses of variance were performed with the factors being scenario and scenario half. A significant main effect of scenario and a significant interaction between scenario and scenario half ( $ps < .001$ ) were obtained, as shown in Figure 4. Vertical separation was greater on average for the 25%-75% scenarios compared with both 50%-50% and 75%-25% scenarios. For the 25%-75% scenario with 8 conflicts, vertical separation increased in the second half. In all other scenarios, average vertical separation decreased. A main effect of scenario and scenario half was obtained for average lateral separation, shown in Figure 5 ( $p < .001$ ). Lateral separation was lower for the 25%-75% equipage scenarios compared with the remaining scenarios. For each scenario, lateral separation increased in the second half, and the interaction with scenario was not significant. Therefore, when the scenario began with 25% equipped AC and this increased to 75% AC, vertical separation was greater and lateral separation smaller than when the scenario contained 50% equipage throughout or began with 75% equipage and this decreased to 25% equipage.

Sector efficiency was measured with average handoff delay and average time working an AC. For handoff delay, only scenario half was significant ( $p < .002$ ); handoff delay increased in the second half of the scenario, presumably because of the increase in traffic. For average time spent working AC, significant main effects of scenario and scenario half were obtained ( $ps < .002$ ). On average, ATCs spent the least amount of time working on AC in the 50%-50% scenario compared with the remaining scenarios. This may not be due to the number of conflicts, however, because the longest average time was for the 25%-75% scenario with 6 conflicts.

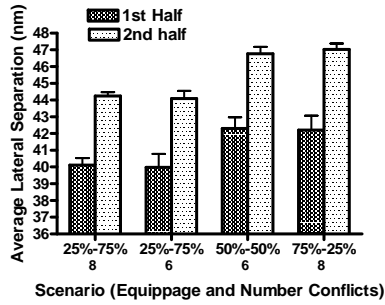


Figure 5. Average lateral separation (nm) as a function of scenario and scenario half.

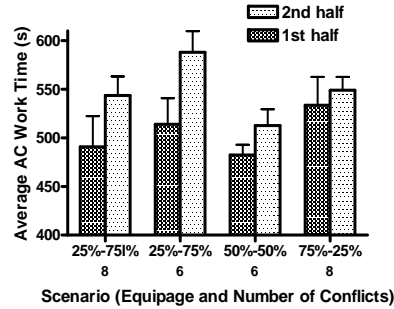


Figure 6. Average time spent working AC as a function of scenario and scenario half.

## Situation Awareness

Situation awareness was measured with the accuracy and latency of responses to online probe queries relating to conflicts or sector status. Probe latencies were analyzed separately for conflict and sector status queries. For sector status queries, a significant interaction between scenario and scenario half was obtained, as shown in Figure 7. For most scenarios, probe latencies were either unaffected by scenario half, or they increased minimally. Probe latencies for the 50%-50% scenarios were nearly identical, meaning that awareness of sector status did not change when the percent of equipped AC remained constant. On the other hand, for 75%-25% equipage, sector probe latency increased in the second half by 3 s on average, suggesting that awareness for this information was lowered by the decreasing percentage of equipped AC. The effect of scenario on conflict probe latency was marginally significant ( $p < .10$ ), but the effect of scenario half was nonsignificant.

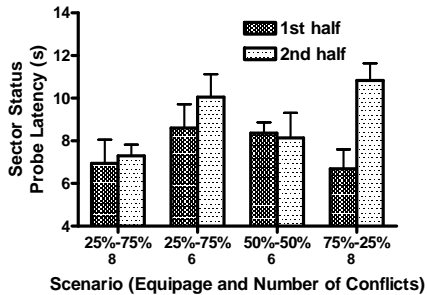


Figure 7. Average probe latency to sector status queries as a function of scenario and scenario half.

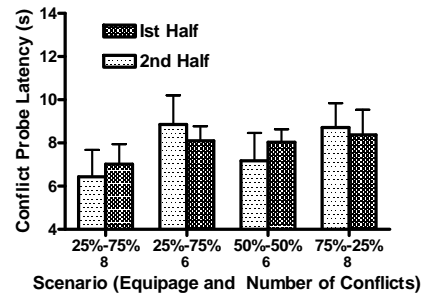


Figure 8. Average probe latency to questions on conflicts as a function of scenario and scenario half.

Conflict probe latencies, shown in Figure 8, were lowest at 25%-75% with 8 conflicts, and slightly higher for the remaining scenarios. To check the validity of our probe latencies, we computed correlations between sector probe latency, conflict probe latency, and performance metrics. Sector probe latencies were significantly correlated with handoff delay ( $r = .29$ ;  $p < .05$ ). Conflict probe latencies and number LOS was marginally correlated,  $r = -.26$ ;  $p < .07$ . The correlation between conflict probe latencies and average time spent working AC was significant  $r = -.35$ ;  $p < .01$ .

## Conclusion

Our preliminary investigation showed that changes in equipage within a scenario affected ATC performance, workload, and situation awareness, but the effects depended on the specific measure. By limiting our analysis to scenarios with roughly the same traffic densities, we determined that when the percentage of equipped AC was initially low and then increased during the scenario, vertical separation was greater and lateral separation smaller than when the percentage of equipped AC was constant or began high and then decreased during the scenario. Combined, these changes in vertical and lateral separation suggest that ATCs changed their strategies for managing traffic based on changes in equipage. The number of LOS was affected only by the number of conflicts,

and not equipage levels. LOS occurred only in the second half of each scenario. ATC perceived workload (ATWIT) increased in the second half of the scenario regardless of equipage levels, because of the increase in traffic during the second half. Situation awareness of sector status information was lower (probe latencies higher) in the second half, and this change was greatest when equipage began at 75% and decreased to 25%. Conflict probe latencies were not dependent on scenario half, because the number of conflicts in each scenario half was equal. In summary, these results indicate that changes in mixed equipage traffic within a scenario should be investigated further to determine the extent to which ATCs change traffic management strategies in response to equipage level.

## Acknowledgements

This simulation was partially supported by NASA cooperative agreement NNA06CN30A: *NRA: Metrics for Situation Awareness, Workload, and Performance in Separation Assurance Systems* (Walter Johnson, TM). Preparation of this paper was supported by NASA cooperative agreement NNX09AU66A, *Group 5 University Research Center: Center for Human Factors in Advanced Aeronautics Technologies* (Brenda Collins, TM).

## References

- Chiappe, D. L., Strybel, T. Z., Vu, K.-P. L. (2010). Review of Situation Awareness Theories from the Perspective of Cognitive Science. Submitted for publication.
- Corker, K., Fleming, K., & Lane, J. (1999). Measuring controller reactions to free flight in a complex transition sector. *Journal of Air Traffic Control*, 4, 9–16.
- Durso, F.T. & Dattel, A.R. (2004). SPAM: The real-time assessment of SA. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp.137-154). Aldersot, UK: Ashgate.
- Flathers, G.W. (1987). Development of an air ground data exchange concept: Flight deck perspective (NASA Contractor Rep. No. 4074). Hampton, VA: NASA Langley Research Center.
- Hallbert, Bruce P. (1997) Situation awareness and operator performance: results from simulator-based studies. *Proceedings of the IEEE sixth annual Human Factors Meeting*, Orlando, Florida.
- Hinton, D.A. & Lohr, G.W. (1988). Simulator investigation of digital data link ATC communications in a single-pilot operations (NASA Tech. Paper 2837). Hampton, VA: NASA Langley Research Center.
- Joint Planning and Development Office. (2007). *Concept of operations for the Next Generation Air Transportation System, Version 2.0*. Washington, DC: Author.
- Kerns, K. (1991). Data-link communications between controllers and pilots: A review and synthesis of the simulation literature. *The International Journal of Aviation Psychology*, 1 (3), 181-204.
- Kiken, A., Rorie, R.C., Bacon, L. P., Billingham, S., Kraut, J.M., Strybel, T. Z., Vu, K.-P. L., & Battiste, V. (in press). Effect of ATC training with NextGen tools and online situation awareness and workload probes on operator performance. *HCI 2011, Lecture Notes in Computer Science*.
- Prevot, T., Smith, N., Palmer, E., Mercer, J. Lee, P., Homola, J. & Callantine, T. (2006). The Airspace Operations Laboratory (AOL) at NASA Ames Research Center. *AIAA Modeling and Simulation Technologies Conference and Exhibit* (AIAA 2006-6112). Keystone, Colorado.
- Salmon, P., Stanton, N., Walker, G. & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics*, 37 (2), 225-238.
- Willems, B., & Hah, S. (2008). *Future En Route Workstation Study (FEWS II): Part I – Automation Integration Research* (DOT/FAA/TC-08/14,II). Atlantic City International Airport, NJ: Federal Aviation Administration William J. Hughes Technical Center.
- Willems, B., Hah, S., & Schulz, K. (2010). *En route data communications: Experimental human factors evaluation* (DOT/FAA/TC-10/06). Atlantic City International Airport, NJ: FAA William J. Hughes Technical Center.