NAVING LANGLEY RESEARCH CENTER

GRANT NUMBER: NAG-1-01002

LEVELS OF AUTOMATION IN THE COMMERCIAL AIRCRAFT COCKPIT AND PILOT SITUATION AWARENESS

FINAL REPORT FOR NORTH CAROLINA STATE UNIVERSITY

Prepared By:
David B. Kaber, Mica R. Endsley, Melanie C. Wright, and Heather L. Warren

Period of Performance: January 1, 2001 to December 31, 2001
### INTRODUCTORY MATERIAL

<table>
<thead>
<tr>
<th><strong>Legal Name of Organization:</strong></th>
<th>North Carolina State University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Address:</strong></td>
<td>Sponsored Programs Services</td>
</tr>
<tr>
<td></td>
<td>North Carolina State University</td>
</tr>
<tr>
<td></td>
<td>P.O. 7514</td>
</tr>
<tr>
<td></td>
<td>Raleigh, NC 27695-7514</td>
</tr>
<tr>
<td><strong>Originating Unit:</strong></td>
<td>Department of Industrial Engineering</td>
</tr>
<tr>
<td><strong>Organization Type:</strong></td>
<td>Educational</td>
</tr>
<tr>
<td><strong>Principal Investigator:</strong></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>David B. Kaber</td>
</tr>
<tr>
<td>Phone</td>
<td>(919) 515-3086</td>
</tr>
<tr>
<td><strong>Project Title:</strong></td>
<td>Levels of automation in the commercial aircraft cockpit and pilot situation awareness</td>
</tr>
<tr>
<td><strong>Technical Contact:</strong></td>
<td>Anna Trujillo/Paul C. Schutte, Crew/Vehicle Integration Branch, Flight Dynamics and Control Division</td>
</tr>
<tr>
<td><strong>Starting Date:</strong></td>
<td>1/1/01</td>
</tr>
<tr>
<td><strong>Closeout Date:</strong></td>
<td>12/31/01</td>
</tr>
<tr>
<td><strong>Elapsed Project Duration:</strong></td>
<td>12 mos.</td>
</tr>
</tbody>
</table>
The Effects of Levels of Automation on Performance, Situation Awareness, and Workload in an Advanced Commercial Aircraft Flight Simulation

DAVID B. KABER¹, MICA R. ENDSLEY², MELANIE C. WRIGHT¹ AND HEATHER L. WARREN¹

¹Department of Industrial Engineering, North Carolina State University, Raleigh, NC 27695-7906, USA
²SA Technologies Inc., 4731 East Forest Peak, Marietta, GA 30066, USA

The objectives of this study were to assess the impact of a broad range of modes, or levels, of automation on pilot performance, situation awareness (SA) and mental workload in the context of a high-fidelity simulation of an advanced commercial aircraft, and, in particular, to provide information on high-level automation for future cockpit interface design. Previous research studied the performance and SA effects of theoretical levels of automation (LOAs) in abstract simulations but results have not been linked to specific types of real-world automated systems. This may be critical for cockpit automation as it involves a wide variety of automation with various capabilities and different premises for design. For example, no research has investigated the effects of different LOAs afforded by the Flight Management System (FMS) in contemporary aircraft on pilot performance.

We applied a theoretical taxonomy of LOAs developed by Endsley and Kaber (1999) to the McDonnell Douglas (MD)-11 aircraft by categorizing modes of automation currently provided through the MD-11 autoflight system in terms of the taxonomy. We also conceptualized new autoflight modes for the MD-11 FMS to model higher LOAs as part of the taxonomy by assuming the availability of an expert system on the flight deck. Select LOAs were modeled in a virtual flight simulator (VFS) in order to empirically evaluate any affects on pilot performance, SA and workload. The flight simulator included all automated subsystems as part of the MD-11 “glass cockpit” and manual attitude and throttle controls. The flight scenario was a generic recreation of the American Airlines 965 incident near Cali, Colombia in 1995 with an approach revision occurring within three waypoints of the Cali airport.

An experiment was conducted with 16 private pilots, who flew the simulator under one of four LOA conditions including: Manual Control with or without automated subsystem displays; an Action Support condition involving pilot flight planning and use of the Flight Control Panel (FCP); and a Supervisory Control condition in which the FMS generated a flight plan and implemented the plan, however, the pilot could intervene at the Action Support LOA and assume control of aircraft flight planning at any time. (Supervisory Control represented a LOA not currently available on commercial flight decks.) Each subject was trained extensively and flew two test trials at their assigned LOA with slightly different approach revisions into the Cali airport. All flights commenced at altitude and speed within Colombian airspace and included three en route waypoints, two approach points and the runway. Performance and SA were observed at random freezes of the simulation during test trials by recording deviations from the flight plan and subject responses to SA queries. One freeze occurred in advance of the revision, another occurred directly subsequent to the revision, and a final freeze occurred at the close of the trial. Mental workload was captured using the NASA-Task Load Index (TLX) after a subject completed a trial.
Results indicated manual flight control to improve subsequent to the approach revision; however, this was not the case for automated control. When high-level automation (involving use of the FMS) preceded use of low-level automation (Action Support involving the FCP) subsequent to an approach revision, performance was never as good as low-level automation in advance of the revision. Situation awareness results across LOAs revealed pilots to possess better overall SA in advance of the approach revision. Directly subsequent to the revision, pilot perception of elements in the flight environment was significantly worse than prior to the revision and at the close of the trial. Workload results indicated Manual Control of the simulator without automated subsystem displays to pose the highest workload in comparison to all other conditions and high-level automation (Supervisory Control) in particular. This corresponded with improvements in manual flight performance and perception as part of pilot SA during the latter portions of simulated flights. Finally, results on pilot flight planning revealed worse performance by subjects assigned to high-level automation (Supervisory Control) in advance of the revision and worse planning after the revision across LOAs.

This work has demonstrated that various modes of commercial aircraft automation and futuristic cockpit automation may differ in terms of effects on pilot performance, SA and workload. In particular manual flight control involving traditional flightdeck displays appears to allow for superior performance after a critical flight event in comparison to low- and high-level automation, but this comes at the expense of high workload. Furthermore, automation, in general, may compromise pilot SA after a critical event as compared to manual flight control, which may allow for recovery of pilot perception. High-level automation not currently available in commercial aircraft appears to compromise pilot flight planning and cause reliance on pre-programmed FMS databases. These results may have direct applicability to flight decks, in comparison to those of previous studies involving low-fidelity simulations.
1. INTRODUCTION

This research was focused on the implications of commercial cockpit automation for pilot performance, situation awareness (SA) and workload. This report presents the results of a second project as part of a multi-year research program conducted by North Carolina State University and SA Technologies, Inc.

1.1. Cockpit Automation

Current implementations of automation in the advanced commercial cockpit have extended far beyond early definitions of cockpit automation, including Wiener’s (1988) statement that portions of flight tasks performed by human crews can be assigned, by choice of the crew, to machinery. Automation has been applied in commercial cockpits in a variety of ways that deviate from Wiener’s (1988) proposition, including:

1. complete replacement of the human in specific flight control functions;
2. provision of autonomous flight control;
3. design of automation to assume cognitive function responsibility, including decision making (i.e., an aspect of information processing to which humans are well suited); and
4. prevention of crew authority over which flight control functions will be automated.

Research has demonstrated that such approaches to automation can lead to serious pilot performance problems and new forms of flight control errors (cf., Sarter & Woods, 1995).

Some of the more common implications of a technology-centered, or “let the machine do it,” approach to automation on a human’s role in complex systems control include:

1. relegation to passive information processing (i.e., operators observe the state of system variables and decide whether to intervene in system control); and
2. substantial increases in monitoring performance requirements.

Consequences to pilot performance associated with such role changes include vigilance decrements (Parasuraman, Molloy and Singh, 1993) attributable to boredom or fatigue, increases in perceived workload, and decreases in SA leading to slower error detection and sluggish system recovery under failure modes (Endsley & Kaber, 1999; Endsley & Kiris, 1995). To expand on this, technology-centered automation has changed the commercial pilot’s role specifically in terms of the tasks that are allocated to an automated Flight Management System (FMS). In second- or third-generation glass cockpits, the majority of flight task functions are allocated to autoflight systems, including flight path control, general navigation, aircraft performance management, and flight progress monitoring. One of the key problems with this situation is that if pilots are essentially removed from the flight control loop, they may lose awareness of the mode in which the autoflight system is operating and become confused as to the concurrent functioning of cockpit interfaces (Sarter & Woods, 1995). This is very likely given the fact that the FMS in, for example, the McDonnell-Douglas (MD) 11 is capable of presenting approximately 17 different modes or levels of automation (LOAs), and within each of these general modes there are anywhere between 2 and 6 sub modes that all have unique
characteristics (Delta Air Lines, Inc., 1998). Although the modes are designed to facilitate flight performance under various circumstances, the majority of modes involve pilot indirect management of aircraft flight (i.e., they act through the automation to control the aircraft) and monitoring of flight control activities invoked by the FMS. A secondary role of pilot under this type of automation is to trouble-shoot apparent errors in flight control when the perceive it to be necessary.

In general, this scenario leads to pilot out-of-the- (control) loop unfamiliarity (Kessel & Wickens, 1982) or performance problems (Endsley & Kiris, 1995). A number of empirical studies of out-of-the-loop (OOTL) performance in laboratory settings have demonstrated that complex system operators have difficulty in forming higher levels of SA, including comprehension of system states in relation to task goals as well as projection of future states, when functioning under high-level automation (Carmody & Gluckman, 1993; Endsley & Kaber, 1999; Endsley & Kiris, 1995). Unfortunately the loss of SA has been identified as a leading factor in aviation accidents (Jones & Endsley, 1995).

As a result of pilots being removed from the flight control loop for extended periods of time, they may eventually loose familiarity with aircraft operations and this loss of SA can negatively influence decision-making capability (Endsley, 1995). In relation to this, Sarter (1996) has said that pilots are posed with the challenge of effectively monitoring multiple cockpit displays simultaneously and recognizing, based on their SA, aircraft behavior that may be normal under other circumstances, but is erroneous under current system settings.

1.2. Research Challenge

The main research question for this project concerned how automation can be configured, or designed, to make a pilot’s job easier by supporting them in addressing this challenge and, at the same time, capitalizing on the obvious advantages of automation, including efficiency and accuracy in flight path control, reduced pilot workload, etc. With this question in mind, Billings (1991; 1997) put forth a theory of human-centered automation and presented it in the context of aviation systems. Billings (1997) theory essentially says that automation should be made a team player. That is, automation should be designed such that human-automation interaction resembles human-human interaction (see Kaber and Endsley (in review) for an expanded definition of human-centered automation).

Several approaches to human-centered automation have been put forth in the literature including the use of intermediate LOAs combining human and computer control over complex system functions. This approach seeks to define an assignment of functions to human operators and automation based on the capabilities of each server and attempts to achieve a team effort (Endsley, 1996). Other approaches have been defined including adaptive automation (Rouse, 1977), which involves dynamic system control allocations to a human or computer over time with the objectives of managing operator workload and facilitating SA. In the context of aviation systems, intermediate LOAs over flight control functions have been advocated for supporting pilot achievement of SA, as a result of the approach of defining meaningful task sets for pilots. This research sought to compare the effects of various LOAs on pilot performance, SA and
workload, including intermediate levels blending human and machine control over flight operations.

1.3. Taxonomies of LOAs

Hierarchies of LOAs have historically been developed with the objective of systematically presenting feasible forms of automation, ranging from manual control to full automation, and serving as a basis for postulating the effects of different levels on human performance, workload, and SA. This work includes Fitts (1962) man-and-machine-allocation lists and Sheridan and Verplanck’s (1978) hierarchy of levels of control for undersea teleoperation. Sheridan and Verplanck focused on how decision-making authority could be allocated between a human operator and computer controller, or who was ultimately in charge of the teleoperation system.

Other early automation research has defined and studied specific modes, or levels, of automation including Sheridan’s (1984) description of supervisory control. Supervisory control has subsequently been empirically and analytically examined (Baron, 1984; Rasmussen, 1984; Sarter & Woods, 1995). Others have studied decision support automation, or automation providing human operators with recommendations on task planning or task implementation (Rouse, 1984; Selcon, 1990).

Unfortunately, at this point in time, few comparisons have been made of the performance effects of various LOAs, such as supervisory control or decision support automation, using comparable experimental paradigms and response measures. One serious shortcoming of the current published research is that it does not provide insight into the potential performance, workload, and SA tradeoffs among various theoretical modes of automation.

Contemporary taxonomies of LOAs have been developed in order to address this research need by providing frameworks for systematically assessing performance and SA affects of various LOAs in synthetic tasks (Endsley & Kiris, 1995; Endsley & Kaber, 1999). Endsley and Kaber (1999) taxonomy (see Table 1) defines a broad range of LOAs spanning from manual control to full automation. The taxonomy was developed on the basis of identification of generic information processing functions found in many different complex systems including:

1. monitoring – observing system status/variable displays;
2. generating – task process planning based on the current state of the system and goals;
3. selecting – deciding among various processing plans by choosing an optimal alternative from the degree of accuracy or productivity required; and
4. implementing – response execution.

These functions are found across various domains including aircraft piloting, air traffic control and teleoperations. They represent human-machine system information processing functions that can be automated.

Each LOA in Endsley and Kaber’s (1999) taxonomy details a function allocation scheme involving these four human-machine information processing functions. In general, the lower LOAs (Action Support, Batch Processing) apply automation to the implementation aspect of a task; whereas, the intermediate LOAs (Shared Control, Decision Support, Blended Decision
Making) define function allocation schemes in which the computer has a role in formulating task processing strategy as well as selecting a strategy to use. Finally, under the high LOAs, the computer has some role in all information processing functions. One of the unique aspects of this taxonomy in comparison to historical works is that it does not strictly focus on how decision making authority is allocated between the human or automation, but it considers the distribution of the gamut of information processing functions between human and/or machine.

Table 1. Endsley and Kaber’s (1999) taxonomy of LOAs.

<table>
<thead>
<tr>
<th>Level of Automation</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monitoring</td>
</tr>
<tr>
<td>(1) Manual Control</td>
<td>Human</td>
</tr>
<tr>
<td>(2) Action Support</td>
<td>Human/Computer</td>
</tr>
<tr>
<td>(3) Batch Processing</td>
<td>Human/Computer</td>
</tr>
<tr>
<td>(4) Shared Control</td>
<td>Human/Computer</td>
</tr>
<tr>
<td>(5) Decision Support</td>
<td>Human/Computer</td>
</tr>
<tr>
<td>(7) Rigid System</td>
<td>Human/Computer</td>
</tr>
<tr>
<td>(9) Supervisory Control</td>
<td>Human/Computer</td>
</tr>
</tbody>
</table>

Endsley and Kaber (1999) compared the impact of the various LOAs described in their taxonomy in terms of human performance and SA in an abstract radar-monitoring task. This work followed earlier research by Endsley and Kiris (1995) who investigated the effect of a limited set of LOAs on the occurrence of OOTL performance problems in complex systems control. The taxonomy of LOAs studied by Endsley and Kiris was primarily applicable to cognitive tasks; whereas, Endsley and Kaber’s (1999) taxonomy has broad applicability to psychomotor and cognitive task performance. Endsley and Kaber’s results supported the use of intermediate LOAs as an approach to human-centered automation. They found that intermediate levels, including Shared Control, moderated operator workload and maintained operator SA leading to improvements in performance. Other findings included low-level automation producing superior performance, but this occurred at the cost of SA. In specific, when operators used the Batch Processing mode allowing them to develop task processing plans in advance of computer implementation, performance was greater than under all other LOAs, including full automation. Unfortunately, this mode did not facilitate operator concentration on the current state of the simulation. Opposite to this result, good SA was observed under higher LOAs, including Supervisory Control, which essentially freed-up operator cognitive resources for accurately perceiving the state of the simulation.

One shortcoming of these studies is inconsistency in the results on SA. This has been attributed to the different types of experimental tasks used in the laboratory settings. Another shortcoming is that the investigations have not been conducted under applied circumstances, which would pose more realistic levels of stress on operators. With this in mind, it is difficult to relate the results on the performance and SA effects of various theoretical LOAs to specific types of real-world systems and functions used in aviation operations. This issue is critical from the perspective of cockpit automation because of the wide variety of automated subsystems and their varying capabilities. An examination of the broad range of theoretical LOAs presented in
Endsley and Kaber’s taxonomy in a more applied context may be helpful in terms of providing insight into the performance and SA impacts of many of the modes of automation currently used in advanced commercial aircraft. For example, no research has investigated the effects of different LOAs afforded by the FMS in contemporary commercial aircraft on pilot performance, workload and SA.

2. OBJECTIVE/HYPOTHESES

2.1. Objective
The objective of this study was to assess the impact of a broad range of LOAs on performance, SA and mental workload in the context of high-fidelity simulation of an advanced commercial aircraft. This study was also intended to provide information on pilot interaction with high-level automation currently not available on commercial flight decks to serve as a basis for future cockpit interface design. In general, the study was to examine the effects of LOAs on performance of realistic simulations of flight tasks versus human performance of abstract laboratory tasks. One of the key steps to achieving this objective was to identify modes of cockpit automation (or FMS operation) representative of the theoretical LOAs defined in Endsley and Kaber’s taxonomy.

2.2. Approach
The approach to this research involved categorizing existing modes of the MD-11 autoflight system in terms of the theoretical taxonomy of LOAs developed by Endsley and Kaber (1999). In order to accomplish this, an in-depth analysis of cockpit systems and flight tasks was conducted, including:

(1) a review of the pilot’s reference manuals for the MD-11 (Delta Air Lines, Inc., 1998);
(2) familiarization with a high-fidelity personal computer (PC)-based flight simulator presenting an MD-11 aircraft;
(3) consultation with a Delta, MD-11 certified pilot; and
(4) development of mode-trees on each mode of the autoflight system in order to establish the capabilities of the system.

Mode-trees list the available functions of each mode and the default functional settings. This analysis resulted in formal descriptions of who (the pilot or autoflight system) does what under each mode of autoflight. The specific flight control functions identified for each mode were categorized in terms of the four, generic information-processing functions represented in Endsley and Kaber’s taxonomy. The ownership of the functions under each mode of MD-11 automation was then considered and the actual mode of aircraft automation was matched to a theoretical LOA in Endsley and Kaber’s taxonomy. These modes were subsequently modeled in a high-fidelity flight simulation.

In general, the actual modes of MD-11 automation corresponded to the lower LOAs defined in Endsley and Kaber’s taxonomy. In order to model the higher levels presented in the taxonomy and to investigate any affects on pilot performance, SA and workload, new modes of autoflight were conceptualized by assuming the availability of an expert system as part of the cockpit automation. These modes were also modeled through the high-fidelity flight simulator. The flight
simulator also included the majority of the autoflight subsystems as part of the MD-11 “glass” cockpit and manual attitude and throttle controls.

3. METHOD

3.1. Simulator Development

The Reconfigurable Flight Simulator (RFS) designed and developed by Pritchet et al. (1998) was used as a basis for the simulator development as part of this project. Figure 1 presents the displays of the original RFS. They include the out-of-cockpit view, the navigation (NAV) display, the primary flight display (PFD), and the RFS attitude and throttle controller. An enhanced Virtual Flight Simulator (VFS) was created including MD-11 autoflight subsystems, such as the multi-control display unit (MCDU), the flight control panel (FCP), and the Flight Mode Annunciator (FMA). The attitude and throttle controller of the RFS was also modified in the VFS. The controls were enlarged and the attitude controller was modified such that pilots could adjust the attitude of the plane and then depress a center button to temporarily “fix” the attitude at that position. Figure 2 presents the VFS displays.

Figure 1. Reconfigurable flight simulator displays.

Figure 2. Virtual flight simulator displays.
A select set of LOAs included in Endsley and Kaber’s taxonomy were modeled through the VFS, including a Manual Control mode without automated subsystem displays, a Manual Control mode with automated subsystem displays, an Action Support mode which required pilot flight planning and use of the FCP, and a Supervisory Control mode in which the FMS generated and implemented a flight plan, however, the pilot could intervene under the Action Support LOA assuming control of the aircraft flight planning at any time. The mode, Supervisory Control, represented a high-level of automation not currently available on commercial flight decks, such as the MD-11.

In order to model these theoretical LOAs using the VFS, the autoflight subsystems that would be used under each mode of automation were identified. The Manual Control mode without automated subsystem displays provided for complete pilot control of the simulated aircraft. Pilots were required to monitor flight operations and aircraft subsystem status in order to account for critical changes in scheduled flight events and/or system failures. Initially, pilots formulated and selected an appropriate flight plan and then manually implemented the flight plan. Under this mode of flight control, the cockpit displays and interfaces included the out-of-cockpit view, the NAV display, the PFD, and the manual throttle and attitude controls.

With respect to the Manual Control condition with automated subsystem displays, the allocation of flight control functions between the pilot and cockpit automation was identical to that for the other Manual Control mode; however, the displays presented through the simulator were augmented with the FCP, FMA and MCDU. The throttle and attitude controls for this mode were slightly different from those provided under the other manual mode, as previously mentioned. Although the FCP and MCDU were presented as part of the Manual Control condition with automated subsystem displays, they were not used.

Under the Action Support LOA, pilots and automation jointly monitored flight control processes and aircraft subsystem status. Pilots formulated, evaluated and selected an appropriate flight plan, and then implemented the selected plan using the FCP. The FCP allowed for indirect pilot control over aircraft flight surfaces. Automation provides some assistance to pilots in implementing a flight plan using the FCP. However, at any point during a flight, pilots could choose to use Manual Control for the purposes of corrective action. Under the Action Support mode, the flight simulator presented the same cockpit display configuration as for the Manual Control condition with automated subsystem displays. Although the MCDU and throttle and attitude controls were presented as part of the flight simulator interface, they were not used under this particular LOA.

The Supervisory Control LOA, defined in the context of the MD-11 flight deck, involved joint pilot and automation monitoring of flight control processes and aircraft system status. Under this mode, the automation generated and selected an appropriate flight plan, which was combined with fully-automated implementation of the plan. The pilot could perform corrective actions in the flight control if he or she intervened in the automated control by shifting the LOA to the Action Support mode and directing the aircraft using the FCP. (It may be possible in future automated aircraft for pilot intervention under a Supervisory Control mode to occur at the Decision Support LOA allowing for pilot reprogramming of a new flight plan using the MCDU.)
As previously mentioned, the Supervisory Control mode represented a LOA that is not currently available on the commercial flight deck but was examined in this study in order to provide a better sense of the impact of a broad range of automation on pilot performance. The simulator display configuration for the Supervisory Control mode was identical to that for the Action Support condition. Although the manual attitude and throttle controls were presented as part of the interface, they were not used under this particular LOA.

Beyond these LOAs, the simulator was also programmed to present a fully-autonomous flight mode in which pilots merely observed the state of the aircraft and flight plan generation, selection and implementation was automated. Under this mode of operation, pilots could not perform corrective actions in the flight control if they perceived them to be necessary. This mode also represented a LOA not currently available on the commercial aircraft flight deck. The same display configuration was used for the fully-autonomous mode of flight, as for the Supervisory Control condition. However, pilots under this mode used none of the control interfaces. It should be noted that only the Manual Control conditions, the intermediate LOA of Action Support, and the Supervisory Control mode were evaluated in this experiment.

3.2. Experimental Design
An experiment was a designed and conducted in which 4 subjects were assigned to each of the four LOA conditions presented through the VFS. Subjects were required to fly two flights under their assigned LOA.

Flight performance and situation awareness were assessed at three stop points during each flight. Workload was assessed using the NASA-Task Load Index (TLX) at the end of each flight.

3.3. Experimental Scenario
The experimental scenario was a generic re-creation of the American Airlines 965 crash in Cali, Colombia. Subjects were presented with 15 waypoints across three potential flight paths through Colombian airspace. All flights began at Barrenquilla, a city along the northern coastline at the Gulf of Mexico. The relevant waypoints were highlighted on enroute flight charts presented to subjects in advance of their test flights.

Prior to departure, subjects were asked to create a flight plan that crossed specific waypoints. Based on pre-determined flight path criteria, all subjects developed the same flight plan for each of the two flights. The required waypoints for the first flight were EJA, ABL, and ULQ. The required waypoints for the second flight were EJA, GIR, and CLO. The resulting flight plans included 3 enroute waypoints, 2 approach waypoints, and 1 runway waypoint (see Figure 3 for a graphic of the test flight plans). Subjects documented heading speed and attitude for each leg of the flight plan. Some subjects also documented latitude and longitude.
After completing a flight plan, subjects in the Supervisory Control condition were allowed to select a plan from a pre-programmed list of available plans. Subjects began all flights (training and experimental trials) at altitude (4000 feet) and at speed (125 knots).

During the flight, approximately 6-7 minutes prior to crossing the third enroute waypoint, an approach revision was given to subjects by an experimenter and they were required to modify their flight plan to a different approach path. In the first flight scenario, the flight path revision required subjects to alter the plan to include the waypoints GIR and CLO. In the second flight scenario, the approach revision required subjects to modify the flight to include the waypoints ABL and ULQ.

3.4. Hypotheses

Detailed hypotheses were formulated at the onset of this research for investigation through the experiment by making performance, SA and workload comparisons across the LOA conditions. In general, it was expected that Manual flight control performance subsequent to an approach revision would be superior to automated performance because of pilot involvement in the aircraft control loop in advance of the critical event. With respect to the two automated conditions, it was expected that pilot performance and SA would be significantly worse when high-level
automation (Supervisory Control) of flight control was integrated with the use of low-level automation (Action Support) subsequent to an approach revision, as compared to the use of Action Support throughout the entire duration of a test flight. This hypothesis was also based on pilot OOTL performance under Supervisory Control in advance of the critical flight event.

With respect to SA, it was expected that pilot perception, comprehension and projection would be significantly affected by the different LOAs on the basis of previous research using abstract synthetic tasks (Endsley & Kiris, 1995; Endsley & Kaber, 1999). More specifically, it was expected that low-intermediate LOAs (Action Support) involving pilots in the flight control loop and, at the same time, moderating pilot workload, would facilitate pilot achievement of SA.

Finally, in regard to pilot workload, it was expected that high LOAs would yield lower ratings of perceived mental demand than the Manual Control modes. It was also expected that workload associated with Manual Control using automated subsystem displays would be greater than Manual Control using the RFS display configuration because of the larger number of displays for pilots to monitor.

3.5. Subjects
Sixteen pilots from the North Carolina area were paid to participate in the study. The selection criteria included a private pilot’s license (commercial and military pilots also were also accepted), 20/20 or corrected to normal visual acuity, and PC experience. Pilots were retained for the experiment based on training in the Multi-Attribute Task (MAT) Battery (Comstock & Arnegard, 1990), a successful training flight in Manual Control mode, and a successful training flight under the assigned LOA. Table 2 presents profiles for each of the subjects in terms of flight experience, license type, and PC experience. All of the subjects were male.
Table 2. Subject profiles.

<table>
<thead>
<tr>
<th>Subject</th>
<th>LOA</th>
<th>Age</th>
<th>Rating of PC Use (1 = “low” to 5 = “high”)</th>
<th>License Type</th>
<th>Flight Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Manual Control (VFS)</td>
<td>19</td>
<td>5</td>
<td>Private</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>Manual Control (VFS)</td>
<td>19</td>
<td>4</td>
<td>Private</td>
<td>200</td>
</tr>
<tr>
<td>3</td>
<td>Manual Control (VFS)</td>
<td>20</td>
<td>5</td>
<td>Private/Military</td>
<td>200</td>
</tr>
<tr>
<td>4</td>
<td>Manual Control (VFS)</td>
<td>22</td>
<td>4</td>
<td>Private</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>Action Support</td>
<td>20</td>
<td>5</td>
<td>Private</td>
<td>80</td>
</tr>
<tr>
<td>7</td>
<td>Supervisory Control/Action Support</td>
<td>42</td>
<td>5</td>
<td>Private</td>
<td>67</td>
</tr>
<tr>
<td>8</td>
<td>Action Support</td>
<td>23</td>
<td>3</td>
<td>Private</td>
<td>72</td>
</tr>
<tr>
<td>9</td>
<td>Supervisory Control/Action Support</td>
<td>26</td>
<td>4</td>
<td>Private</td>
<td>120</td>
</tr>
<tr>
<td>10</td>
<td>Action Support</td>
<td>21</td>
<td>4</td>
<td>Commercial</td>
<td>650</td>
</tr>
<tr>
<td>11</td>
<td>Action Support</td>
<td>21</td>
<td>4</td>
<td>Commercial/Instructor</td>
<td>300</td>
</tr>
<tr>
<td>14</td>
<td>Action Support</td>
<td>21</td>
<td>5</td>
<td>Private</td>
<td>80</td>
</tr>
<tr>
<td>15</td>
<td>Manual Control (RFS)</td>
<td>21</td>
<td>5</td>
<td>Private</td>
<td>70</td>
</tr>
<tr>
<td>16</td>
<td>Manual Control (RFS)</td>
<td>20</td>
<td>5</td>
<td>Private</td>
<td>150</td>
</tr>
<tr>
<td>17</td>
<td>Manual Control (RFS)</td>
<td>21</td>
<td>4</td>
<td>Commercial</td>
<td>600</td>
</tr>
<tr>
<td>18</td>
<td>Supervisory Control/Action Support</td>
<td>23</td>
<td>5</td>
<td>Private</td>
<td>70</td>
</tr>
<tr>
<td>19</td>
<td>Supervisory Control/Action Support</td>
<td>30</td>
<td>5</td>
<td>Commercial</td>
<td>600</td>
</tr>
<tr>
<td>20</td>
<td>Manual Control (RFS)</td>
<td>21</td>
<td>5</td>
<td>Private/Military</td>
<td>140</td>
</tr>
</tbody>
</table>

3.6. Apparatus
A high-performance Intergraph graphics workstation was used to run the flight simulator. The computer included dual-550 MHz Pentium III processors with 0.5 GB of RAM and Dual-Intense3D Wildcat 4105 Graphics boards. Each graphics board was used to drive a 21” graphics monitor at 1024 x 768 resolution in a dual-monitor configuration. The system was integrated with a mouse and a standard keyboard. A separate standard PC with a 17” monitor was used to run the MAT battery during training sessions.

3.7. Procedures
The experiment was completed during two sessions on two separate days. The procedures for the first day of the experiment were as follows:

(1) Familiarization with the experiment and initial questionnaires and forms - Subjects were required to complete a consent form, payment form, and background questionnaire. They were then familiarized with the equipment including computers, displays, and controls to be used in the experiment.

(2) Training in the MAT battery - Subjects were given instructions on the systems monitoring, tracking, and fuel management subtasks and then completed two ten-minute training trials.

(3) Training in the flight simulator - This training session lasted approximately 90 minutes. Subjects were given instructions on the flight deck interfaces. They were then presented with aviation charts, associated measurement techniques, and information on flight constraints. The experimenter demonstrated both manual and automated piloting using the simulator.
Subjects were given an opportunity to practice manual and automated piloting using the simulator. During this session, subjects were allowed to ask questions about the features or capabilities of the flight deck interfaces.

(4) Training flight under the assigned LOA - Finally, subjects were asked to complete a 30-minute training flight under their assigned LOA. As part of this session, pilots were required to generate a flight plan based on given flight constraints and using the flight charts.

The procedures for the second day of the experiment were as follows:

(1) Re-training on simulator displays and controls - Subjects were given a 10-minute refresher on the displays and controls. They were also allowed to ask questions if they were unsure about any functions of the simulator.

(2) Manual flight practice - Subjects practiced manual flight control following a pre-determined flight plan for approximately 30 minutes. They were required to strictly adhere to speed and altitude constraints. Subjects that were unable to complete the flight under the Manual Control mode within 3 trials were terminated from the experiment.

(3) Experimental trial instruction - Subjects were given instructions on the use of the NASA-TLX and Situation Awareness Global Assessment Technique (SAGAT) queries to be administered at each stop point. Subjects were asked to complete the NASA-TLX factor comparison form. They were also shown sample SAGAT queries.

(4) Automated flight practice - Subjects generated a flight plan and practiced flight under an automated condition for approximately 30 minutes.

(5) Experimental trials – An experimenter explained the procedure for the experimental trials including control responsibility, pre-departure procedure etc. Subjects then completed two 45-minute experimental trials under the assigned LOA. A 10-minute break was provided between trials.

3.8. Response Measures

The response measures recorded during the experiment included performance measures, SA measures, and workload measures. Performance was recorded in terms of speed deviations from planned speed, heading deviations from planned headings, altitude deviations from planned altitudes, and the completeness of a pilot’s flight plan. Pilot SA was recorded using the SAGAT method (Endsley, 1988; 1995). Finally, workload was recorded using the NASA-TLX (Hart & Staveland, 1988).

3.8.1. Simulation Freezes for Measuring SA and Flight Performance

At three points during each flight, the simulation was frozen to collect both performance and SA data. The freezes occurred at the second en route waypoint (prior to the approach revision), the fourth waypoint (directly after the approach revision), and the sixth waypoint marking the close of the flight (see Figure 4).
3.8.2. Flight performance measures

Deviations from planned flight were recorded at the three stops during each of the two experimental trials. Each time the simulation was frozen, the experimenter recorded the current altitude in feet, the heading in degrees, and the speed in knots. The absolute value of the difference between the actual and planned headings, speeds, and altitudes was calculated. The heading deviation also considered the direction of rotation (clockwise or counter clockwise) from the planned value to the actual heading that yielded the smallest offset. Since pilots were generally correct in their flight planning (unless they failed to complete a plan), the performance measures generally reflected deviations from ideal performance as well as from their plan. In the case of an incomplete flight plan, the ideal values for aircraft heading, speed and altitude were used as bases for comparison with actual flight parameters.

3.8.3. Flight planning

Performance was also measured in terms of completeness of the flight plan before the flight and completion of a flight plan at the approach revision. Flight plans were evaluated by checking to see that pilots recorded waypoint names and the speed, heading, and altitude for each leg of the simulated flight. The percentage of complete entries in a plan was calculated by dividing the number of data points that a pilot recorded by the number of data points that should have been
recorded both for the pre-flight plan and for the revised approach plan. Separate scores were recorded for the two experimental trials.

3.8.4. SA Measures
With respect to SA, an adaptation of the Commercial Airline version of the SAGAT was used (SA Technologies, Inc., 2000). At each of the simulation stops, the VFS display was blanked and subjects were posed with 9 SA questions. Each question was presented on a separate sheet of paper and subjects simply wrote their responses. Three of the questions were classified as Level 1 SA, or perception, questions; three questions were considered Level 2 SA, or comprehension, questions; and three questions were classified as Level 3 SA, or projection of future system state, questions. The questions presented to subjects were randomly selected from larger pools of questions for each of the levels of SA. In total, there were 11 Level 1 SA questions, 15 Level 2 SA questions, and 5 Level 3 SA questions used in the experimental tests. Subjects were given unlimited time to answer the questions. Once all of the questions were answered at a simulation stop, subjects returned to the simulation and the flight began at the point at which it was frozen. An experimenter recorded the correct responses to the SA questions based on the state of the flight simulator at the time of the freeze while subjects were not looking at the VFS displays and were answering the SA questions.

3.8.5. Workload
Workload was recorded using the NASA-TLX at the end of each of the two trials. The NASA-TLX factor comparison form was completed immediately after the flight training. The NASA-TLX rating forms were completed immediately after the finish of each experimental trial.

4. DATA ANALYSIS

4.1. Level of Automation
In all analyses, LOA was treated as a between-subjects factor. The four settings of the independent variable included Manual Control of the RFS (MRFS), Manual Control of the VFS (MVFS), Action Support (AS), and Supervisory Control permitting pilot intervention in flight direction under the Action Support mode (SC/AS). Subjects in the SC/AS condition used Supervisory Control prior to the approach revision and Action Support after the revision. This condition was treated as a LOA unique from the AS condition. With respect to comparisons of flight performance and SA observations after the approach revision (at the second and third simulation stops), subjects in the AS and SC/AS conditions actually operated under the same LOA (Action Support), however, this was preceded by performance under different LOA conditions at the outset of experimental trials.

4.2. Flight Performance Measures
All flight performance measures were analyzed through a two-way Analysis of Variance (ANOVA) with LOA and simulation Stop as between- and within-subjects variables, respectively. To ensure that the underlying assumptions of the ANOVA were upheld by the data sets, transforms were applied to the various response measures according to the procedures described by Neter et al. (1990, pp. 142-146). In the case of the flight performance measures, a square-root transformation of the heading, speed and altitude deviations was used to address potential non-normality and the non-constant variance violations due to the lower bounded data sets.
4.3. **Flight Planning Performance Measure**

Observations on flight planning were grouped into three separate data sets. One data set included both the pre-flight and post-revision flight plan data. The second data set included only the pre-flight plan data, and the third data set included only the post-revision flight plan data. The first data set containing both pre-flight and post-revision data was analyzed using a two-way ANOVA with LOA and plan (pre-flight or post-revision) treated as between- and within-subjects variables. The two data sets containing only pre-flight or post-revision observations were analyzed with using a one-way ANOVA with LOA as the independent variable. Because of a large number of perfect (100% complete) flight plans, the data in all three sets exhibited non-constant variance and non-normality. An arcsine transform of the flight planning data improved the nature of the data with respect to the ANOVA assumptions.

4.4. **Situation Awareness Measures**

Situation Awareness was measured using the SAGAT technique described previously. The percent correct responses to queries on each level of SA were analyzed along with overall SA, which was calculated as the percent correct responses to all nine questions at any stop. All scores were calculated using two different methods. First, an answer was considered correct only if the subject’s answer exactly matched the answer recorded by the experimenter. In the second method, the answer was considered correct if it was within an accepted industry tolerance of the answer recorded by the experimenter. Consequently, two scores were calculated and analyzed for each trial, for each subject, for Level 1 SA, Level 2 SA, Level 3 SA, and Overall SA at each of the three SAGAT stop points.

Because the MRFS condition did not present the same displays as the other three LOA conditions, several of the SA queries designed for the experiment were not applicable to that condition. Three of 11 possible Level 1 SA questions were not applicable to the MRFS condition, and three of five possible Level 3 SA questions were not applicable to the MRFS condition. All of the Level 2 SA questions were applicable to all conditions. With this in mind, the SA data was analyzed using two different approaches. The first approach includes the MRFS data along with the SA data recorded under the other LOA conditions. Subject responses to queries that were not applicable to the MRFS condition were removed from the data set. Consequently, the mean SA scores for the MRFS condition were likely based on a smaller sample of responses. The second approach to the SA data analysis excluded the MRFS data. Since there were a relatively large number of observations on Level 1 SA and the Overall SA measure recorded under the MRFS condition, this data was analyzed using both approaches. As well, since all the Level 2 SA questions were applicable to the MRFS condition, pilot comprehension data was analyzed including observations under the MRFS condition and, to be consistent with the other data analyses, it was also examined without the observations on the MRFS condition. However, since so few of the Level 3 SA questions were applicable to the MRFS condition, pilot projection of future aircraft states was analyzed without considering the observations under the MRFS condition.

Consequently, seven data sets using the exact scoring method, and seven data sets using the tolerance scoring approach, were analyzed using a two-way ANOVA with the between-subjects variable LOA and the within-subjects variable Stop. An arcsine transform was performed on the
Overall SA data in order to address potential constant variance and normality assumption violations. With respect to the SA data that was analyzed according to level, the arcsine transform did not improve the nature of the data in terms of non-constant variance and non-normality.

4.5. Workload Measures
Pilot perceptions of cognitive workload were captured using the NASA-TLX (Hart and Staveland, 1988). Overall workload was computed based on the individual ratings for mental demand, physical demand, temporal demand, performance, effort, and frustration and the paired comparison weightings for the individual factors. Overall workload and a subset of the individual demand components (mental demand, temporal demand, and frustration) were analyzed with a one-way ANOVA with LOA as the independent variable.

5. RESULTS
All statistical results reported here are based on the transformed responses described in the previous section. The graphs, however, are presented in the original response units to promote ease of interpretation and understanding (Neter et al., 1990, p. 147).

5.1. Performance Measures
5.1.1. Flight performance measures
Analysis of variance results on the arcsine transform of heading deviation revealed no significant effect due to either the LOA condition or Stop. However, the ANOVA on the arcsine transform of altitude deviation revealed a significant effect of LOA \((F(3,12) = 12.85, p<0.001)\). The average altitude deviations at each stop are shown in Figure 5. Figure 6 shows the performance of individual subjects in terms of managing aircraft altitude according to planned flight levels throughout a flight under the various automation conditions. Duncan’s Multiple Range (MR) tests showed that the two manual conditions produced significantly greater altitude deviations than the two automated conditions \((p<0.05)\). In addition, altitude deviations under the SC/AS condition were significantly greater than those in the AS condition \((p<0.05)\).

The data analysis on altitude deviations also revealed a marginally significant interaction of LOA and Stop, when using the Mean Square Error (MSE) as a denominator in the F-statistic \((F(6,48) = 2.18, p<0.10)\). (However, the interaction effect only appeared to be a trend when individual differences were accounted for in the error term of F-test.) Duncan’s MR test on the interaction effect of LOA and Stop revealed that Manual Control of the RFS was worse in the early stages of flight (Stops 1 and 2) than it was for Stop 3 \((p<0.05)\). This indicates that performance improved after the approach revision under the Manual Control mode, whereas performance showed no improvement under the automatic conditions. In addition, performance under the Action Support mode, as part of the SC/AS condition, towards the close of the experimental flights (Stop 3) was significantly worse than performance under the Action Support condition at the beginning of flight (Stop 1).
In general, the altitude deviation results revealed that computer assistance in the form of the two automation conditions led to better performance than Manual Control. In addition, performance under the MRFS condition improved throughout the trial. This may have been due to pilots becoming more engaged in the flight task following the approach revision. Another possible reason for improvement in altitude deviation under the MRFS condition was that by the third stop, the pilots were beginning their approach to Cali and they may have refined their internal performance criteria as they approached the landing. However, it is unclear why similar effects were not seen in the MVFS condition. With respect to both of the automated flight control conditions, AS and SC/AS, altitude deviations appeared to gradually increase through the trial, though the change was not significant.

The difference in altitude deviation performance in the automated conditions is difficult to interpret. In general, it appears that computer assistance provided by the FCP under the AS condition may have been more accurate for flight path control than the computer assistance provided by the MCDU under the SC/AS condition. However, this does not explain why differences in performance under the AS and SC/AS conditions persisted (or increased) at Stops 2 and 3 when subjects assigned to both conditions were using the FCP to control the simulator. It is possible that subjects in the Action Support condition were more engaged in the task at the time of the revision and made more accurate inputs to the FCP in order to handle the revision. Or, subjects in the SC/AS condition may have suffered from OOTL performance problems as a result of being required to operate the aircraft under SC in advance of the approach revision and then to shift to the Action support mode after the critical flight event. This may have caused slightly poorer performance over the last two stops in comparison to the AS condition.
The ANOVA on the arcsine transform of speed deviation revealed a significant effect due to Stop ($F(2,24) = 4.24$, $p<0.05$). According to Duncan’s MR test, performance at Stop 3 was significantly worse than performance at Stop 1 ($p<0.05$) (see Figure 7). This effect indicates that performance with respect to speed control was affected by the approach revision.

Figure 7. Mean speed deviation at each stop under each LOA condition.
There were no significant differences in speed control performance across the LOA conditions. In general, speed appeared to be easier to control under both the manual and automatic modes of flight control, as compared to altitude and heading control.

5.1.2. Flight Planning Performance

Figure 8 presents the mean percent completion of flight planning pre-flight and post-approach revision across the four LOAs. The ANOVA on the arcsine transform of the percent completion of both pre-flight and post-revision plans revealed a significant effect due to plan (F(1,12) = 5.10, p<0.05). Pre-flight planning appeared to be more complete than plans prepared subsequent to the approach revision (p<0.05). There were no significant differences in planning due to the LOA.

However, the ANOVA on the arcsine transform of the percent completion of only pre-flight plans revealed a significant effect due to LOA, when using the MSE as a denominator in the F-statistic (F(3,16) = 7.18, p<0.01). (However, this effect only appeared to be a trend when individual differences were accounted for in the error term of F-test (F(3,12) = 1.85, p=0.19).) Duncan’s MR tests on the data indicated that pre-flight planning under the SC/AS condition was significantly worse (p<0.05 using the MSE as an error term) than under all other conditions. Similar analyses on the post-revision flight planning showed no significant differences in the degree of plan completion across the different LOAs.

These results indicate that pre-flight planning was poorer when pilots initiated the flight under high LOAs. Pilots were aware that the flight plan data was stored in the FMS and may have placed less emphasis on the importance of completing the flight plans manually. The results also indicate that, in general, flight planning was poorer when completed during the flight after the approach revision than it was when completed pre-flight.

Figure 8. Mean percent complete flight plans prepared during pre-flight procedures and subsequent to the approach revision under each LOA.
5.2. **Situation Awareness**

The analysis of the SA data when the percentage of correct responses was determined using accepted industry tolerances revealed no significant effects due to LOA or Stop. However, some significant results and trends were observed when query grading required subjects’ answers to exactly match the correct answers recorded by the experimenter.

5.2.1. **Overall SA**

The mean Overall SA scores across LOA and Stops are shown in Figure 9. The ANOVA on the arcsine transform of the Overall SA response, including observations on the MRFS condition, revealed a significant effect due to Stop ($F(2,12) = 4.67, p<0.05$). Duncan’s MR tests on the Stop main effect indicated that Overall SA at the beginning of flights (Stop 1) was significantly better than the Overall SA directly following the approach revision (Stop 2) ($p<0.01$). When Overall SA was analyzed without considering the MRFS data, a similar effect was observed ($F(2, 12) = 3.86, p<.05$) and the Duncan’s MR tests indicated that performance at Stop 1 was significantly better than performance at both Stops 2 and 3 ($p<0.05$).

![Figure 9. Mean overall SA at each stop under the various LOA conditions.](image)

This data indicates that, for all LOA conditions, SA following the approach revision was worse than SA prior to the revision. The data also appears to indicate that recovery from any post-revision SA deficit was better under the MRFS condition than under any other condition; however, there was no significant LOA by Stop interaction effect to support this assertion.

5.2.2. **Level 1 – Perception**

The mean percent correct responses to Level 1 SA queries across LOA and Stops is shown in Figure 10. The ANOVA on the Level 1 SA data, including the MRFS observations, revealed a significant effect due to Stop ($F(2,12) = 3.86, p<0.05$). According to Duncan’s MR tests, Level 1 SA at Stop 2 (directly subsequent to the approach revision) was significantly worse than Level 1 SA at Stops 1 and 3 ($p<0.05$). A similar trend was present when the MRFS data was excluded from the analysis, although the effect was only marginally significant ($F(2,12) = 2.94, p<0.10$).

Pilot perception of aircraft states decreased from 44% correct at the beginning of flight (Stop 1) to only 26% correct at Stop 2. Pilots appeared to recover from any SA deficit at Stop 2 by the...
time of Stop 3 (the close of the flight) with an average 43% correct responses to queries. Although not significant, this decrease in perception, followed by recovery, appeared to be much greater under the Action Support condition than under the MVFS and SC/AS conditions.

Finally, post-hoc tests also revealed that Level 1 SA under the SC/AS condition was significantly better (p<0.05) at the close of experimental flights (Stop 3), when pilots were using the FCP to control the aircraft, than at the beginning of flight (Stop 1), when they were using the MCDU for planning, programming and implementation purposes. In fact, performance under the SC/AS condition at Stop 3 was approximately equivalent to performance under the SC/AS condition by the close of test trials.

The general trend on the Level 1 SA data (somewhat reflected in the Overall SA data) reveals heightened perception at the onset of flight and just prior to landing compared to the time frame after the in-flight approach revision. The approach revision and the need to modify the flight plan likely diverted subjects attention from the simulator displays. Improvements in SA toward the close of the flight reflected the preparation of pilots for landing by attending closely to flight parameters and displays.

![Figure 10. Mean Level 1 SA at each stop under the various LOA conditions.](image)

5.2.3. **Level 2 – Comprehension**

The mean percent correct responses to Level 2 SA queries across LOA and Stops are shown in Figure 11. The ANOVA on the Level 2 SA response, including observations on the MRFS condition, revealed no significant effects due to LOA or Stop. When the MRFS condition was removed from the data set, the ANOVA revealed a marginally significant effect due to Stop (F(2,12) = 3.05, p<0.10) and Duncan’s MR tests indicating that Level 2 SA at Stop 3 was significantly worse than pilot comprehension at Stop 1 (p<0.05).
The results on the Level 2 and Overall SA response measures indicate that pilot comprehension of aircraft states and SA, in general, were better prior to the approach revision than afterwards. One explanation for this is that subject workload at the final stop was higher due to landing activities and may have subtracted from cognitive resources available for comprehending the state of the aircraft.

Although there were no significant interactions of LOA and Stop, there appeared to be a decrease in Level 2 SA under the Action Support condition over the course of the flight (a drop of about 40%) that was not present in the other conditions.

In comparison to the SC/AS condition, the AS condition yielded lower Level 2 SA. It was expected that SA for the SC/AS condition at the final SAGAT stop during the course of flight would be worse than under all other experimental conditions due to pilot OOTL performance in advance of the approach revision. It may be possible that the lower pilot SA observed under the AS condition was due to pilot fatigue or boredom during the test trials.

5.2.4. Discussion of SA measure

Although performance measures were impacted by the LOA, the SAGAT did not appear to be sensitive to the experimental manipulation. One possible reason for this may be that the laboratory condition did not subject pilots to the same stress levels associated with flight control that accompany a real approach revision. Another possibility is that the SAGAT queries target information that commercial pilots need to know for flight control, but, since naïve pilots were used in the study, the queries may not have been relevant to their experiences. If this is the case, the measure may be more sensitive if actual MD-11 pilots with accurate understanding of the use of FMS interfaces and information displays were used in an experiment. It is also possible that the flight trials were not long enough for effects of complacency, vigilance decrements, or the OOTL performance problems to manifest themselves in terms of pilot SA.
5.3. **Workload**

The ANOVA results on overall workload revealed a significant effect due to LOA when the MSE was used as a denominator in the F-statistic ($F(3,48) = 30.63, p<0.0001$). This effect appeared only as a trend ($F(3,12) = 2.07, p=0.16$), when individual differences were accounted for in the error term of the F-test. Duncan’s MR test (using the MSE as an error term) revealed that the overall workload for the MRFS condition was significantly greater ($p<0.05$) than the overall workload perceived under the other three conditions. The post-hoc procedure also indicated that the overall workload for the MVFS condition was significantly lower ($p<0.05$) than the overall workload for the other three conditions.

An ANOVA on the mental demand component of the TLX revealed a marginally significant effect due to LOA ($F(3,12 = 3.32, p<0.10$), when individual differences were accounted for in the error term of the F-test. Duncan’s MR test revealed that perceived mental workload under the MRFS condition was significantly greater than the mental load for the MVFS and SC/AS conditions ($p<0.10$). Figure 12 shows the mean demand ratings across the four LOAs. All other demand components of the NASA-TLX showed no significant effect due to LOA.

![Figure 12. Mean mental workload ratings across the four LOAs. (Condition means with lines above them appearing at the same level are not significantly different each other.)](image)

These results indicate that manual control of flight without access to the automated subsystem displays posed the highest workload. The control interface for the MRFS condition appeared to be more difficult to operate than the control interface for the MVFS condition. The MVFS condition allowed for bank and pitch to be temporarily fixed, offloading the pilot for other activities. Also, it was easier under the MVFS condition to find flight parameter information and to obtain feedback on the state of the aircraft than it was under the MRFS condition. The highest LOA (SC/AS) posed the lowest workload, however this was not significantly different from the loads posed by the MVFS or AS conditions.

These results corresponded with the flight performance and SA effects of LOA. In specific, the increase in perceived workload under the MRFS condition corresponded with improvements in
manual flight performance and perception as part of pilot SA during the later segments of simulated flights (subsequent to the approach revision).

6. CONCLUSIONS

This work has demonstrated that various modes of commercial aircraft automation and futuristic cockpit automation may differ in terms of effects on pilot performance, SA and workload. In particular manual control involving traditional flightdeck displays appears to allow for superior performance after a critical flight event in comparison to low- and high-level automation, but this comes at the expense of high workload. Furthermore, automation, in general, may compromise pilot SA after a critical event as compared to manual control, which may allow for recovery of pilot perception.

High-level automation not currently available in commercial aircraft appears to compromise pilot flight planning and cause reliance on pre-programmed FMS databases. This and the other results of the present study may have direct applicability to flight decks, in comparison to those of previous studies involving low-fidelity simulations.

7. REFERENCES


