Analysis of Pilots Monitoring and Performance on Highly Automated Flight Decks

Authors:
R. Mumaw[1]
N. Sarter[2]
C. Wickens[3]
S. Kimball[1]
M. Nikolic[2]
R. Marsh[3]
W. Xu[1]
X. Xu[3]

[1]Boeing Commercial Airplane
P.O. Box 3707, MC 67-TC
Seattle, WA 98124-2207

[2]Ohio State University
Columbus, OH

[3]University of Illinois
Champaign-Urbana, IL

NASA Ames Contract NAS2-99074
Final Project Report

December 15, 2000

Table of Contents

List of Tables

List of Figures

Executive Summary

A key issue for enhancing the safety of flight operations is to support more effectively the interaction between flight crews and automated flight deck systems (in particular, autopilot, autothrottle, and the Flight Management Computer), which exist in a large percentage of the current U.S. fleet. Breakdowns in this interaction have been discussed in great detail in the human factors literature (e.g., Abbott et al., 1996; Sarter & Woods, 1997; Woods & Sarter, 2000), and they have played a role in a number of recent aircraft incidents and accidents. These breakdowns have often been attributed to a combination of inadequate automation feedback and pilots buggy mental models; together, these two factors have been assumed to lead to monitoring failures. However, only subjective and anecdotal data are available on the actual monitoring behavior of flight crews on glass cockpit aircraft. It is not known what monitoring strategies flight crews adopt to track automation behavior, how effective these strategies are, and under what circumstances they tend to break down. Consequently, it is difficult to identify appropriate counter-measures to observed failure in pilot-automation interaction.

The present study directly addressed this gap: eye-tracking, as well as performance, data were collected while pilots were flying a challenging scenario in a high-fidelity simulator of a modern automated aircraft (Boeing 747-400). In addition, during debriefing, pilots were asked questions about various features and functions of the autoflight system to identify gaps or misunderstandings in their model of the system to help us explain their behavior and performance. The combination of these data allowed us to better understand when and why monitoring and coordination break down. It also helped identify possible design changes for enhancing the performance of the joint human-machine team.

We recruited 20 747-400 pilots (10 Captains, 10 First Officers) from two U.S. airlines. In collaboration with one of the participating airlines, we developed a 1-hour scenario (SFO to LAX) that allowed us to evaluate monitoring of the flight deck interface. In general, the scenario events were situations that required a thorough understanding of flight deck automation, and they were designed to both broadly assess and challenge pilots monitoring skills. It is important to note that these situations represent routine flight operations that is, we were not relying on uncommon failures or unusual clearances.

The performance data revealed that many of these experienced 747-400 pilots had difficulty with a number of the scenario events involving proficient use of automation or automation monitoring. In some cases, subjects failed to detect meaningful indications that were present (e.g., a change in a waypoint altitude constraint) or to look for those indications (e.g., FMA changes). In other cases, subjects failed to understand the implications of the indications they monitored (e.g., the implications of VNAV ALT during cruise).

These findings bolster the literature suggesting that the flight deck interface fails to provide a complete picture of the automation s authority and control actions, and moreover, they strengthen the call for more effective schemes for data-driven monitoring. Further, these findings reveal the importance of pilots having strong expectations of the automation s behavior to support knowledge-based monitoring. Because much of the burden of automation monitoring falls on the pilot, pilots need a more complete and accurate mental model of automation behavior.

We used an open-ended approach to exploring subjects mental models. Our results show that when subjects offer statements about automation behavior, they usually articulate correct statements (about 92% of the time). However, subjects either do not know about certain aspects of automation behavior or were unwilling to offer other statements in this context. Further, even when subjects made a correct statement about which mode to expect in a certain situation, they did not always detect inappropriate modes in those situations. Therefore, their knowledge may be poorly linked to the operational context.

The monitoring data showed some predictable trends at a high level of analysis that can be used to characterize the statistical properties of glass cockpit scanning. For example, the Flight Mode Annunciations (FMAs) do not seem to be part of the pilot s routine scan pattern. Interestingly, there was considerable diversity of scan patterns across pilots, and it was difficult to find scanning patterns that were clearly better than others as had been possible to do in the studies of conventional aircraft scanning (Bellenkes et al., 1997).

Analysis of Pilots Monitoring and Performance on Highly Automated Flight Decks

1. Introduction

The overall objective of this research is to contribute to further improvements of the already high level of safety of flight operations in commercial aviation. In particular, our work focuses on possible ways to support more effective communication and coordination between pilots and highly automated autoflight systems on modern glass cockpit aircraft. This communication and coordination rely primarily on pilot monitoring of relevant indications and the interface s presentation of indications.

Breakdowns in the interaction between pilots and the autoflight system have been discussed in great detail in the human factors literature (e.g., Abbott et al., 1996; Sarter & Woods, 1997; Woods & Sarter, 2000), and they have played a role in a number of recent aircraft incidents and accidents. It appears that improvements are needed in at least two areas to address the problem. One area of improvement are new forms of pilot training that support the formation of a more complete and accurate mental model of the automation and thus improve flight crews ability to anticipate, track, and understand the status and behavior of the automation (Mumaw et al., 2000). Current pilot training rarely emphasizes mental model development. The second area is more effective automation feedback that better alerts system-initiated changes and events and aids pilots in visualizing the implications of system status for aircraft behavior (Norman, 1990). Improvements in both training and system feedback can be expected to contribute to more effective monitoring of highly complex automated systems.

It is important to note that, even though monitoring failures are widely assumed to be responsible for most of the observed problems with pilot-automation coordination, only subjective and anecdotal data are available on the actual monitoring behavior of flight crews on glass cockpit aircraft. It is not known what monitoring strategies flight crews adopt to track automation behavior, how effective these strategies are, and under what circumstances they tend to break down. Consequently, it is difficult to identify appropriate countermeasures to observed failure in pilot-automation interaction. The present study directly addresses this gap: eye-tracking, as well as performance, data were collected while pilots were flying a challenging scenario in a high-fidelity simulator of a modern automated aircraft. As part of a debriefing, pilots were asked questions about various features and functions of the autoflight system to identify gaps or misunderstandings in their model of the system that might explain their behavior and performance. The combination of these data allows us to better understand when and why monitoring and coordination break down. It also helps identify possible design changes for enhancing the performance of the joint human-machine team.

1.1 Flight Deck Automation and Mode Awareness

A key issue for enhancing the safety of flight operations is to support more effectively the interaction between flight crews and automated flight deck systems (in particular, autopilot, autothrottle, and the Flight Management Computer), which exist in a large percentage of the current U.S. fleet. The introduction of automation to the modern flight deck has led to a number of anticipated benefits, such as an increased precision and efficiency of operations. However, these benefits appear to accrue primarily in situations where the automation performs tasks on its own that is, without pilot involvement. In circumstances that require cooperation and coordination between pilots and the system, unexpected problems are being encountered (Sarter, 2000).

A number of recent studies (e.g., Lyall & Funk, 1998; Sarter & Woods, 1992, 1994, 1995, 1997; Wiener, 1989) have demonstrated that pilots can become confused about the state and behavior of the flight deck automation. One of the most often observed and widely acknowledged consequences of breakdowns in pilot-automation coordination is a loss of mode awareness on the part of the pilot (Abbott et al., 1996; Sarter & Woods, 1994, 1997, in press; Sarter, Woods, & Billings, 1997). Mode awareness refers to the knowledge and understanding of the current and future status and behavior of the automation. A loss of mode awareness can lead to mode errors and automation surprises. Mode errors, generally speaking, occur when a pilot performs an action that is appropriate for the assumed system state but not for the actual state. Or, a mode error can refer to the omission of a required action or intervention with automation actions. Mode errors lead to automation surprises when the pilot notices that the automation is engaged in activities that were not commanded or intended. Both mode errors and automation surprises have played a role in recent incidents and accidents involving glass cockpit aircraft. Several reports (Abbott et al., 1996; BASI, 1998) have established that, at a minimum, mode errors and automation surprises can lead to poor or slow compliance with air traffic control (ATC) clearances (in particular, deviations from assigned altitudes).

Several factors can contribute to a loss or lack of mode awareness:
- an incomplete and/or inaccurate mental model of the flight deck automation.
- automation feedback that is inadequate because it fails to support pilots in predicting, assessing, and understanding the current state (and changes in state) and behaviors of the system.
- a highly complex logic underlying flight deck automation behavior that differs from pilots reasoning about their flying tasks and which, in addition, differs considerably across manufacturers, aircraft types, and in some cases, across individual planes within type (due to software upgrades).

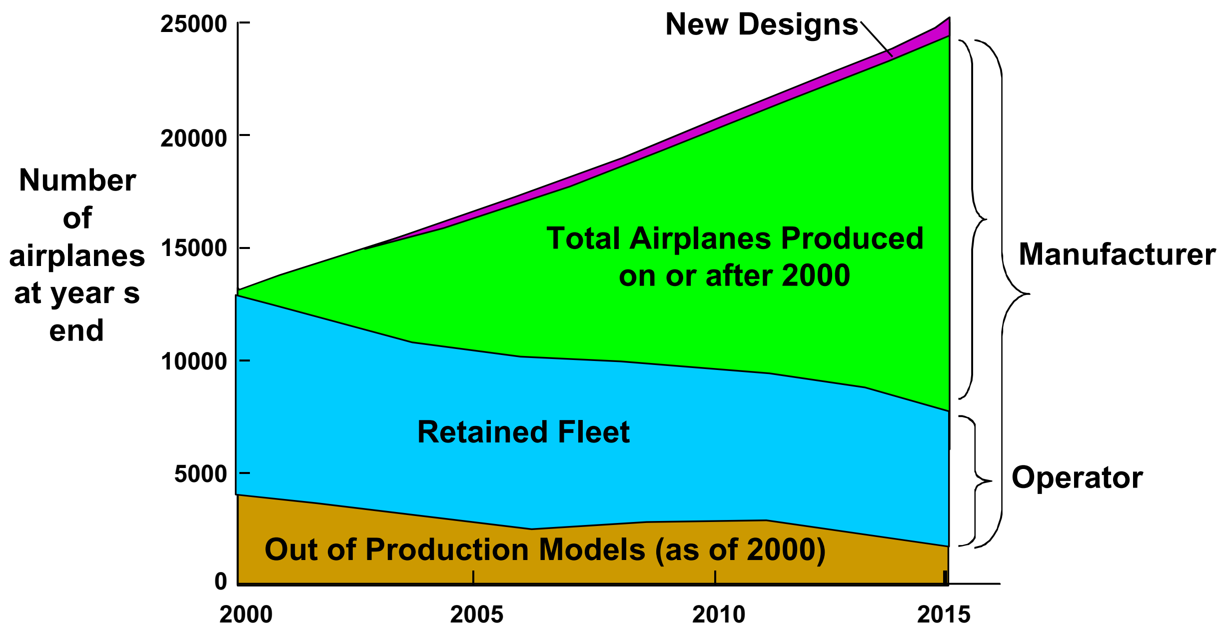
One avenue for addressing these problems is to modify the flight deck interface design to highlight more effectively the changes and events that can occur independently of explicit pilot commands and are triggered instead by sensed environmental conditions and/or designer input to the system. Also, changes that occur a considerable time after the crew has commanded them need to be indicated in a more salient fashion. In addition, pilots need better support for interpreting the indicated automation status in terms of its implications for current and future aircraft behavior. New flight deck interface designs are being developed with these requirements in mind. However, design changes take considerable time to find their way into the existing fleet (see Figure 1-1), and solutions are needed for the existing fleet. Therefore, new approaches to automation training also need to be developed and implemented.

1.2 Areas of Concern for Training and Design

There is considerable evidence that, despite airlines considerable training efforts and investments, pilots are leaving current training programs with incomplete and inaccurate mental models of the automation (e.g., Wiener, 1989; Sarter and Woods, 1992, 1997). Some airlines have explicitly stated that they do not expect their pilots to gain a complete understanding of the automated systems until after their first 6-12 months of operating experience.

A number of researchers (e.g., Feltovich, Spiro, & Coulson, 1991; Sarter & Woods, 1994, in press) have suggested that training could be enhanced in at least two ways. First, information on automation modes and functions in training and reference documents is typically presented as sets of facts isolated from either an operational context or framework that reveals the underlying system design or logic. Pilots are provided with recipes for how to use the automation for simple tasks but they do not necessarily know when and why these procedures are appropriate and how they can be modified or adjusted when unusual circumstances require a deviation from the standard. Several groups (e.g., Mitchell et al., 2000; Mumaw et al., 2000) have been working with U.S. airlines to develop approaches for supplementing training with the why and when information.

Figure 1-1. Impact of new designs on the existing fleet.



A second area for improvement relates to the fact that pilots are not sufficiently supported in learning how to monitor automation-related indications effectively. The accepted wisdom on scanning cockpit indications for years was based on the T pattern of primary indications (airspeed, attitude, altitude, heading). However, with the advent of integrated flight deck displays and highly complex automated systems on glass cockpit aircraft, the pilot needs to monitor a larger, more diverse, and more distributed set of indications. Further, many indications are presented in a different format (e.g., the change from round-dial airspeed and altitude gauges to tape instruments). The primary flight display (PFD) and the navigation (Nav) display integrate most of the primary indications, but the pilot also needs to monitor the mode control panel (MCP) and the flight management computer (FMC), which is accessed through the control data unit (CDU). There are no documented strategies for effectively monitoring this diverse set of indications, and, as a result, pilots often develop their own (not necessarily appropriate) approaches to the task.

Another factor that can make it more difficult for pilots to monitor the automation is the introduction of multi-function interfaces such as the CDU. The CDU contains a large number of pages, each containing a considerable amount of data, but only one page can be viewed at a time on each CDU. Thus, monitoring relies completely on the pilots ability to determine what information is relevant, where it is located, and how the relevant page can be accessed in a timely manner. Yet another complicating factor is the fact that the interfaces that are used to select automation modes are physically separated from the interfaces that indicate the engaged modes. For example, pilots manipulate controls on the MCP to select an autopilot mode but they have to look at the PFD to verify that that mode has been engaged (or is only armed) and will perform the intended action. This is a rather unintuitive procedure, and a violation of principles behind a more promising approach called direct manipulation interfaces (Hutchins, 1997).

In general, monitoring is either data-driven or knowledge-driven. Data-driven monitoring occurs when a pilot re-orients his/her attention towards a stimulus that has attracted the pilot s attention because it is salient or unexpected. Knowledge-driven monitoring is guided either by formal practices or procedures, or by the pilot s mental model of the system. In the first case, pilots monitor indications that are expected to be relevant at certain points in time e.g., monitor engine parameters on take-off. The cue for monitoring may be taken from a procedure or other formal practice. The traditional wisdom about a T-pattern scan falls in this category of a formal practice. In the second case, the pilot has a mental model of the system that allows him/her to generate expectations about system behavior. This mental model aids the pilot in deciding what indications are relevant at various points in time. Knowledge-driven monitoring tends to be based on knowledge that was acquired during training and practice, whereas data-driven monitoring, which is equally important, needs to be supported through effective interface design. Based on the literature on breakdowns in pilot-automation interaction it appears that neither form of monitoring is adequately supported by current glass cockpit interfaces.

1.3 Concerns with Reported Pilot Monitoring Behavior and Strategies

Earlier research has identified a number of problems with pilots monitoring behavior and performance (e.g., Abbott et al., 1996; Eldredge et al., 1992; Sarter & Woods, 1994, in press). For example, it has been shown that

- pilots miss mode transitions induced by the automation
- pilots fail to recognize that a target (e.g., altitude) will not be obtained
- pilots fail to recognize that a programmed restriction was deleted
- pilots are unaware that they have lost the vertical navigation path (VNAV PTH) mode
- pilots are unaware of mismatches between information on a CDU cruise page vs. a descent page

Note that the above problems refer to performance outcomes. Little is known about the reasons underlying these breakdowns in monitoring performance and about the monitoring process in general.

There are few studies that have systematically examined pilot scanning with adequate sample size to draw statistical generalizations across pilots, and there are fewer still that have done so within the context of the glass cockpit. In the former category, Fitts, Jones, and Milton (1950), Carbonell, Ward, and Senders (1968), and Harris and Christhilf (1980) provided earlier examination and analysis, establishing that pilot fixation frequencies and durations provided useful approximations of information extraction and relevance. These studies also established the attitude indicator as the primary center of visual interest, and observed the sensitivity of visual scanning to different flight phases. Harris and Christhilf also distinguished between short dwells (less than around 400 msec) used for confirming the reading of instruments, and longer dwells that were used to extract new information from an instrument. Bellenkes, Wickens, and Kramer (1997) employed similar procedures but used a large sample of both high-time instructor pilots (experts) and low-time students (novices). They observed the systematic changes of flight instrument importance across different phases of flight (i.e., climbing vs level, turning vs straight), and also systematic differences between experts and novices. In particular, they found that experts had a more refined mental model, which would lead them to scan more to instruments that anticipated changes in flight path parameters.

In contrast to the work of Harris and Christhilf and Bellenkes et al., a series of two studies by Wickens and his collegues (Wickens, Xu, Helleberg, Carbonari, & Marsh, 2000; Helleberg and Wickens, 2000; see Wickens, 2000 for a summary) examined scanning during visual flight, when pilots had a two screen projection viewable beyond the high-fidelity cockpit simulator. They examined the influence on general aviation pilot scanning of two aspects of advanced cockpit technology Datalink and a cockpit display of traffic information (CDTI) for free flight. Both studies showed that these technologies represented a substantial sink of visual attention, capturing approximately 20% of scan time for the CDTI and 15% for the Datalink interface. Both showed the instrument panel to occupy around 50-60% of the attention, and the outside world to receive a relatively small percentage of visual attention (around 30%), given the potential importance of visual contact with traffic, in the see and avoid environment of much of general aviation.

The only study of visual scanning done in a glass cockpit appears to be one carried out by Heuttig, Anders, and Tautz (1999), using an Airbus A340 simulator, with six professional line pilots (three crews), on flights to Munich. They observed that around 40% of the time was spent scanning the primary flight display, of which most was distributed equally between instruments with the basic T (attitude, airspeed, altitude and heading). The heading indicator was undersampled, but compensated by heavy interest in the navigation display (20%). Only 10% of visual attention was allocated to the outside world, and, of particular interest to the investigators, only 5% was allocated to the FMAs (note that this value dropped to 2% on flights that did not contain any abnormalities). The investigators observed that the FMAs did not appear to be part of any regular scan path, although they noted a greater-than-expected frequency of visiting two FMAs consecutively. They also reported a bimodal distribution of FMA dwells, akin to the description offered by Harris and Christhilf (1980). That is, there were both some very short FMA dwells (used, presumably, to confirm an expected observation) as well as some longer ones (used to extract new information). Thus, while some data exist, knowledge of information sampling strategies in the glass cockpit airplane is far from complete.

Several questions need to be answered, including:
1) How do pilots monitor cockpit interfaces and indications on automated flight decks?
2) What information do they access? When, in what sequence, and for what purpose do they access the information?
3) How does the interface design support or hinder effective monitoring ?

The present study serves to address the above questions. Pilots were asked to fly a scenario that involved a number of situations and events that are known to challenge pilots monitoring skills (e.g., the automation performs actions that were not commanded and not expected by the pilot, or the automation fails to take an anticipated action see Sarter & Woods, in press). More specifically, Sarter and Woods (1994) identified three ways in which flight deck automation can surprise pilots:

1) the automation fails to show expected behavior e.g., fails to slow to 240 kts below 10,000 ft.
2) the automation takes an action that wasn t expected e.g., losing flight plan information when the departure runway is changed.
3) an automation failure occurs that is not annunciated with salient indications e.g., loss of glideslope signal.

We developed a scenario that allows us to present and evaluate each of these types of situations. Our goal was to identify subjects monitoring skills and strategies and relate those to performance outcomes, their understanding of the automation, and the feedback provided by the system.

2. Methodology

2.1 Subjects

Line pilots were recruited from two U.S. airlines, here referred to as Airline A and Airline B.

Recruiting. We contacted management at each airline and requested permission to recruit their 747-400 pilots. For Airline A, we had an announcement placed in each pilot s company mailbox. For Airline B, we had an email sent to all eligible pilots. Interested pilots then contacted Boeing directly to volunteer for the study. Pilots were not paid for their participation. Each volunteer arranged a time to participate and arranged his own transportation to Boeing.

Demographics. Table 2-1 lists important characteristics of the 20 subjects. All subjects were male. Note that subject numbering is from 4 to 23. Subjects 1, 2, and 3 were used for ensuring that the scenario and data collection ran smoothly; data from these subjects were not analyzed.

S#	Airline	Position	Age	747-400 hrs*	Hrs on prior glass cockpit*
4	A	FO	47	5400	0
5	A	FO	56	2000	0
6	A	Capt	58	1000	1800 (757/767)
7	A	FO	50	2500	0
8	A	FO	47	800	3500 (767,737-300)
9	A	Capt	59	1200	3000 (757/767)
10	A	Capt	56	1750	3500 (767)
11	A	FO	45	3500	0
12	B	Capt	59	100	5000 (757)
13	A	FO	53	4000	0
14	B	Capt	59	700	3000 (757)
15	A	Capt	56	960	2500 (757)
16	A	FO	45	3000	4500 (767, 737-300)
17	B	Capt	58	9000	0
18	B	FO	51	2500	0
19	B	Capt	53	5000	1400 (757)
20	B	Capt	55	3200	1000 (757)
21	A	Capt	59	300	700 (757/767)
22	A	FO	54	3600	0
23	A	FO	46	2000	1500 (737-300)

Table 2-1. Subject demographics

*Subjects estimated their hours in the 747-400 and previous experience in other glass cockpit airplanes. We did not use log book hours. Subjects typically used 600-700 hrs per year to arrive at an estimate.

2.2 Experimental Session Overview

The following sequence was carried out for each subject:
1) review informed consent and demographics forms
2) briefing and review of clearance, charts, and dispatch papers
3) eye-tracker calibration
4) simulator familiarization
5) scenario/data collection

6) break
7) debrief on pilot actions
8) mental model test

The following provides a more detailed description of these events:

Each subject, upon arriving at the Boeing simulator facility, was met at the security gate and escorted to the 747-400 simulator. The subject was then taken to a conference room where he read and completed the informed consent form (this form is reproduced in Appendix A), provided demographic information (age, experience, etc.), and was briefed on the study (see the briefing narrative in Appendix B). At the conclusion of this briefing, each subject was given the following:

- the initial clearance (written out)
- the relevant Jeppeson charts for the route
    - SFO ILS 28L
    - SFO ILS 28R
    - PORTE THREE departure
    - SADDE SIX arrival
    - LAX ILS 25R
    - LAX ILS 25L
- a set of dispatch papers/weight manifest in the form that his airline would use.

Each subject was then given as much time as he wanted to review these documents. When the subject completed this review, he returned to the simulator to be fitted and calibrated with the eye tracker.

The eye tracker calibration was done in the simulator with the subject sitting in the appropriate seat that is, either Captain or FO seat. When the calibration was completed, the Team Pilot (our confederate pilot) joined the subject in the simulator and reviewed the clearance, flight plan, and CDU/MCP entries. The Team Pilot (TP) had set up the flight deck while the subject was being briefed, so this was an opportunity to show the subject what information had been entered. The subject also took time to become familiar with the simulator. The TP then discussed his role and how they would operate together (e.g., managing MCP entries). When the subject was comfortable, there was a final check of the eye tracker calibration, and the scenario was started.

Appendix C contains the procedure that was used to guide the scenario. It shows the timing of events and the ATC script. When the subject landed at LAX and came to a complete stop on the runway, the scenario was stopped and the eye tracker was removed. The subject was then given a 10-minute break.

After the break, R. Mumaw, S. Kimball (observer), and the TP sat with the subject near the simulator and reviewed the scenario. This was primarily an opportunity to make sure that the observer notes were complete and accurate. The subject was asked about his response to each event (how he handled it; why he took that approach; etc.). When this debrief segment was complete, everyone left except for R. Mumaw and the subject. R. Mumaw then asked the set of questions that comprised the mental model test (see Appendix E).

When the mental model test was completed, the study ended. Each subject was given the opportunity to ask questions, but certain information about the mental model test and scenario events was kept from subjects. We were interested in preserving ignorance of subsequent subjects. Each subject was asked not to talk about the scenario to other pilots. Finally, each subject was given a few Boeing trinkets (mug, poster, etc.) before being escorted out of the building.

2.3 Simulator

The study was carried out in a 747-400 fixed-base simulator. Each pilot s front window view covers 45° horizontally and 34° vertically, with a 2° look-down angle. The image monitor has a resolution of 1024 x 768, and uses a wide-angle columnation (WAC) optical technique to simulate focus at infinity. The image generation system is an Evans & Sutherland ESIG 3350.

The flight deck interface was different from an actual aircraft in the following ways:
- some functions that were not pertinent to the scenario (e.g., TCAS, fuel flow) were inaccurate.
- The PFDs were programmable displays that replicated the functionality of standard hardware displays. Primarily, the only difference from the hardware displays was that we were able to overwrite the flight mode annunciation (FMA) area, so that we could alter the annunciated mode.

2.4 Simulator Scenario

In collaboration with one of the participating airlines, we developed a scenario that allows us to evaluate monitoring in situations where the flight deck interface can create impediments to a full assessment of current and future automation behavior. It is important to note that these situations represent routine flight operations that is, we are not relying on uncommon failures or unusual clearances. We describe below what types of interface monitoring issue are raised for each event.

As the scenario began, each subject was given the following clearance (see Appendix C): Airline A/B 400, cleared to LAX airport via PORTE 3 departure, AVENAL transition, direct DERBB, then via Los Angeles SADDE6 arrival, DERBB transition. Maintain flight level 350. Cross PORTE intersection at 9,000, then resume climb. Contact Bay departure 135.1

The 747-400 was set up with a zero fuel weight of 470,000 lbs and a fuel weight of 100,000 lbs, for a gross weight of 570,000 lbs. This is a fairly light weight for a 747-400. We used a cost index of 100. Cruise altitude was FL350. We used a derated climb (climb 2), and flaps at 20 for take-off.
We used daylight operations. Weather at SFO was clear. LAX had a 5-mile visibility. There were minimal winds.

Figure 2-1 shows the scenario flight plan and CDU LEGS pages; Figure 2-2 shows a profile view. This clearance is a fairly standard clearance, with the added altitude constraint at PORTE. During the scenario, the following events were implemented:

<u>Event 1. Runway change</u>
Action: ATC requested a change from 28L to 28R during initial taxiing. This is not an uncommon event.
Operational impact: With this FMC, changing runways requires reselecting the departure (SID), which, in turn, produces several changes in the FMC: the take-off speeds (v-speeds) are deleted and need to be reselected, a route discontinuity is created after the SID (between AVE and DERBB), and the hard restriction for 9000 ft at PORTE is lost (reverting to 9000A) since it wasn t part of the standard SID. The CDU information no longer matches the ATC clearance and the subject must somehow restore this information. Specifically, he must cross PORTE at 9000 and must close the discontinuity.
Monitoring issue: The side effects of this runway change are distributed across the interface the CDU (LEGS page, Departure page) and the Nav display and may be hidden from view. The interface does not fully inform the pilot of these changes; they must be discovered. We believed that subjects would not anticipate these side effects, and thus, the automation takes an action that wasn t expected.

Figure 2-1 Scenario flight plan.

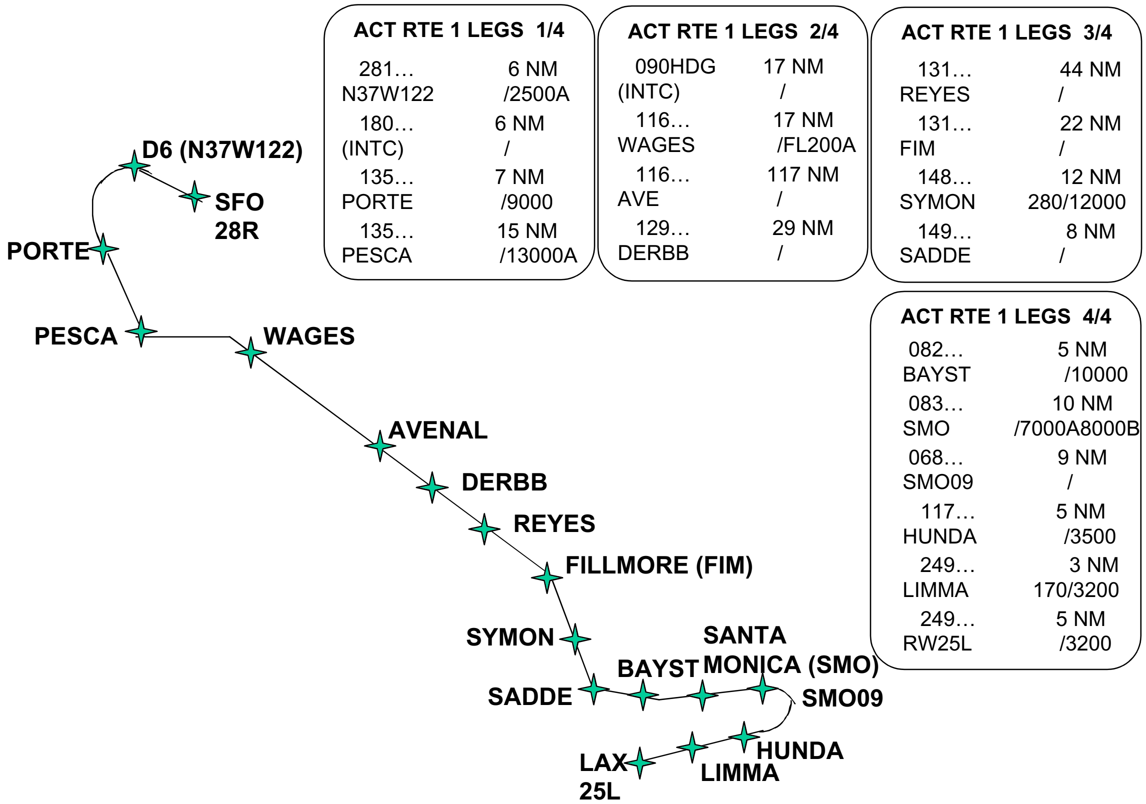
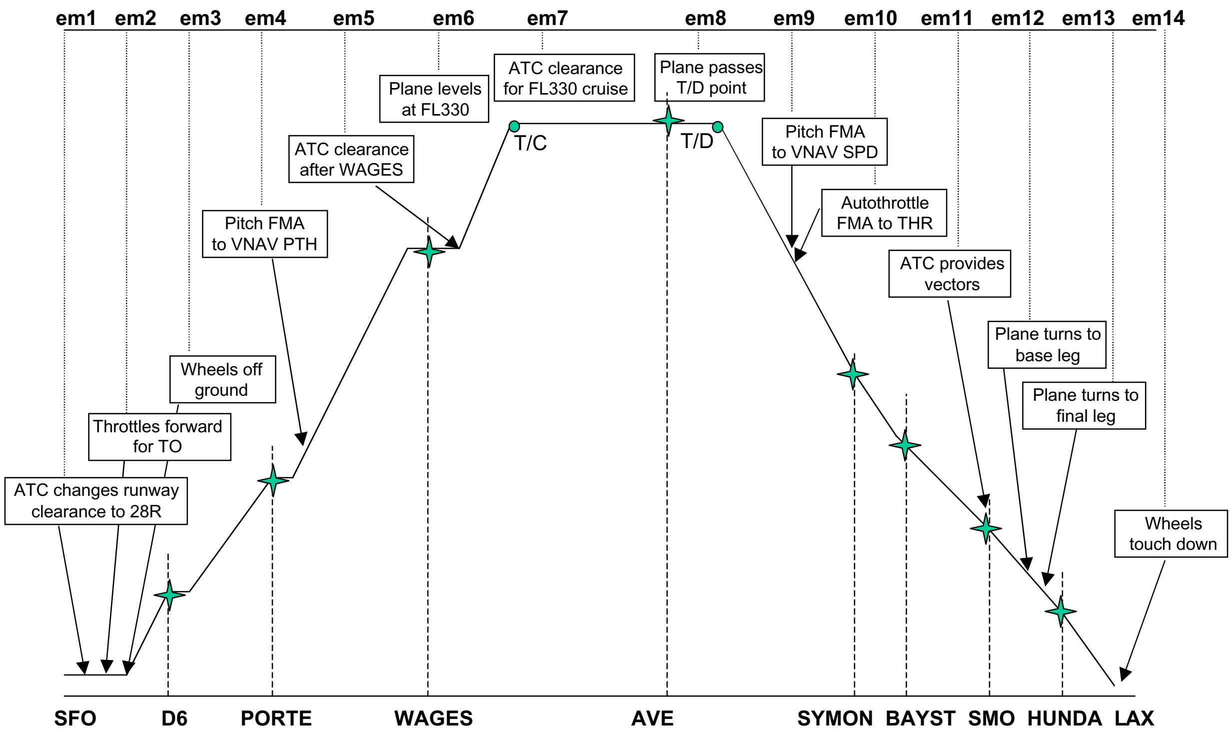


Figure 2-2 Scenario events on schematic profile.



<u>Event 2. Expedite to cross SFO 6 DME at 4000 ft</u>
Action: When the airplane climbed to 1500 ft AGL, ATC requested an expedited climb to cross D6 at 4000 ft. This ATC request is not generally made of heavies (e.g., 747s) but can occur.
Operational impact: This ATC request can compel the pilot to leave VNAV (into a lower-level mode) in a high-workload situation and then recover it later. Also, this clearance leads to setting the MCP altitude down to 4000 ft (from 9000 ft). Thus, all reminders of a hard restriction at 9000 ft are removed (if the 9000 at PORTE in the FMC had changed to 9000A).
Monitoring issues: The first effect is to remove information about the 9000 ft restriction. A second issue is tied to the interface failing to show expected behavior of the automation. The green arc on the Nav display initially makes it appear that capturing 4000 ft will be no problem until the airplane reaches 3000 ft and then noses over to accelerate. At that point the green arc can move out beyond the 6 DME point.

<u>Event 3. Inappropriate pitch mode</u>
Action: After the PORTE waypoint was passed (with the 9000 ft restriction), and the airplane began climbing, we altered the pitch mode annunciation. With VNAV engaged on climb, the airplane pitches to a speed target (not to a path target) to climb to the cruise altitude. In this phase, VNAV SPD is annunciated on the FMA. Just after the plane started climbing, we artificially changed the pitch mode annunciation to VNAV PTH. Obviously, this never occurs during actual operations.
Operational impact: This mode annunciation was changed without the use of a green box that typically shows mode changes. Other aspects of airplane performance did not change.
Monitoring issue: An element of assessing the state of the automation is understanding the implications of an indication. VNAV modes behave differently, and pilots may not have strong expectations about which VNAV mode should appear. No green box was used because we wanted to see the extent to which monitoring is tied to expectations about specific modes.

<u>Event 4. Loss of LNAV/VNAV and visual airplane target</u>
Action: When the airplane reached FL200, ATC requested the subject to level at FL210 and to turn to 150° for traffic. Then, eventually he was vectored to a 90° heading to re-intercept the course to AVENAL (AVE). This clearance and request to leave the flight plan for traffic is not uncommon.
Operational impact: Re-intercepting the course is a little tricky for the following reasons. When he turns to 150° around WAGES, there is a conditional waypoint at the top of the LEGS page for a 90° intercept prior to WAGES. WAGES and then AVE are the next waypoints on the LEGS page. Generally, one shouldn t re-engage LNAV until AVE becomes the active waypoint (moves to the top of the LEGS page and becomes magenta on the Nav display). That is, you shouldn t select a guidance mode (LNAV) until you have ensured you are being guided to the right place. In this case (because of the conditional waypoint), engaging LNAV earlier won t take you back to WAGES, but in other situations, it could. Engaging LNAV early in this case leads to a CDU message that says not on intercept heading. This message indicates that even though the airplane is visually past WAGES, LNAV thinks WAGES is the desired waypoint. WAGES does cycle off eventually (with no pilot action), and he can recapture the course to AVENAL, but in other situations, the airplane might turn around and head for WAGES.
Monitoring issue: Indications are distributed across the CDU, Nav display and PFD. It can be difficult for pilots to assess the state of the automation before they arm LNAV.

<u>Event 5. Revise CRZ altitude</u>
Action: When the airplane was at 31,500 ft, the pilot received an ATC request to level at FL330 (the FMC cruise altitude was FL350). Then, at some later point ATC indicated that FL330 would be the final CRZ altitude. This request is not uncommon.
Operational impact: Going through this sequence with this FMC putting an altitude lower than the FMC cruise altitude on the MCP, then changing the FMC cruise altitude to match results in a VNAV ALT pitch mode. The airplane would typically transition to VNAV PTH on cruise. Other sequences of actions would have put it in VNAV PTH, but in this case, the subject needs to take an additional action

(push the MCP altitude selector) to transition from VNAV ALT to VNAV PTH. The implications are that if he is in VNAV ALT, the airplane will not descend at the T/D point (and even DES NOW will not start the plane down early).
Monitoring issue: The automation fails to take the expected behavior, which is to transition to VNAV PTH. This event allows us to determine whether subjects have an accurate expectation about which VNAV mode is used in cruise and whether that expectation influences monitoring. In addition, if the subject ends up in VNAV ALT on cruise, the airplane will not start down at the T/D point. Again, the automation fails to take the expected behavior (but because subjects have inappropriate expectations).

Event 6. Give speed and altitude restrictions at SYMON
Action: The subject was asked to cross SYMON at 280/12000. This clearance is expected.
Operational impact: This restriction puts the pilot in a situation that requires him to monitor how well VNAV is working on descent. If VNAV transitions to VNAV SPD and he gets above the path, he will not cross SYMON at 12,000 ft. He needs to be in VNAV PTH on descent to guarantee that VNAV will make the restriction (assuming the descent profile is programmed correctly).
Monitoring issue: none

Event 7. Reduction in airspeed to 260 kts
Action: Late in cruise, ATC requested a speed reduction to 260 kts. This request is not uncommon.
Operational impact: This speed reduction can be handled with MCP speed intervention or through reprogramming the FMC. If the pilot puts 260 only on the CRZ page, his airspeed will revert to Econ speed (about 280 kts) after the T/D point. To avoid this, the subject needs to know that CRZ speeds do not propagate to the DES page, and he needs to program the DES page separately. Also, if the pilot only uses speed intervention (and no FMC reprogramming), the FMC does not adjust the top-of-descent (T/D) point, which makes it more difficult to regain the VNAV descent path.
Monitoring issues: One issue is that subjects need to monitor airspeed targets that can come from the MCP or from the FMC. They have to manage the MCP airspeed window and the CRZ and DES page speeds. A second issue is that, if done poorly, the automation fails to show expected behavior and descend at the assigned airspeed.

Event 8. Inappropriate pitch mode
Action: After the airplane was established on the VNAV descent path, it showed VNAV PTH as the pitch mode annunciation. At this point, we artificially changed the annunciation to be VNAV SPD, even though the vertical path indicator on the Nav display showed it was on path. Obviously, this never occurs during actual operations.
Operational impact: This mode annunciation was changed without the use of a green box that typically shows mode changes. Other aspects of airplane performance did not change.
Monitoring issue: An element of assessing the state of the automation is understanding the implications of an indication. VNAV modes behave differently, and pilots may not have strong expectations about which VNAV mode should appear. No green box was used because we wanted to see the extent to which monitoring is tied to expectations about specific modes.

Event 9. Inappropriate autothrottle mode
Action: After the change to VNAV SPD, while the airplane was still actually on the VNAV descent path, we artificially changed the autothrottle annunciation to THR. THR is not a mode one would see in this situation. Obviously, this never occurs during actual operations.
Operational impact: This mode annunciation was changed without the use of a green box that typically shows mode changes. Other aspects of airplane performance did not change.
Monitoring issue: see Events 3 and 8.

Event 10. Loss of glideslope diamond and glideslope
Action: We failed the ground signal for the glideslope. As a result, the glideslope diamond on the PFD never filled in and centered itself. While this event is rare, it does occur.
Operational impact: The subject is required to do a LOC only approach.
Monitoring issue: This failure was introduced because it relies on the disappearance on an indication, which is more difficult to notice than a failure that is associated with a positive indication, such as an aural warning.

2.5 Performance Measurement

Eye Fixations. Eye scan measures were made using an ASL" series 4000 head-mounted eye tracker (Applied Science Laboratory, Waltham, MA). The ASL Model 4000 is a complete eye-tracking system for use in situations where the subject can wear lightweight, head-mounted optics and must have unrestricted head movement (Figures 2-3 and 2-4 provide illustrations). The system is designed to measure a subject's eye line of gaze with respect to the head. This measurement is displayed as a cursor (set of cross hairs) superimposed on the image from the scene camera. When combined with an optional head tracking device and eye/head integration software, the eye tracker can also measure a subject's eye line of gaze with respect to stationary surfaces in the environment. Generated data include time, x and y eye position coordinates, and pupil diameter. The EYENAL analysis package enables us to process the recorded data and create measures such as dwell sequence and dwell duration.

With certain constraints described below, we were able to associate each fixation with a particular *area of interest* (AOI) on the flight deck. These AOIs, described in detail below, involve such regions as a particular instrument (e.g., the airspeed tape), or a more general area (e.g., the CDU or outside scene).

Prior to the experiment, each subject was given a briefing regarding the calibration procedure for the eye tracker. Following this briefing, the eye tracking apparatus was placed on the subject's head, and the calibration procedure was conducted.

During data collection, we placed 14 marks (called eye marks) in the eye tracking data file for the following events:
eye mark 1 — ATC gives clearance to change runway
eye mark 2 — throttles forward for TO
eye mark 3 — wheels off the ground
eye mark 4 — pitch mode is forced to VNAV PTH
eye mark 5 — ATC gives clearance to resume flight plan after WAGES
eye mark 6 — plane levels at FL330
eye mark 7 — ATC gives final clearance for CRZ altitude to be FL330 and it is executed in the FMC
eye mark 8 — plane passes top-of-descent point; enters descent phase
eye mark 9 — pitch mode is forced to VNAV SPD
eye mark 10 — autothrottle mode is forced to THR
eye mark 11 — ATC gives clearance for vectors at SMO
eye mark 12 — plane turns onto base leg
eye mark 13 — plane turns onto final leg
eye mark 14 — wheels touch down

Figure 2-2 shows these eye marks (em) on the profile view of the flight plan.

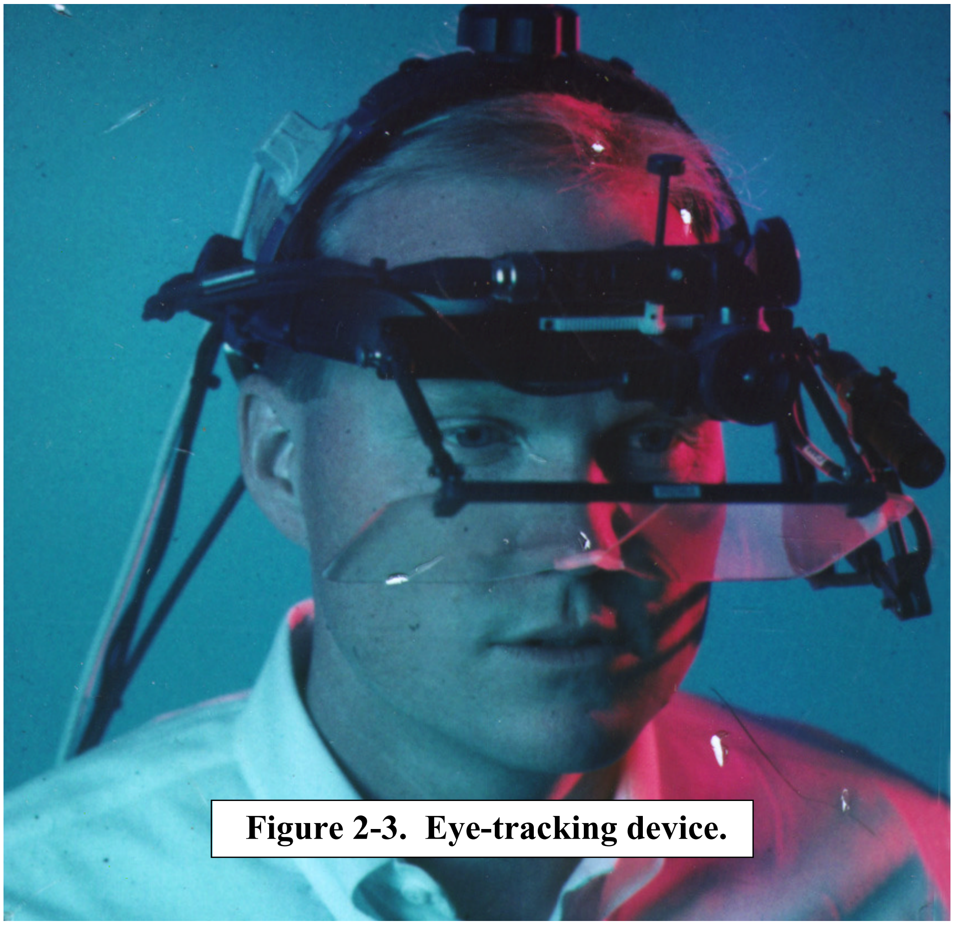


Figure 2-3. Eye-tracking device.

Figure 2-4 here

Figure 2-4. Subject wearing eye-tracking device in simulator.

Simulator Variables. Appendix D lists the simulator variables that were recorded during the scenario. Each variable was recorded at a rate of once per second (1 Hz). However, note that because we are asking pilots to use the automation (VNAV/LNAV), we were not focused on many aspects of flight performance.

Pilot Behavior. Data sheets were developed to record subject behavior during the scenario. An observer sat behind the subject pilot and recorded his responses to the scenario events. Specifically, we were interested in:
- how he handled the runway change and restoring lost FMC entries
- his method for expediting climb to 4000 ft at D6
- whether the restriction for 9000 ft at PORTE was met, and how it was met
- whether he noticed the pitch mode change to VNAV PTH on climb
- how he met the ATC request to hold at FL210 for traffic and then return to the programmed flight path
- how he handled the change in cruise altitude (to FL330)
- how he handled the ATC request for airspeed reduction (to 260 kts)
- whether the plane descended at the top of descent point
- whether he noticed the pitch mode change to VNAV SPD on descent
- whether he noticed the autothrottle mode change to THR on descent
- how he handled the loss of glideslope signal

These data were reviewed and edited during a debriefing session with the subject and TP.

In addition to the observer s data, two videotapes were made for each subject. One video captured the input from the eye-tracker scene camera. This camera was mounted on the eye-tracker head gear, moving as the subject s head moved, and shows the general area that the subject is viewing. Crosshairs superimposed on this scene show the eye-fixation point. The second video integrates four separate cameras: the scene camera, a camera on the PNF s CDU, a camera on the MCP, and a direct feed of the PF s PFD. Audio of ATC, the PF, and PNF was captured through lapel mics on the PF and PNF. ATC transmissions were channeled to an overhead speaker in the simulator cab, and these were then picked up through the pilots lapel mics. Debriefing sessions with each subject were recorded as audio-only on the videotapes.

Further behavioral analysis was done from the videotapes.

Mental Model. After the subject debriefing, each subject was asked a series of questions about how the automation functions. The questions and answers are in Appendix E. This session was carried out as a discussion to explore each subject s understanding, and a few topics were not covered with some subjects. Each subjects answers were scored to identify both correct and incorrect assertions.

3. Results

3.1 Overview

The results first address pilot performance issues, then we discuss general monitoring results, and finally the mental model results. The last section discusses larger trends in the data.

3.2 Pilot Performance

In this section, we characterize the performance of the subjects in carrying out the simulator scenario. In some cases, eye-tracking and mental model data are used to further describe pilot performance. All data tables are in Appendix F.

Operational differences between airlines A and B.

It is worth noting the differences we encountered between the two sets of airline pilots. First, we found that neither airline uses derated take-offs anymore. The take-off out of SFO was set up as a derated take-off. These subjects were familiar with derated take-offs but both airlines had recently removed them from operations.

Also, we found differences in the following operational practices:

Airline A — when flying with autopilot, the pilot flying (PF) indicates he wants an altitude set on the MCP, and the pilot not flying (PNF) sets the altitude. The PF enters airspeed, heading, vertical speed, and engages modes.
Airline B — when flying with autopilot, the PF enters altitude, airspeed, heading, vertical speed, and engages modes.

Airline A — they arm the autothrottle only when ready for take-off.
Airline B — they arm the autothrottle when they are taxiing for take-off.

Airline A — they ask for altitude call-outs 1000 ft prior to MCP altitude.
Airline B — they ask for altitude call-outs 2000 and 1000 ft prior to MCP altitude.

Loss of FMC altitude restriction with runway change.

Soon after beginning to taxi, each subject received a clearance for Runway 28R instead of 28L (original clearance). The Team Pilot (TP) entered and executed the new runway clearance. The TP then re-inserted the take-off speeds. However, the reselection of the standard instrument departure (SID) caused the FMC altitude restriction at PORTE to change from 9000 to 9000A. Table 3-1 (in Appendix F) shows how each subject responded to this change and how well they met the altitude restriction.

Simple Table Summary:
- 9 of 20 Ss noticed that FMC altitude value was different from the clearance; 11 of 20 were unaware of the reversion.
- of the 9 who noticed, 5 noticed after take-off; 4 before take-off.
- 5 of the 9 who noticed the reversion changed the PORTE LEGS altitude constraint back to 9000.
- 3 of 20 failed to cross PORTE at 9000 ft; BUT these failures to meet the restriction at PORTE were not due to missing the change to 9000A.
- S17, after crossing D6 at 4000 ft, was going to set the MCP altitude window to FL350, but asked the TP what the clearance was. The TP correctly told him the clearance was to cross PORTE at 9000.

Discussion:
The reversion of 9000 to 9000A that was created by the runway change was available on the LEGS page and on the Nav display. The LEGS page was typically displayed during taxiing and early climb, and the Nav display was continuously present. Only one subject (S10) anticipated that the PORTE restriction would be lost with the change of runways. Eleven of the remaining 19 subjects failed to notice this side effect. Thus, despite having indications available, more than half of the subjects failed to notice the automatic reset of the altitude constraint at PORTE.

<u>Creation of flight plan discontinuity with runway change.</u>

The runway change also created a route discontinuity between the termination of the SID (AVENAL or AVE) and the next waypoint on the route (DERBB). Table 3-2 shows how each subject responded to the discontinuity.

Simple Table Summary:
- all 19 Ss noticed and closed the discontinuity; data is missing for 1 S (S21).
- 3 of 19 Ss anticipated the discontinuity and closed it before take-off.
- of the 16 Ss who closed it after take-off; 13 closed it between WAGES and AVENAL.
    - 6 of these 13 seemed to notice from the Nav display
    - 5 from the LEGS page
    - 2 from acting on the descent clearance
- the remaining 3 of 16 noticed the discontinuity between PORTE and WAGES as they were climbing before being vectored around WAGES. Note that all 3 of these seemed to be cued by the LEGS page.

Discussion:
All subjects either anticipated this change or noticed it before it affected their route of flight. The LEGS page shows a very prominent message that there is a route discontinuity and the Nav display removes the flight path between the two waypoints. Thus, when the cues are strong enough, pilots will notice them and take appropriate action.

For the three subjects who noticed the discontinuity between PORTE and WAGES, note that they all seemed to be cued by the LEGS page. Interestingly, the discontinuity after AVE appears on the bottom of the LEGS page (LEGS pg 1) just after the plane crosses PORTE, and this is the first opportunity to see it.

For the subjects who noticed the discontinuity between WAGES and AVE, we believe there is a link to receiving the clearance to resume the flight plan after being vectored around WAGES (eye mark 5). Table 3-2 shows the elapsed time from that eye mark. During the planning to resume the flight plan to AVE, subjects are likely to set the range on the Nav display to show AVE and to focus on AVE on the LEGS page. This will reveal the discontinuity after AVE. The two subjects who noticed the discontinuity later may have been aided by receiving the ATC clearance to DERBB from AVE, which occurred early in cruise.

<u>Expedite climb to 4000 at the SFO 6 DME.</u>

After take-off (with derated climb, CLB 2) and at about 1500 ft (RA), ATC requested the subject to expedite to cross the SFO 6 DME (D6) at 4000. The TP dialed down the MCP altitude from 9000 to 4000, and the subject was required to meet this clearance. Table 3-3 shows how each subject responded to this request.

Simple Table Summary:
- 7 of 20 took no initial action; 5 of these 7 eventually used speed intervention.
- 2 of 20 used FLCH initially.
- 1 of 20 used V/S; he later used speed intervention and a thrust increase.
- 2 of 20 used speed intervention.
- 8 of 20 increased thrust to full CLB initially; 2 of these 8 eventually used speed intervention.

Discussion:
It is interesting that such a variety of techniques was used to meet this clearance. This may, to some extent, reflect the fact that this type of clearance, although not unheard of, is not common for the 747.

No subject who went to full climb power had any trouble achieving 4000 ft. Among the various approaches, it is interesting that one subject took no action in response to the clearance and made 4000 ft although the other subject who took no action did not make 4000 ft. Note that many of these subjects, in addition to the techniques listed in the table, also delayed flap retraction to aid climbing. Although we didn t capture the details of flap retraction timing, it was a help to several subjects in climbing better.

Also interesting are the effects of the automatic transition at 3000 ft where the airplane pitches down to emphasize acceleration over climbing. Eight subjects used speed intervention in response to the new situation presented after 3000 ft. All three of the subjects who failed to meet the clearance took no action after 3000 ft.

Two of the three subjects who fell well short of 4000 ft used speed intervention with speeds that were apparently too high. In fact, one selected a speed that is inappropriate for flight below 10,000 ft. (pilots are not to exceed 250 kts below 10,000 ft without a clearance). This subject commented on how he had selected a bad technique. Some of these subjects struggled with identifying an appropriate airspeed for best rate of climb.

<u>Leaving 4000 ft and meeting the clearance at PORTE.</u>

To make the altitude restriction at PORTE, subjects had to start climbing from 4000 ft just after D6. Table 3-4 shows how subjects managed this climb.

Simple Table Summary:
- 7 of 20 Ss started climbing late
- all 3 Ss who failed to meet the 9000 ft restriction at PORTE started up late from D6.
- the 2 Ss who remembered the clearance on their own started up sooner, and generally the sooner they started climbing, the more likely they were to make 9000.

Discussion:
The reason that three subjects failed to cross PORTE at 9000 was either confusion about the clearance or a failure to act on the clearance promptly. Although the clearance was clearly stated, some subjects misinterpreted it. The TP answered correctly if asked about the clearance, but did not offer help when the subject failed to act on the clearance.

<u>Detecting the artificial pitch mode while climbing in VNAV.</u>

After PORTE, we changed the pitch mode from VNAV SPD to VNAV PTH. Table 3-5 shows whether subjects detected this change and how frequently they fixated the pitch mode.

Simple Table Summary:
- For 2 of the 20 Ss (S9 and S16), the actual pitch mode was restored very soon after we introduced the artificial pitch mode. No analysis is done for these subjects.
- None of the 18 subjects indicated that he noticed that the pitch mode was unusual.
- 12 of the 18 subjects did fixate the pitch mode during this period, some of them made multiple fixations.
- 7 of those 12 subjects also made a later statement that indicated they had correct expectations about the VNAV mode during climb.

Discussion:
All subjects exposed to the artificial pitch mode failed to notice it. This outcome suggests that either
- subjects knew what pitch mode should be present in climb, but didn t check pitch mode
- subjects didn t know what pitch mode should be present
- subjects noticed but chose not to mention the inappropriate mode

We found that at least seven subjects seemed to understand that VNAV SPD should be annunciated during climb (mental model results) and also fixated the pitch mode while it was artificially showing VNAV PTH. Why didn t these subjects notice the inappropriate mode? One possibility is that these subjects have inert knowledge about VNAV modes. That is, they can offer the correct VNAV mode when asked in an interview, but cannot apply that knowledge in line operations. The other logical possibility is that these subjects DID notice but chose not to mention it. Reasons for not mentioning it are
- they were reluctant to mention perceived simulator malfunctions. However, in our instructions, subjects were asked to identify any incorrect or inappropriate indications they saw, and subjects did generally identify spurious EICAS messages when they appeared.
- the airplane was performing correctly.
- they lacked confidence in their belief about the VNAV mode in climb.

<u>Resuming the flight plan after WAGES.</u>

Each subject was held at FL210 and vectored around WAGES for traffic. When they were cleared to resume the initial flight plan, they could take several approaches for meeting this clearance. Table 3-6 shows how subjects managed this clearance.

Simple Table Summary:
- 1 S (S13) had problems with his eye tracking head gear and was being calibrated, which interfered with arming LNAV.
- 8 Ss set up the intercept to AVE; moving AVE to the top of the LEGS page.
- 10 of 13 Ss who set heading and altitude and armed LNAV ended up with an earlier active waypoint.

Discussion:
Pilots shouldn t select a guidance mode (LNAV) until they have ensured they are being guided to the correct waypoint. Ten subjects violated this rule. The active waypoint was behind them when they armed LNAV. More detailed eye fixation analysis would be required to understand this behavior more fully.

<u>Changing the cruise altitude to FL330; VNAV mode in cruise.</u>

Each subject was asked to level at FL330 before reaching cruise altitude. Eventually, he was told that FL330 would be his new cruise altitude. Table 3-7 shows how subjects managed this clearance.

Simple Table Summary:
- 7 of 20 Ss put 330 in the CDU before they got ATC clearance; 12 of 20 Ss put 330 in the CDU after they got ATC clearance; 1 S never entered a new cruise altitude into the FMC.
- 18 of 19 Ss who entered a new cruise altitude did so by going to the VNAV page in the CDU, which was the CLB page. 1 S tried to enter the new altitude on the PERF INIT page, but found this was an invalid entry, so he put it on the CLB page.
- 6 of 20 Ss noticed that the pitch mode that resulted (VNAV ALT) was not what they wanted. 2 Ss used the MCP altitude knob to transition to VNAV PTH; 2 Ss cycled out and back in to VNAV to transition to VNAV PTH (note that 1 S got this technique from the TP); and 2 Ss never followed up to change the pitch mode.
- 16 of 20 Ss, during the mental model interview, indicated generally that they knew the VNAV mode during cruise should be VNAV PTH. 6 of 20 stated that it is important to ensure VNAV PTH is engaged.

Discussion:
It is not unusual to get a clearance for a new cruise altitude. Further, cruise altitude is a major determinant of the top of descent (T/D) point. So, these subjects seem to understand the importance of having a correct cruise altitude in the FMC and establishing it early in cruise.

However, few subjects seem to understand the importance of transitioning to VNAV PTH on cruise. While most subjects fixated the pitch mode during cruise, only six subjects mentioned that VNAV ALT wasn t what they expected, and only four subjects elected to change the mode to VNAV PTH. It seems likely that those subjects who did not change VNAV ALT to VNAV PTH do not understand the implications of having a VNAV mode other than PTH. This conclusion is strengthened in the next table.

Concerning the mental model findings shown in the Table, the second, stronger statement is more closely tied to behavior than is the first statement. In the second statement subjects offered the belief that it is important to see VNAV PTH on cruise because it ensures starting down at the T/D point. Four of the six subjects who said this also made sure they transitioned to VNAV PTH. Also, as the next table shows, one of the two who didn t transition to VNAV PTH started down early and had no trouble at T/D.

<u>Descending at the FMC T/D point.</u>

The FMC generates a top-of-descent (T/D) point, and ideally, the pilot begins descent at that point. Table 3-8 shows how well each subject initiated descent.

Simple Table Summary:
- 11 of 20 Ss started down late; 8 of these 11 used the MCP altitude knob; 3 of these 11 engaged a new pitch mode (V/S or FLCH).
- 2 of 20 Ss tried to start down early using DES NOW but this feature does not work while in VNAV ALT mode.
- 5 of 20 Ss elected to start down early using the MCP altitude knob.
- 4 of 20 Ss started down with no action (note that one of these Ss should NOT have started down, but did so because of a simulator glitch).
- 2 of the 20 Ss began descent with the incorrect airspeed.

Discussion:
This table further reveals the misconceptions of subjects who were in VNAV ALT on cruise. Eleven of the 16 subjects in VNAV ALT failed to start down at the T/D point. Moreover, two of these eleven tried to use the FMC DES NOW feature to start down, not understanding that this function would not work.

We conducted an additional analysis to look at monitoring for those who descended early or at T/D (descenders) versus those who flew through the T/D point (late). Because data were missing for three of the 20 subjects, there were only seven subjects in the descenders group and ten in the late group.

For each group, we looked at how scanning changed from the four minutes prior to T/D to the four minutes just after T/D. Two types of analyses were conducted. First, we looked for differences between the two groups in their scanning during the four minutes after T/D (after analysis). Second, we looked for differences between the two groups in the change in scanning behavior from before to after (difference analysis).

For the after analysis, we found that the only relatively stable difference was related to the interest in the MCP. Here we found that the late subjects fixated a lot more frequently (6.5/minute versus 1/minute; $t15=1.90$, $p<.10$), and for more total time ($t15 = 1.55$, $p=.15$) than did descenders. There is also a hint that the late descending pilots spent less total time looking at the FMA-roll than did the descenders ($t15 = 1.55$, $p=.15$), although this trend is less meaningful and less interpretable than is the difference in MCP scanning.

For the difference analysis, we did statistical comparisons within each AOI, of the before-after differences. That is, which AOIs showed a marked change in scanning behavior as the T/D point was passed. We then found which of these showed a differential change between the two groups, and identified the following two trends:

- MCP: for the descenders, the number of fixations dropped from 9.9 to 2.0, ($t6 = 2.175$, $p<.05$), and the total time of fixation (within the four-minute window) dropped from 8.9 sec to 1.8 sec ($t6 = 2.76$, $p<.02$). However for the late subjects, both of these parameters showed modest increases (not statistically significant).
- Out window: for the descenders the mean duration of each out-the-window scan dropped from 1.8 seconds down to 0.9 seconds (i.e., mean duration cut in half. $T6=3.58$, $p<.02$), whereas for the late subjects, out-window dwell duration maintained a fairly constant value of around 1.00 seconds.

Thus, in summary, the primary difference appears to be in the MCP, where the descenders lost their interest after T/D, but the late subjects sustained or slightly increased that interest. The lost interest in the MCP by the descenders was coupled with shorter fixations on the outside world. The interest in the MCP by the late group is likely tied to wondering why they are not descending. They are likely to be checking the MCP altitude setting (that it is at 12,000) and the mode (VNAV). A secondary difference between the two groups during cruise appeared to relate to the primary flight instruments of the ADI and heading indicator, for which the descenders showed slightly greater interest, and to the CDU, where the descenders showed slightly less interest (3.7% vs 6.8%, $t15=2.0$, $p=.07$).

Managing airspeed restrictions

During cruise, the subject is given an airspeed restriction to maintain 260 kts. Sometime after T/D, ATC gives the clearance to resume normal speed. Table 3-9 shows how each subject managed these airspeed restrictions.

Simple Table Summary:
- all 20 Ss initially used speed intervention to set 260.
- 4 of 20 put 260 on the CRZ page and closed the speed window; 1 S put 260 on the CRZ page but didn t close the speed window.
- 7 of 20 put 260 on the DES page right away; 2 put 260 on the DES page after the T/D point; 5 Ss had 260 on the DES page but didn t close the speed window.

- 17 of 20 Ss knew that putting an airspeed on the CRZ page does not propagate to the DES page.

Discussion:
In the case where one is given a clearance for a reduced airspeed on cruise, it is wise to enter that airspeed into the DES page since the descent airspeed is a determinant of the geometry of the descent path. It establishes a shallower path and correspondingly closer T/D point. Pilots want to avoid descending late and getting above the VNAV descent path. Some subjects started down early for this reason.

Also, the use of MCP speed intervention has different effects during cruise and descent. In cruise, you can open the MCP airspeed window without losing VNAV PTH. In descent, opening the speed window transitions you from VNAV PTH to VNAV SPD. Thus, having the speed window open in cruise even though you have entered 260 into the CRZ page isn t a concern, but keeping the MCP airspeed window open on descent even though you have entered 260 into the DES page (as five subjects did) is poor use of the automation.

Note that subjects 20 and 21 indicated in the mental model interview that they knew the airspeed from the CRZ page does not propagate to the DES page. However, they both put 260 on the CRZ page only initially, and transitioned to DES with the reversion to 282 kts (see Table 3-8).

Detecting the artificial pitch mode while descending in VNAV

After the subject obtains the VNAV descent path, the pitch mode is changed from VNAV PTH to VNAV SPD. Table 3-10 shows whether subjects detected this change and how frequently they fixated the pitch mode.

Simple Table Summary:
- 1 of the 20 Ss (S14) never reached VNAV PTH for a stable period on descent, and we were unable to introduce the artificial pitch mode. No analysis was done for this S.
- 1 of the 19 Ss indicated that he noticed that the pitch mode was unusual.
- 10 of the 19 Ss did fixate the pitch mode during this period, most of them made multiple fixations.
- 7 of those 10 Ss also made a later statement that indicated they had correct expectations about the VNAV mode during descent.

Discussion:
Similar to the results in Table 3-5, this table shows that subjects (with one exception) failed to notice the artificial pitch mode. As with that table, we also find here that many subjects did fixate the pitch mode during that time, and that many subjects volunteered statements reflecting correct expectations of VNAV in descent.

S13 alone noticed the artificial pitch mode. This subject is set apart from the other pilots in two ways. First, he showed a larger number of fixations to the pitch mode FMA (see column 7 of table 3-5). Second, and perhaps more importantly, this subject also showed a very long dwell duration (1.2 seconds) on the first fixation to the pitch mode FMA following the unannounced transition. No other pilot had a dwell on this FMA longer than 600 msec, suggesting that all other pilots may have been engaged in quick fixations to confirm expectations. The 1.2-second dwell of S13 appears to be a qualitatively different sort of activity, involved in extracting new information.

Detecting the artificial autothrottle mode while descending

After the subject obtains the VNAV descent path, the autothrottle mode is changed to THR. Table 3-11 shows whether subjects detected this change and how frequently they fixated the pitch mode.

Simple Table Summary:
- 1 of the 20 Ss (S14) never reached VNAV PTH for a stable period on descent, and we were unable to introduce the artificial autothrottle mode. No analysis was done for this S.
- No S indicated that he noticed that the autothrottle mode was unusual.
- 10 of the 19 Ss did fixate the autothrottle mode during this period; many of them made multiple fixations.
- None of these subjects volunteered any strong expectations about the correct autothrottle mode in descent.

Discussion:
In this case (unlike in Tables 3-5 and 3-10), there is no conflict between subjects professed expectations about the autothrottle mode and what they saw. In general, subjects had few expectations about autothrottle modes, although Table 3-29 shows some areas in which subjects have correct expectations.

Meeting altitude constraints on descent

There are a number of altitude constraints to meet on descent. Table 3-12 shows how well subjects met these constraints.

Simple Table Summary:
- 1 S was high at SYMON. We re not sure why that happened.
- 3 Ss were high at SMO. In each case, the S started down late from BAYST.

Discussion:
Each of the three subjects who failed to get below 8,000 ft at SMO had dialed down the altitude for that clearance, but each one forgot to push the MCP altitude knob to start down. Or perhaps, they didn t understand that they had to push the altitude knob.

Noticing glideslope failure

The glideslope indication was missing. Table 3-13 shows how each subject managed the approach with automation.

Simple Table Summary:
- 8 Ss mentioned that the glideslope was missing (with no help from the TP) prior to intercepting the final leg.
- 6 Ss mentioned the glideslope after intercepting the final leg (with no help from the TP).
- 2 Ss armed the LOC only mode initially.
- 12 of 20 Ss decided to maintain APP mode even though there was no glideslope. These responses are coded as incorrect.

Discussion:
We were conservative in implementing this event. We decided to ensure that subjects got set up for the LOC only approach in time to prevent the need for a go-around. We didn t want the scenario to require a go-around and increase the scenario length. So, we told subjects about the glideslope if they didn t notice the failure.

The reason that subjects should arm LOC instead of APP after being told the status of the glideslope signal is that if they have flight directors (FDs) on, they are getting pitch mode guidance with a bad glideslope signal.

Shutting off automation on approach

Each subject was allowed to fly the approach and landing as he saw fit, meaning it was acceptable to shut off autoflight. Table 3-14 shows how each subject managed the automation on approach.

Simple Table Summary:
- 2 Ss did not shut off the autothrottle prior to touchdown.

Discussion:
Subjects typically waited until they could see Runway 25L before they shut off the autopilot and flew flight directors. Because of the reduced visibility at LAX (5 miles), some pilots had trouble picking up the runway until quite late.

MCP Alt Knob Use: Case #1 — Climbing and setting altitude lower

The MCP altitude knob is used to set new altitudes and initiate climbs and descents (along with other functions). Table 3-15 shows cases where the subject is climbing and ATC requests a hold at an altitude lower than what the subject was expecting:
- he is expecting 9000 and is asked to level at 4000
- he is expecting FL350 and is asked to level at FL210
- he is expecting FL350 and asked to level at FL330

In each case, there is no need to push the altitude knob after setting the new altitude.

Simple Table Summary:
- 6 of 20 Ss were correct in all 3 cases (avoided pushing the MCP alt knob in all 3 cases).
- only 1 of 20 got all 3 wrong.
- 5 of 20 got it wrong twice.
- Note that 2 Ss pushed the alt knob a second time after setting FL330; their comments suggest they were trying to change the CRZ altitude in the FMC with the second push.
- only 1 S got the 4000 ft case wrong.
- 8 of 20 got the FL210 case wrong. This was probably the most urgent/complex clearance of the three.
- 12 of 20 got the FL330 case wrong.

Discussion:
While the MCP altitude knob is a complex device, this case seems to be fairly straightforward: the airplane will not fly through the MCP altitude (except in approach). The mental model interviews revealed that subjects generally had some confusion about this knob.

It is not clear why the level off at FL330 produced the most errors.

MCP Alt Knob Use: Case #2 — Level and setting altitude higher to climb

Table 3-16 shows the cases where the subject is level and there is already a clearance to continue climbing, or ATC requests a higher altitude:
- cross D6 at 4000 and continue climbing
- cross PORTE at 9000 and continue climbing
- after holding level at FL210 continue on up to FL350

In each case, if the subject is level and in VNAV ALT or ALT, he needs to push to climb. If he is in VNAV PTH (at 9000), he doesn t need to push the MCP alt knob.

Simple Table Summary:
- 10 of 20 were correct in all 3 cases.
- 10 of 20 got 1 wrong; no one got more than 1 wrong.

Discussion:
Subjects generally made fewer errors for these cases than for the others. If they do not push the altitude knob, they won t leave the altitude.

<u>MCP Alt Knob Use: Case #3 — Level and setting altitude lower to descend</u>

Table 3-17 shows cases where the subject is level or descending to an altitude and there is a clearance to continue descending. Clearances that were used in this scenario were:
- descend from cruise to 12,000 to cross SYMON
- then cleared to 10,000 to cross BAYST
- then cleared to cross SMO between 7,000 and 8,000
- then cleared to 3,500
- then cleared to 2,200
- and finally cleared to 700(660)

Note that not all subjects used all clearances. If the subject is level and in VNAV ALT or ALT, he needs to push the alt knob to descend. If he is in VNAV PTH or not yet level, he doesn t need to push the MCP alt knob to continue down.

Simple Table Summary:
- 4 of 20 were correct for all 6 cases.
- 6 got 1 wrong; 6 got 2 wrong; 4 got 3 wrong; no one got more than 3 wrong.

Discussion:
Many subjects began using MCP modes (FLCH, V/S) below 10,000 ft and so there is less information here about how well subjects were using the MCP altitude knob. These can be complex cases.

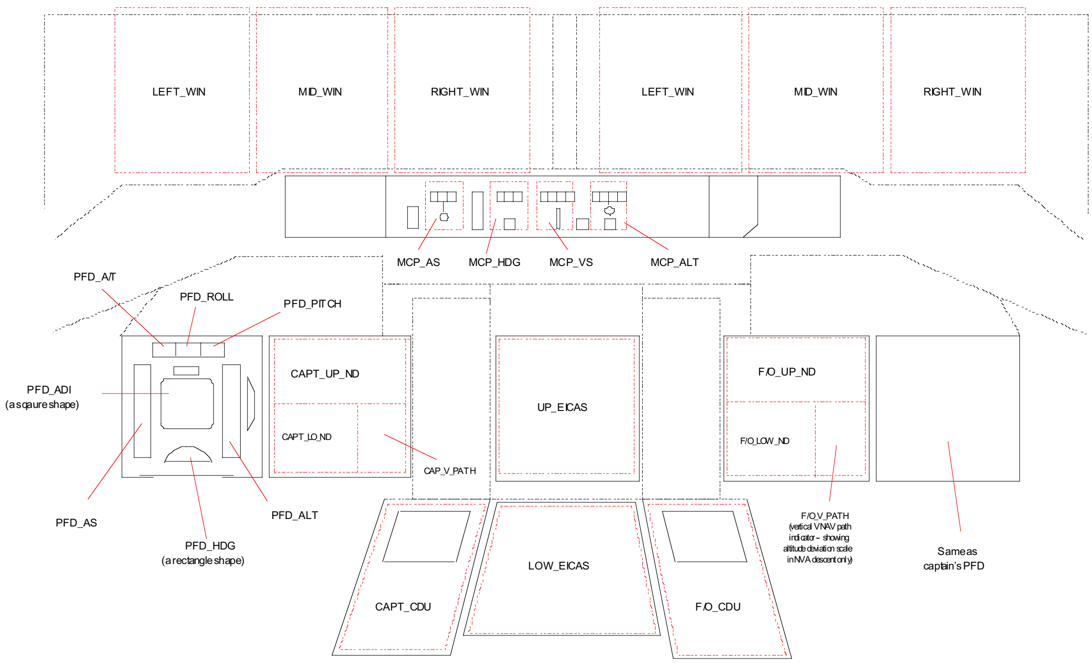
In general, Tables 3-15,16, and 17 provide a detailed record of how poorly pilots understand the altitude knob. Pilots seem to overuse it (there were 48 unneeded pushes but only 14 omissions), and may believe it is better to push when you are unsure. However, under certain circumstances, an extra push can eliminate an altitude constraint from the FMC. Therefore, it is important for pilots to understand the implications for each push. The flight deck interface does a poor job at revealing the consequences of each push of the altitude knob.

3.3 Pilot Monitoring

For each subject, we recorded eye fixation location and the dwell time of each fixation. The primary analysis of these data was to determine into which area of interest (AOI) each fixation fell. Figure 3-1 shows the set of AOIs for the analysis of a Captain s data (the mirrored set is used to analyze a First Officer s data).

A single data set could be analyzed at two levels, depending on the precision of the data. At the highest level, data could be analyzed into the following set of seven larger AOIs: PFD, Nav display, out the window, CDU, upper EICAS display, lower EICAS display, and the MCP. We were able to reliably analyze the data from 17 subjects with this set of AOIs. Thus, eye fixation data were not available for three subjects.

Figure 3-1. Areas of interest (AOIs) - Captain s side



At a more fine-grained level, data could be analyzed into the following smaller set of seven AOIs within the PFD: attitude, altitude, airspeed, heading, and each of the three FMAs. We were able to reliably analyze the data from 14 subjects with this set of AOIs. Thus, eye fixation data at this level were not available for an additional three subjects.

<u>AOI sampling.</u>

The first analysis was to determine which AOIs were being fixated for each phase of flight. Phases of flight were defined as follows (see also figure 2-2):

1. take-off	throttles forward (em2) — wheels off (em3)
2. climb	wheels off (em3) — level at FL330 (em6)
3. cruise	level at FL330 (em6) — T/D point (em8)
4. VNAV descent	T/D point (em8) — clearance for vectors at SMO (em11)
5. vector descent to landing	vectors at SMO (em11) — wheels touchdown (em14)

Our first analysis looks at percent of dwell time in each of the AOIs. Figure 3-2 shows the percent dwell time for each of the seven larger AOIs by phase of flight (n = 17 subjects). Figure 3-3 shows the percent dwell time for each of five PFD AOIs by phase of flight (n = 14 subjects) (these values represent the total of dwells within the PFD AOI from Figure 3-2. Each figure is accompanied by a data table with the plotted values. Note that in these two figures the AOI percentages do not add up to 100% because the data do not include those times in which pilots were fixated on regions other than the designated AOIs e.g., looking down at a clipboard or across the cockpit at the other pilot.

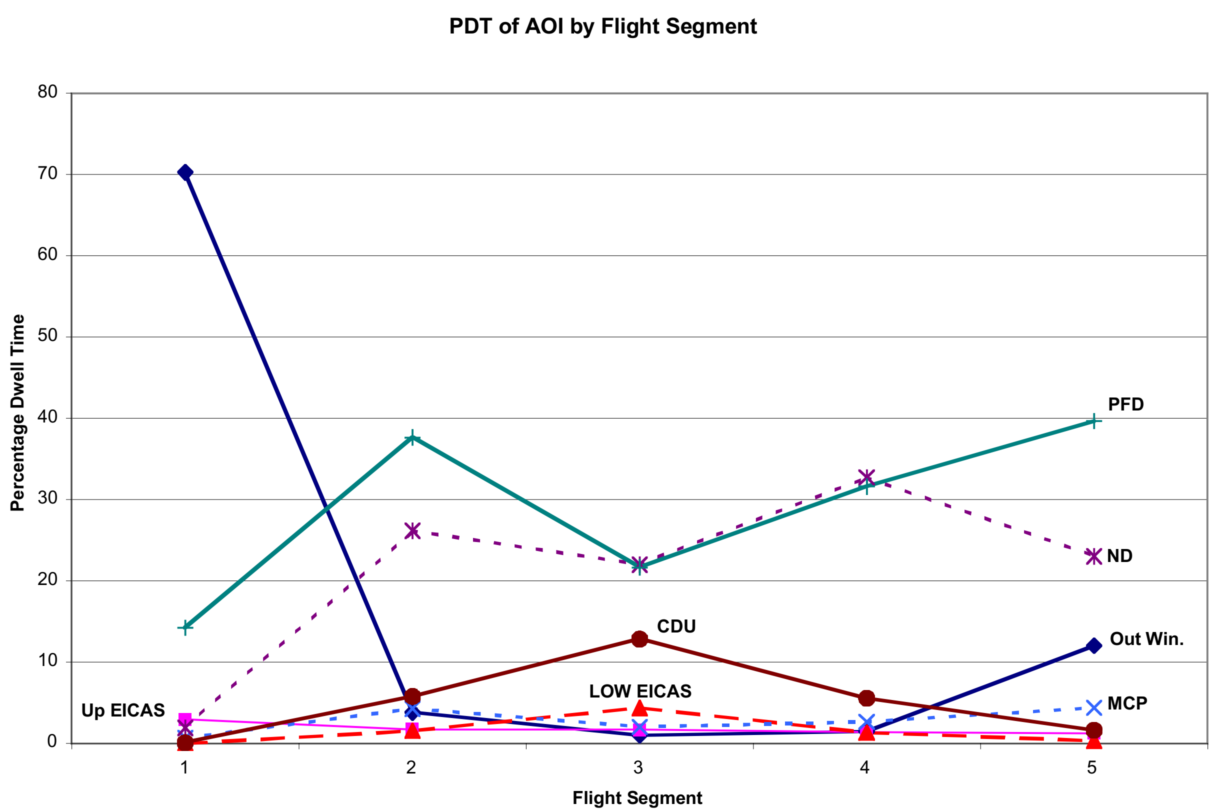


Figure 3-2. Percent dwell time for each major AOI by phase of flight.

Data plotted in Fig 3-2

Flight Phase:	1	2	3	4	5
PFD	14 %	38	22	32	40
ND	2	26	22	33	23
MCP	1	4	2	3	4
CDU	0	6	13	6	2
Upper EICAS	3	2	2	1	1
Lower EICAS	0	2	4	1	0
Out Window	70	4	1	1	12
Total Percentage:	90	82	66	77	82

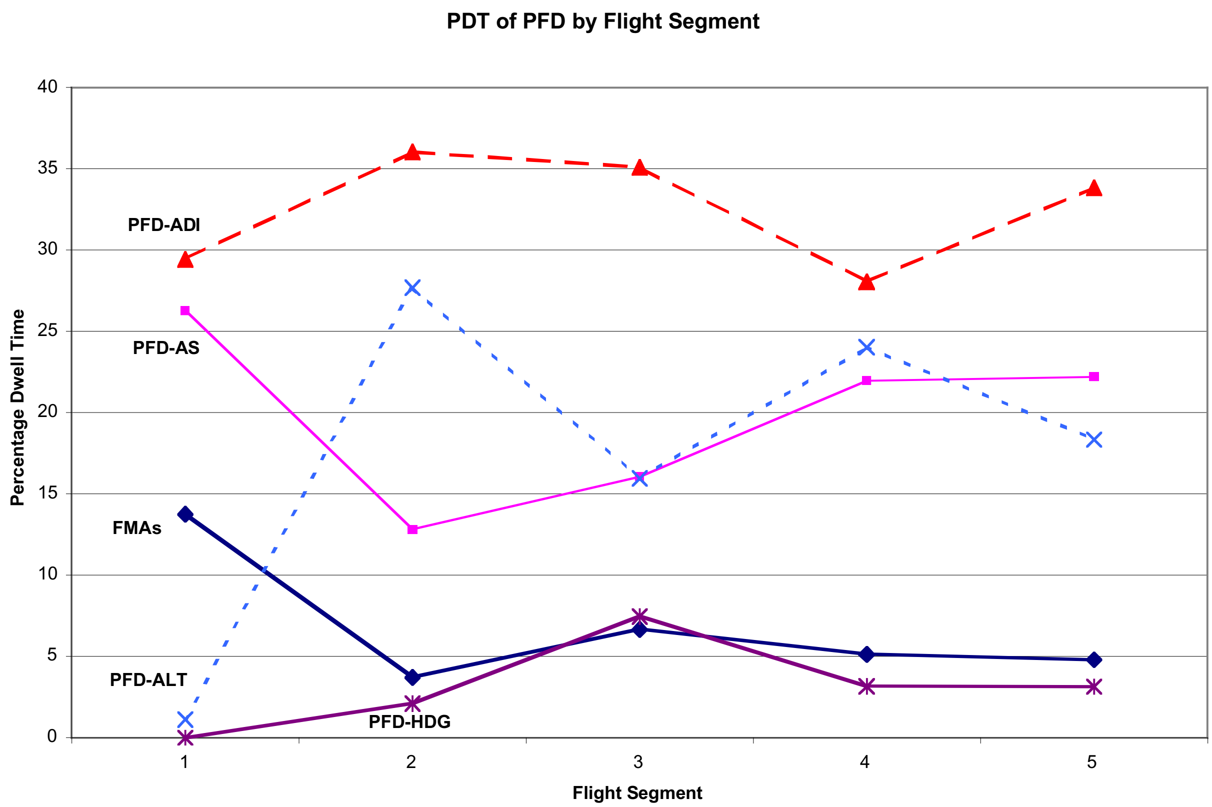


Figure 3-3. Percent dwell time for each PFD AOI by phase of flight.

Data plotted in Fig 3-3

Flight Phase:	1	2	3	4	5
PFD airspeed	26	13	16	22	22
PFD attitude	29	36	35	28	34
PFD altitude	1	28	16	24	18
PFD heading	0	2	7	3	3
PFD FMAs	14	4	7	5	5
Total Percentage:	70	83	81	82	82

Figure 3-4 presents the mean dwell duration (MDD) on each area of interest (AOI) across the full cockpit (n = 17 subjects). That is, when the fixation enters the AOI, how long does it remain there before departing. Figure 3-5 presents analogous data for the AOIs within the primary flight display (n = 14 subjects).

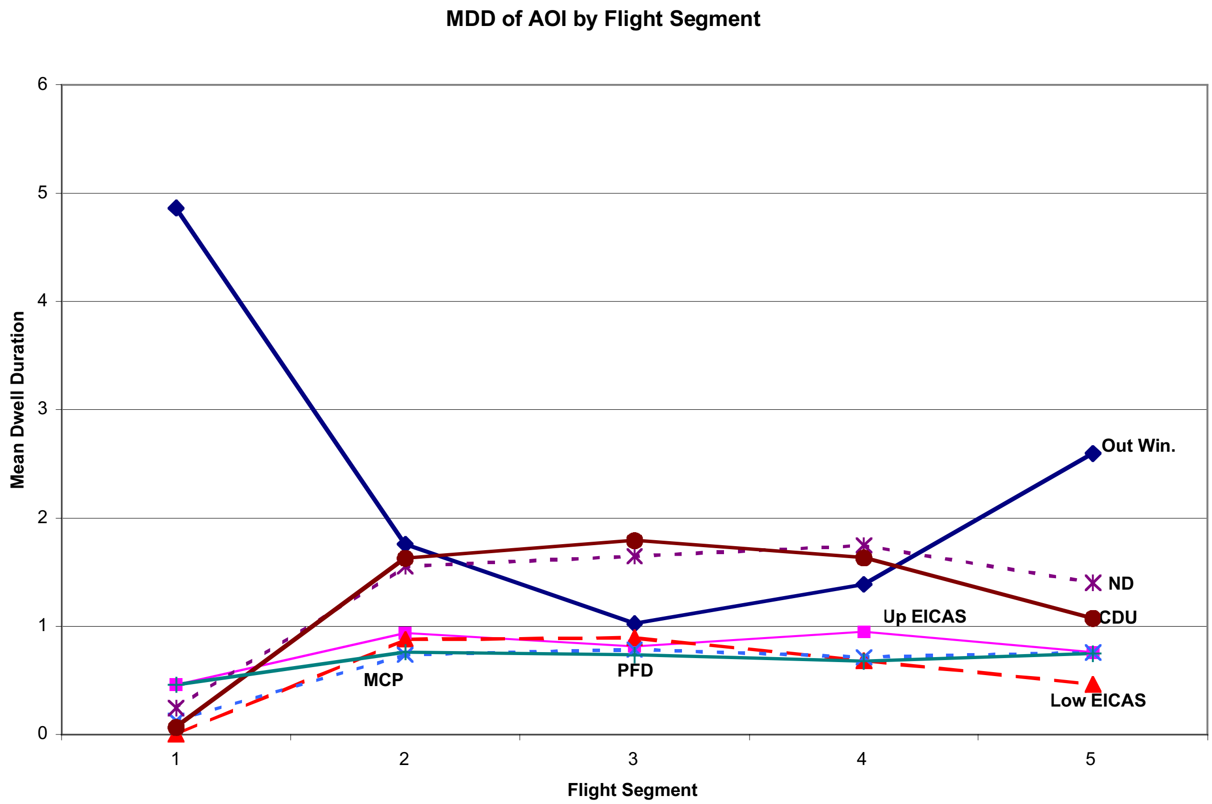


Figure 3-4 Mean dwell duration (in secs) for each major AOI by phase of flight.

Data plotted in Fig 3-4

Flight Phase:	1	2	3	4	5
PFD	0.46	0.76	0.73	0.68	0.75
ND	0.25	1.55	1.65	1.75	1.40
MCP	0.13	0.74	0.79	0.72	0.76
CDU	0.07	1.63	1.79	1.63	1.07
Upper EICAS	0.46	0.94	0.82	0.95	0.76
Lower EICAS	0.0	0.88	0.89	0.68	0.46
Out Window	4.86	1.76	1.02	1.38	2.60

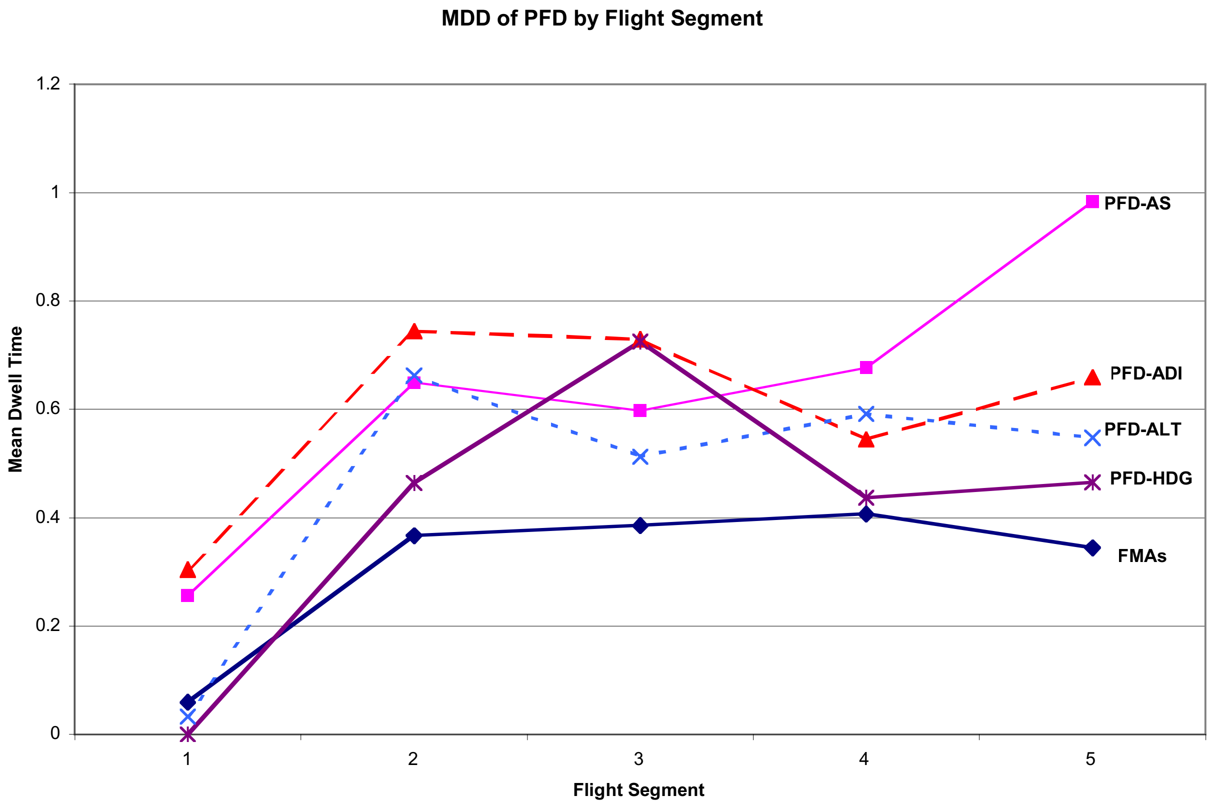


Figure 3-5 Mean dwell duration (in secs) for each PFD AOI by phase of flight.

Data plotted in Fig 3-5

Flight Phase:	1	2	3	4	5
PFD airspeed	0.26	0.65	0.59	0.68	0.98
PFD attitude	0.30	0.74	0.70	0.54	0.66
PFD altitude	0.03	0.66	0.49	0.59	0.55
PFD heading	0.00	0.46	0.74	0.44	0.47
PFD FMAs	0.06	0.37	0.39	0.41	0.34

The data shown in Figures 3-2 through 3-5 can be examined from two perspectives:

1) How do the values and percentages relate to those found in other studies of glass cockpit (Heuttig et al., 1999), and non-automated cockpit (Wickens, 2000; Wickens et al., 2000, Helleberg & Wickens, 2000; Bellenkes et al., 1997) visual scanning?
2) How do the scanning parameters change across the five major phases of flight: take-off, climb, cruise, FMS descent and ATC vectors descent?

In our description of the prominent effects below, we only describe those that reached a statistical significance ($p<.01$). In each perspective, we use the percentage dwell time as an index of the relative attention allocation, and therefore the relative importance to the pilot of a given AOI.

A key feature of Figure 3-2 is the dominance of the PFD as an area of interest, followed closely by the Nav display. In this case the PFD value of around 35% (averaged across the in-flight phases: 2-5) is consistent with the other glass cockpit study of Heuttig et al. (1999) who reported a value of around 40%. (correspondingly the Nav display value of around 25% is not far different here from the 20% value reported by Heuttig et al., 1999). The PFD value of 35% observed in the current data is substantially less than the 50-60% PFD values reported by Wickens (2000) in the general aviation (GA) studies involving CDTI and data link technology. In addition to possible differences in pilot training and expertise, a major reason for the difference between these studies can be attributed to the presence of the Nav display in the glass cockpit airplanes of the current study, which contains much of the information that would otherwise be presented in the heading display of the GA aircraft. Indeed, when pilot attention allocated to the Nav display and the PFD are summed, the total is around 60%, a value that agrees closely with that observed in the GA studies.

A second parallel with the glass cockpit data of Heuttig et al. is the very small percentage of time spent attending to the FMAs (Figure 3-3). Except in phase 1, these values hovered reliably around 5%. Also consistent with the data of Heuttig et al. is the relatively small amount of time fixated out the window. Heuttig et al. found this to be approximately 10%. In the current study, this value was also about 10% during the vector final approach (and less during earlier phases in the air). Interestingly, the relative smallness of this out-of cockpit scanning value is consistent with the conclusions drawn by Wickens (2000) in the two GA technology studies, that those pilots also undersampled the outside world, relative to the desired value (and that specified in the Federal Air Regulations). While Wickens GA pilots sampled outside approximately 30% of the time, this higher value (than the present 10%) is consistent with the habits of VFR flying, and the see and avoid prescription, characterizing much of the flight of the GA pilots used in that study.

Figure 3-4 indicates some important conclusions about the dwell duration as well: during flight (phases 2-5), the Nav display and the CDU both engage the eyes for a substantially longer duration than does the PFD (1.5 sec and 0.73 sec respectively). Within the PFD, it is interesting to note that the FMA dwells are shorter (significantly so) than all other instruments. The FMA dwell times of under 400 msec are consistent with the view that these are primarily checks , confirming a configuration or entry that is expected to be there, rather than used for extracting new information. Finally, like the expert pilots observed by Bellenkes et al. (1997), but unlike the novice pilots observed by those investigators, the pilots in the present study dwelt on the ADI no longer than on the other primary flight instruments within the basic T.

The second way of examining the data in figures 3-2 to 3-5 is to consider the relative modulation of scanning behavior (and hence perceived AOI importance) across phases. Here there are a number of predictable, and expected, effects. The most obvious of these is the large drop in Out the window scanning after the plane leaves the ground (Figure 3-2: phase 1- phase 2) and its subsequent but smaller increase after the plane approaches the ground in the final segment (phase 5). Also of interest in Figure 3-2, although small in magnitude but still statistically significant, is the increase in the MCP attention during climb and descent, relative to cruise. In Figure 3-2, the Nav display receives its greatest attention during the FMS portion of the descent, whereas the CDU receives greatest interest during the cruise phase (3). (Note that cruise may be considered the safest phase of flight, and hence the most plausible portion for pilots to be engaged in head-down programming). It is interesting to observe that across the four in-flight phases, there is a nearly reciprocal relation between attention to the CDU (automation concerns) and attention to the PFD (flying concerns).

Within the PFD (Figure 3-3), attention to the dominant ADI remains relatively unchanged across flight phases. The altimeter shows an appropriate increase of interest during vertically changing flight (climb and descent) with slightly greater interest on climb than descent. Airspeed, in contrast, shows the reversed relationship, with greater attention received during the descent than the climb. Finally, as we have noted before, interest in (attention to) the FMAs is greatest during take-off (phase 1).

A final observation, when comparing the Percentage Dwell Time (attention allocation) plots of Figures 3-2 and 3-3 with the Mean Dwell Duration plots of Figures 3-4 and 3-5, one notices that changes across segments are more pronounced with the former than the latter. That is, pilots appear to modulate their behavior more by how frequently they scan, than by how long they dwell on an AOI when it is fixated.

In conclusion, the general pattern of monitoring is consistent with the smaller glass-cockpit data set obtained by Heuttig et al. (1999), and also shows a consistent pattern of responding to the data collected in GA flight, whether pilots are given an outside scene (Wickens, 2000; Wickens et al., 2000; Helleberg & Wickens 2000) or not (Bellenkes et al., 1997).

Monitoring FMA changes.

To aid pilots in maintaining an awareness of the current flight mode annunciations (FMAs), a green box appears around each FMA (roll, pitch, and autothrottle) whenever a change occurs to that FMA. This box is present for 10 seconds and then disappears. One concern is that pilots engage a mode on the MCP but may be unaware of the actual status of the mode, which is reflected on the FMA. The green box is intended to enhance the salience of FMA changes. Pilots are also instructed to monitor the FMA after making mode changes.

We decided to see how well our subjects were monitoring FMA changes. We classified mode changes into three groups:

- Manual (M) — This case occurs when the pilot manually selects a new pitch or roll mode (e.g., FLCH, HDG SEL) by engaging a switch on the MCP. The pilot should monitor to observe that the mode selected is actually engaged.
- Automatic-Expected (AE) — This case occurs when a mode change is initiated by the automation, but it is a change that the pilot expected. For example, when the pilot is climbing in FLCH to an altitude he set in the MCP altitude window, the pitch and autothrottle modes will change as he begins to capture and level off at that altitude. The pilot should monitor to observe that the level-off actually occurs and the new modes are engaged.
- Automatic-Unexpected (AU) — This case occurs when a mode change is initiated by the automation and the pilot does not have expectations that this change will occur. For example, as the airplane descends from a level altitude, the autothrottle mode will start in IDLE and eventually will transition to HOLD mode. Pilots probably are not concerned about these changes to the autothrottle mode when they occur.

The following table shows how frequently subjects monitored the FMA for each of these classes of mode transitions (percentage is shown in parentheses).

- the first column shows the number of cases in which the subject monitored a changed FMA within the 10 seconds that the green box was present.
- the second column shows the number of cases in which the subject monitored an unchanged (irrelevant) FMA within the 10 seconds that the green box was present around the changed FMA.
- the third column shows the number of cases in which the subject monitored a changed FMA within the 10 seconds after the green box was removed (within 20 seconds of box appearing).
- the fourth column shows the number of cases in which the subject failed to monitor a changed FMA within the first 20 seconds after the green box appeared.

	changed — within 10 secs	unchanged — within 10 secs	changed — within 20 secs	failed to monitor	total cases
Manual	16 (47)	1 (3)	6 (18)	11 (32)	34
Automatic-Expected	17 (55)	0	5 (16)	9 (29)	31
Automatic-Unexpected	18 (38)	4 (8)	7 (15)	19 (40)	48

We included fixations up to 20 seconds because we felt it was possible to be alerted by the offset of the green box and then decide to fixate the FMA a few seconds later. If we take the liberal view that any fixation within 20 seconds of green box onset is tied to monitoring for that change, we see that failure to monitor is expressed by the last column. Thus, FMA monitoring occurs roughly 70% of the time for manual and expected automatic transitions, but only 60% of the time for automatic transitions that were not expected.

Note that if we take a more conservative look, and include only those fixations of the actual changed FMA during the time the green box appeared, we get a different level but similar relative conclusions. We would conclude that FMA monitoring occurs roughly 50% of the time for manual and expected automatic transitions, but only 40% of the time for automatic transitions that were not expected. A further difference between these three classes was revealed in the distribution of dwell durations that is, length of first fixation of the mode. For the AU cases, 30% of these dwells were longer than 700 msec, suggesting information extraction. However for the pooled data from the M and AE changes, only 20% of the dwells were longer than 700 msec.

3.4 Pilots Mental Model of the Functional Structure of the FMS

Pilots tend to monitor the status and behavior of the FMS in a knowledge- or expectation-driven manner (see Sarter & Woods, 1997). They combine their understanding of the functional structure of the system with their knowledge about input to the FMS to predict its behavior and allocate their attention accordingly. Therefore, gaps and misconceptions in their model of the FMS, which tend to be fairly widespread as suggested by earlier research (e.g., Sarter & Woods, 1992, 1997; Wiener, 1989), may, in part, explain poor monitoring performance.

The present study represents a first attempt to relate data about pilots understanding of the FMS to their monitoring behavior and performance data, both of which were collected throughout a simulated flight. After completion of the scenario and the debriefing, pilots were asked a number of questions about various aspects of the system s logic and operation. This section provides an overview of our findings.

Pilots were asked open-ended questions (see Appendix E), sometimes with follow-up. For example, a question might be What are VNAV PTH, VNAV SPD, and VNAV ALT? (note that a few pilots were not explicitly asked about every aspect of the system). We scored each subject s response by comparing it to a set of factual statements that we deemed to show a fairly complete understanding. For example, there were six statements associated with knowledge of the VNAV PTH mode. We indicated (see Tables 3-18 through 3-32 in Appendix F) when a subject made one of the correct statements (coded as 1) and when a subject made a statement that contradicted one of the correct statements (x). Further, a (?) means the subject explicitly said he didn t know the answer, and (NA) means the question was not asked. However, the absence of a correct statement may only mean that the subject didn t chose to comment; he may have the knowledge but have failed to offer it in this setting.

The first set of questions examined subjects understanding of the three vertical navigation (VNAV) modes: VNAV PTH, VNAV SPD, and VNAV ALT. These modes and their transitions are highly complex and difficult to comprehend for many pilots. In addition, VNAV is usually not covered in sufficient detail during pilot training due to time constraints. Instead, pilots tend to learn about the intricacies of these modes during line operations and form their mental model of these modes based on those experiences, with few opportunities for uncovering and correcting erroneous assumptions.

Vertical navigation modes: VNAV PTH.

First, the subjects were asked to explain the targets and strategies of the automation in the three different VNAV modes: VNAV PTH, VNAV SPD, and VNAV ALT. Table 3-18 (in Appendix F) shows how subjects responses compared with the following six statements about VNAV PTH (the number of subjects correctly making this statement and the number making an incorrect statement are shown after each statement):
1. the FMC flies to an FMC-calculated geographically fixed path (12 correct)
2. the FMC flies path on elevator and airspeed on autothrottle (1 correct)
3. FMC altitude and airspeed constraints will be met during descent (2 correct)
4. occurs when in cruise at FMC cruise altitude (15 correct, 1 incorrect)
5. occurs when on the FMC-created descent path (15 correct, 1 incorrect)
6. occurs during climb when there is an FMC altitude constraint tied to a waypoint (3 correct, 1 incorrect)

Discussion:
Subjects tended to offer that VNAV PTH is tied to a geographically fixed path (from the FMC) and note that VNAV PTH is typically active in cruise and descent. Three subjects made an incorrect statement with respect to the conditions under which VNAV PTH is active. Few statements were made concerning the targets and strategies of VNAV PTH.

Vertical navigation modes: VNAV SPD.

Table 3-19 shows how subjects responses compared with the following five statements about VNAV SPD:
1. the FMC controls to an airspeed target but is not linked to a geographic fixed path (4 correct)
2. the FMC flies speed on elevator (3 correct, 1 incorrect)
3. FMC altitude and airspeed constraints may be missed on descent (1 correct, 1 incorrect)
4. occurs during climb, when not level at an FMC altitude constraint (13 correct)
5. occurs during descent, when deviating sufficiently from FMC path or when using speed intervention (8 correct)

Discussion:
Again, subjects tended to note when VNAV SPD was typically active: climb and descent (off path). No false statements were made with respect to the circumstances under which the mode is active. Two subjects made an incorrect statement with respect to the operational logic of the VNAV SPD mode.

Vertical navigation modes: VNAV ALT.

Table 3-20 shows how subjects responses compared with the following five statements about the third vertical navigation mode, VNAV ALT:
1. VNAV ALT active only when level at MCP altitude (that is not in the FMC) (15 correct, 2 incorrect)
2. flies altitude on elevator and airspeed on autothrottle (0 correct)
3. will not lead to an automatic start of descent at T/D (1 correct, 1 incorrect)
4. to leave VNAV ALT, dial in a new MCP altitude above or below current altitude and push MCP

altitude button (0 correct)
5. becomes active when the aircraft levels off at an MCP altitude (9 correct)

Discussion:
As with the other two VNAV modes, subjects tended to note when VNAV ALT is active. The one subject (S14) who thought that VNAV ALT would lead to an automatic start of descent at T/D was among several subjects who were in VNAV ALT at cruise altitude and started their descent late. The other subjects did not make any statements with respect to this item.

Vertical navigation modes: VNAV mode transitions.

Next, subjects were asked about the mechanisms that can trigger a transition between the various VNAV modes. Knowing about these different possibilities is important to be able to anticipate changes in system status and behavior and thus avoid "automation surprises," which sometimes result from uncommanded (and subtly announced) mode transitions (e.g. Sarter & Woods, 1997). The following tables give an overview of subjects awareness of the various triggering mechanisms.

Table 3-21 shows how subjects responses compared with the following three statements about transitions from VNAV PTH to VNAV SPD:
1. when deviating significantly from VNAV PTH FMC-calculated path (11 correct)
2. when climb is resumed after leveling off at FMC altitude constraint (0 correct)
3. when speed intervention is used on descent (4 correct)

Table 3-22 shows how subjects responses compared with the following two statements about transitions from VNAV SPD to VNAV PTH:
1. leveling off at either an FMC intermediate level-off or at the FMC cruise altitude after climbing in VNAV SPD (5 correct, 1 incorrect)
2. recapturing the VNAV descent path on descent (transition back to VNAV PTH) (8 correct, 1 incorrect)

Table 3-23 shows how subjects responses compared with the following statement about transitions from VNAV PTH to VNAV ALT:
1. descending in VNAV PTH and using altitude intervention to level at an MCP altitude (12 correct, 4 incorrect)

Table 3-24 shows how subjects responses compared with the following two statements about transitions from VNAV ALT to VNAV PTH:
1. entering an MCP altitude lower than the CDU cruise altitude, then changing the CDU cruise altitude accordingly, and pushing the MCP altitude knob (2 correct, 1 incorrect)
2. after leveling off at MCP altitude (below FMC path) on descent, then intercepting path and descending (3 correct, 1 incorrect)

Discussion:
Subjects tended to focus on VNAV transitions during descent and they offered less information about less-common transitions. A total of eight incorrect answers were given by five subjects. One of these subjects made incorrect statements about three of the transitions, and a second subject was mistaken with respect to two types of transitions. Most incorrect responses were related to transitions to/from VNAV ALT. This supports the impression from the performance data that subjects do not have a very good understanding of this VNAV mode.

Vertical navigation modes: VNAV PTH management of path and speed.

The next question regarding VNAV asked subjects about how VNAV assigns priority to maintaining the flight path or the target airspeed on VNAV descent. Table 3-25 shows how subjects responses compared with the following two statements about VNAV priorities:

1. VNAV will give up airspeed initially to maintain the VNAV descent path, and will give up on the path and transition to VNAV SPD if the aircraft speed deviates by more than 15 kts and the aircraft deviates by more than 150ft from the flight path (15 correct, 3 incorrect; note that subjects were not required to make this statement verbatim, but clearly conveyed the basic concept)
2. Automation will maintain the path if you are low and slow, but not if you are high and fast (1 correct)

Discussion:
Fifteen subjects knew generally that VNAV PTH will abandon the airspeed target first to maintain the intended flight path but will give up on the path and transition to VNAV SPD if the aircraft speed deviates significantly in speed and distance. Three subjects made incorrect statements about the logic and priorities of this mode.

Vertical navigation modes: VNAV modes tied to phases of flight.

The last question about VNAV probed subjects knowledge of which VNAV modes should be engaged for specific phases of flight. Table 3-26 shows how subjects responses compared with the following two statements about VNAV modes:

1. it is important to verify that VNAV PTH is active at cruise altitude to make sure that the aircraft starts its descent when it reaches the T/D point (6 correct)
2. it is important to verify that VNAV PTH is active on descent because this is a prerequisite for meeting the FMC speed and altitude constraints (7 correct)

Discussion:
This is one of the strongest statements about expectations for VNAV PTH that we looked for, and it certainly relates to the performance data (see Table 3-7). Six subjects mentioned that it is important to verify that VNAV PTH is active at cruise altitude to make sure that the aircraft starts its descent when it reaches the T/D point. Two of those subjects and another group of five subjects also mentioned that they make sure to be in VNAV PTH on descent because this is a prerequisite for meeting the FMC speed and altitude constraints.

Speed protection

Table 3-27 shows how subjects responses compared with the following four statements about when speed protection is provided by the automation:

1. Speed protection is available in VNAV (17 correct)
2. Speed protection is available in FLCH (16 correct)
3. Speed protection is not available in V/S (17 correct, 2 incorrect)
4. Speed protection is not available if the autothrottle is engaged (14 correct)

Discussion:
Generally, subjects understood the limits of speed protection.

Airspeed propagation across VNAV pages.

The next question asked about data propagation throughout the CDU page architecture. In particular, we were interested to find out whether subjects understood that an airspeed that is entered on the CDU CRZ

page does not propagate to the DES page (see scenario event 7). Table 3-28 shows how subjects responses compared with the following two statements about airspeed propagation:
1. An airspeed target entered on the CRZ page does not propagate to the DES page and is not used in VNAV descent (17 correct, 2 incorrect)
2. If new airspeed is not entered on the DES page, VNAV descent reverts to the Econ descent speed (11 correct)

Discussion:
Seventeen subjects knew that they had to enter the airspeed on the DES page, also. Only two subjects believed that this was not necessary. A review of the performance data (see Table 3-9) shows that, accordingly, these two subjects failed to enter the descent speed on the DES page during the scenario. Note, however, that another eight subjects who had given the correct answer to this question performed the same actions: they used speed intervention and never entered the speed on either the CRZ or DES page of the CDU.

<u>Pitch-autothrottle mode combinations.</u>

Another area of interest was the possible combinations of pitch modes and autothrottle modes during each phase of flight, and linked to the three VNAV modes. Subjects were asked about the autothrottle modes that the automation uses. Table 3-29 shows how subjects responses compared with correct answers to the following six questions about autothrottle modes:
1. autothrottle modes during takeoff and climb can be what? (5 correct, 1 incorrect)
2. autothrottle modes during cruise can be what? (8 correct)
3. autothrottle modes during descent can be what? (6 correct, 1 incorrect)
4. autothrottle modes with VNAV PTH can be what? (8 correct, 3 incorrect)
5. autothrottle modes with VNAV SPD can be what? (1 correct, 2 incorrect)
6. autothrottle modes with VNAV ALT can be what? (7 correct)

Discussion:
Subjects, in general, were reluctant to make statements about autothrottle modes. Quite a few subjects protested that they knew little about this or paid attention to it. However, when subjects did offer statements, they seemed to understand at least two ideas. First, autothrottle SPD mode is used with VNAV PTH and VNAV ALT. Second, autothrottle modes IDLE, HOLD, and SPD are common in descent and with VNAV PTH.

<u>LNAV capture criteria.</u>

After deviating from a programmed flight route (see, for example, scenario event 4), certain prerequisites need to be met to be able to re-capture LNAV. Table 3-30 shows how subjects responses compared with the following three statements about LNAV capture:
1. the plane must be pointed at the active leg, inbound to the active waypoint (2 correct)
2. LNAV will capture when the cross track error is 2.5 NM (generally) (1 correct; note that subjects were not required to make this statement verbatim, but clearly conveyed the basic concept)
3. the distance at which LNAV will capture depends on several factors, including wind, ground speed, and course intercept geometry (2 correct)

Discussion:
Few subjects offered much on LNAV capture criteria. Two subjects correctly explained that the airplane must be pointed towards the active leg, and a third subject mentioned that the cross track error must be less than 2.5 NM. Another two subjects stated that the distance at which LNAV capture is possible depends on factors such as wind and ground speed. No incorrect statements were given.

Use of the MCP altitude knob.

During the scenario, numerous subjects were observed pushing the MCP altitude knob after selecting a new target altitude even though this action was not necessary. To better understand this behavior, subjects were asked about the circumstances under which selecting a new altitude on the MCP was not enough for that altitude to become the new active target.
Table 3-31 shows how subjects responses compared with the following two statements about pushing the MCP altitude knob:

1. the MCP altitude knob needs to be pushed when you are level at an MCP altitude and have set a higher or lower altitude and want to climb or descend; or when you are at an FMC altitude constraint but failed to reset the MCP altitude above or below you until after you pass the waypoint (or T/D) (10 correct; note that subjects were not required to make this statement verbatim, but clearly conveyed the basic concept)
2. the MCP altitude should not be pushed when setting an MCP altitude above you or below you; or when you set the MCP above you or below you and the FMC will initiate the climb or descent (5 correct, 1 incorrect; note that subjects were not required to make this statement verbatim, but clearly conveyed the basic concept)

Table 3-32 shows how subjects response compared with the following two statements about general rules for MCP altitude knob use:

1. pressing the MCP altitude knob, generally speaking, removes a restriction, which can be your current altitude, or an FMC (LEGS page) waypoint restriction, or FMC cruise altitude, or the MCP altitude you have captured (11 correct; note that subjects were not required to make this statement verbatim, but clearly conveyed the basic concept)
2. the FMC will initiate an altitude change without pushing the MCP altitude knob, generally speaking, when you are in VNAV and the MCP altitude is not your current altitude (unless you have passed the constrained waypoint or T/D point) (0 correct)

Discussion:
Ten subjects correctly stated that the button has to be pushed

a) to make sure that the aircraft begins a climb or descent when it is level at the MCP altitude and a new target altitude is entered, and
b) when the aircraft is level at an FMC altitude constraint and the MCP altitude is not reset until after passing the corresponding waypoint.

Five subjects (four of the above ten and one other subject) correctly explained that they do not have to push the MCP altitude button when they are entering a new MCP altitude and when they set a new MCP altitude and the FMC will initiate the climb or descent. Only one subject made an incorrect statement to the latter question.

Eleven subjects correctly described the circumstances and consequences of pushing the altitude select button for altitude intervention. No explicit incorrect statements were made.

Summary of mental model results

Of the 299 statements that pilots made in response to the above questions, only 24 answers were incorrect. It is important to note, however, that a total of 940 responses would have been possible (20 pilots could comment on 47 aspects of system functioning). In other words, in the majority of cases, we cannot tell whether or not pilots knew the correct answer. They may have forgotten to mention facts that they understood (e.g., they may have mentioned some, but not all, circumstances under which a mode transition can occur), or they may have avoided those areas that they were not familiar with or certain about. This limits our ability to relate the mental model data to the performance data in this study. It also highlights the need for more effective approaches to probing subjects' mental model and understanding of a highly complex system. One possible improvement may be to ask questions in the context of specific scenarios to avoid the problem of inert knowledge. Thinking through a specific case may help pilots activate and express relevant knowledge.

3.5 Results Summary

The performance analysis showed that a set of 20 experienced 747-400 pilots can have difficulty in using the flight deck automation proficiently. We were able, with some fairly commonly occurring events, to reveal misses and misunderstandings. Primary performance issues were loss of a hard altitude restriction at PORTE, mild confusion in meeting a clearance to expedite climb to 4000, poor proficiency in re-intercepting the LNAV route after vectoring off of it, setting up VNAV to descend at the T/D point, managing airspeed restrictions, and using the MCP altitude knob. In addition to the errors in using automation, there were also a handful of errors more tied to aviation safety e.g., missed altitude restrictions.

The monitoring data showed some predictable trends at a high level of analysis, and these results fit with those from other studies. Interestingly, there was considerable diversity of scan patterns across pilots, and it was difficult to find scanning patterns that were clearly better than others as had been possible to do in the studies of conventional aircraft scanning (Bellenkes et al., 1997). Due to time and resource constraints, we were able to complete only a few more-focused analyses on monitoring patterns but hope to continue to pursue these types of analysis.

Finally, taken as a set, the performance, monitoring, and mental model data provide evidence that pilots sometimes are unaware of indications that are available (i.e., present).

- The altered altitude restriction at PORTE (9000A) was present on the CDU LEGS page and on the Nav display. However, eleven subjects were unaware of the change, and many of the other subjects were confused about why the entry didn t match the clearance.
- In three separate cases, we placed an inappropriate FMA on the PFD, and only one subject commented on the oddity. We know that many of the subjects fixated the FMA but still failed to notice the mismatch. We also know that some of the subjects accurately stated the mode that should be present in an interview after the scenario was complete.
- Half of the subjects armed LNAV without seeming to be aware of the active waypoint at the time. Again, the LEGS page and the Nav display indicated the active waypoint clearly.
- Most of the subjects flew the cruise phase of the flight in VNAV ALT, apparently not realizing that this compromised the automation s ability to start descent at the T/D point.
- Subjects did not consistently check the FMA after selecting a mode or making a major transition, such as a level-off.

These failures seemingly were not failures in scanning indications (although scanning could be improved). Instead, subjects failed to integrate information taken from the world with a thorough understanding of the automation s behavior. We discuss this idea further in the Discussion section.

4. Discussion

4.1 Subject Performance

Before analyzing specific performance issues, it is important to discuss overall subject performance in the scenario. First, it is important to distinguish between performance errors that were unsafe and behaviors that reflect a poor understanding of automation use. While our Results section reveals many behaviors that we have coded as incorrect, there is a smaller set of errors that are typically classified as unsafe. Safety is directly challenged when subjects don t meet altitude or airspeed constraints that were requested by ATC for example, not meeting the altitude restrictions at 4000 ft or 9000 ft (within 300 ft) or failing to leave an altitude when you have been cleared higher. Also, exceeding 250 kts below 10,000 ft or descending at the incorrect airspeed (282 kts) can be called unsafe behaviors. Across the 20 subjects, there were 22 errors of this type:

- 10 missed altitude crossing restrictions
- 9 delays in leaving an altitude after being requested to climb
- 1 exceedence of 250 kts below 10,000 ft
- 2 violations of assigned airspeed on descent

However, this number needs to be viewed in the context that subjects were handicapped in two ways in performing the scenario. First, each subject flew with a co-pilot who offered as little help as possible. The critical PNF duty of monitoring the PF was not being carried out. When the PF misunderstood clearances and missed altitudes, the PNF did not jump in to correct the misconception. Second, we placed the subjects in an unfamiliar setting, making them wear a tight headgear and sticking an observer over their shoulder with pen and paper. This situation can engender a level of distraction that may have degraded performance, and this context needs to be considered when looking at safety violations in this scenario.

A stronger focus of the study, however, was automation use and monitoring. The 747-400 seems to prompt less automation use than many other airplanes in the fleet (e.g., 737, 757). Long-haul pilots get few opportunities to hand-fly and few landings, and so most subjects told us that, on the line, they use LNAV/VNAV primarily only above 10,000-18,000 ft. One might expect them to be less proficient with the automation level we imposed on them. On the other hand, each pilot indicated that there are conditions on the line under which they will fly LNAV/VNAV with autopilot for virtually the entire flight. Further, these pilots are all certified to fly the 747-400 with full automation, so a certain level of proficiency is expected. And, the scenario we used was a fairly quiet one. Events were separated by long periods of routine flight, and no real emergencies occurred.

The performance problems we are focussing on concern situations such as starting down late from cruise, which by itself is not a significant threat to aviation safety. However, we strongly believe that a high level of proficiency with an automated tool is a safety issue. A Boeing review of automation-related accidents and incidents (Mumaw et al., 2000) indicates that poor automation use can lead to significant accident precursors such as path deviations, missed altitudes, loss of control, etc. It is critical for pilots of glass cockpit airplanes to understand and be proficient in using the automated tools available to them.

4.2 Performance and Outcomes

Another general issue tied to performance measurement in a realistic scenario is the too-often loose coupling between process and outcome. It is almost impossible to execute a scenario the same way every time. As much as we tried to script every action, there is still a dynamic quality to how events play out. For example, we sometimes struggled to get each subject back on the VNAV descent path after the ATC request to resume normal speed. For one subject (S14), our inability to do this prevented us from introducing the artificial VNAV SPD and THR modes on descent. Also, in at least one case, a subject s error failed to produce a bad outcome: S20 was in VNAV ALT on cruise but the simulator erroneously started his descent at the T/D point.

The point is that there is sometimes a loose coupling between the subject s behavior and larger performance measures. Because of this, we have tried to look at performance in individual events and subject s behaviors in those events. We chose not to rank subjects as being generally good or bad across the set of events in the scenario.

4.3 Contributors to Performance: Interface and Alerting

Monitoring can be supported by alerting schemes in the interface. That is, effective interfaces can identify meaningful events and alert the pilot (through a salient visual or auditory cue) about changes to the system state. Or, the interface can offer intuitive, integrated views of system performance that allow pilots to evaluate performance with brief examinations. However, today s glass cockpits are not effective monitors of changes to automation state. In fact, the current flight deck interface design creates barriers to effective monitoring in a number of ways.

First, there is a separation of mode engagement controls, primarily located on the MCP, and mode annunciations, which are on the PFD. There is a natural tendency for pilots to assess mode status by looking at the location where selections are made. However, a mode may be selected on the MCP but not actually engaged, or it may be selected and armed, but not engaged. Knowing which mode is in effect can only be determined by monitoring the PFD. Also, selecting VNAV leads to a specific VNAV mode being engaged (e.g., VNAV SPD), which is not revealed by the MCP. For example, as we saw in this study, subjects were in cruise, the MCP showed VNAV was engaged, but subjects had to monitor the PFD to determine that it was VNAV ALT instead of VNAV PTH. Likewise, autothrottle modes can change in ways not shown on the MCP (e.g., going from IDLE to HOLD). Although pilots are trained to monitor the PFD for mode status, they don t always do it. In fact, our data suggest that, at best, they only do it 60 to 70% of the time.

Second, indications are widely distributed to the MCP, PFD, and CDU. Although the interface is designed so that the information on CDU pages is not required for assessing the current situation, CDU pages provide valuable supplemental information (e.g., transition points, track error values, etc.). The CDU is problematic in that only two pages (of the more than 70 available) can be viewed at a time. Pilots must occasionally change the selected page to make sure the most useful information (for the phase of flight) is being displayed. And, other tasks, such as selecting a radio frequency, can force the removal of a page that is supporting monitoring. If the crew fails to keep the most appropriate CDU page displayed, valuable information may not be monitored in a timely way. We found that our subjects were not disciplined in maintaining the optimal set of CDU pages at all times.

Third, some modes (i.e., system states) are not annunciated and yet change the automation system s behavior. For example, there is an early descent zone that is defined as the cruise segment within 50 NM of the FMC T/D point. The outcome of certain early descent actions can change depending on whether the airplane is in this zone. Other unannunciated system states are VNAV approach (where the airplane can fly through the MCP altitude target) and altitude capture. Altitude capture is a transitional pitch mode occurring between climb or descent and a level segment. Certain automation actions (e.g., removing FMC altitude constraints with the MCP altitude selector) do not work in this transitional phase.

Fourth, transition points are not always shown. The automation determines transition points (for airspeed targets, level-offs, descents, holds, etc.). Few of these points are communicated through the interface, and pilots cannot determine when the airplane behavior should change. One example is that the acceleration height (3000 ft) was only shown on the CDU page, and the transition point is not shown on the Nav display. Because of this, some of our subjects failed to anticipate the transition from climb to acceleration that threatened their ability to expedite climb to 4000 ft after take-off. An important special case of unannunciated transitions are automatic mode transitions in particular, VNAV PTH to VNAV SPD reversions. It is up to the pilot to monitor effectively enough to pick up these transitions.

Fifth, the interface can present two targets each for airspeed and altitude. The MCP altitude window always contains an altitude target. The FMC may also provide a relevant altitude target. However, the FMC target is only meaningful if VNAV is engaged (and even with VNAV engaged, the FMC altitude target may not be met e.g., if you are above the VNAV descent path). A similar story exists for airspeed targets, although the MCP airspeed window does not always show an airspeed target. A general rule is that the MCP target overrides the FMC target. The primary problem is that pilots have trouble evaluating whether the FMC target will be honored. The FMC shows what the autopilot may do but it gives poor feedback on what it actually will do. Pilots need a good understanding of the automation to correctly anticipate which FMC targets will be honored.

Sixth, there are cases where an action taken can have unanticipated side effects. In our scenario, subjects were asked to change departure runways, and this action removed a hard altitude restriction at PORTE and created a discontinuity after the SID. Subjects were largely unaware of these changes unless they noticed their effects for another reason.

Finally, there are a few cases of cueing by absence, which means a state change is indicated by removing an indication. In particular, we saw that a certain glideslope failure removes the glideslope cue from the PFD. If pilots are not monitoring for the glideslope, they will be unaware that the signal is missing.

The point of listing out these characteristics of the flightdeck interface is to show that much of the burden of automation monitoring rests entirely with the pilots who, based on their understanding and anticipation of system performance, will actively search for information. The system interface does not support the other complementary component of effective human-machine coordination: data-driven monitoring. It does not alert the pilot to changes, current targets, upcoming transition points, system side effects, etc.

4.4 Contributors to Performance: Pilot Mental Model and Knowledge-Based Monitoring

Because of the failure of the system interfaces to support sufficiently data-driven attention allocation, operators (pilots) need to know on their own where and when to look for important indications. More specifically, pilots need to anticipate changes that will occur and monitor for them e.g., anticipate a transition to VNAV PTH in cruise and monitor for that change. This type of monitoring (knowledge-based monitoring) needs to be supported by a fairly complete and accurate mental model of the system. However, the results suggest that our subjects sometimes failed to engage in this type of monitoring.

Failures take two forms. First, the subjects could fail to look for the indication or change in indication. If they do not anticipate it, they will not look. Second, even when the subjects looked, they could fail to understand the implications of an indication. We found both types of failures in this study.

4.4.1 Fail to Monitor

In many cases, indications are present (are available) but are not observed by the subject, as illustrated by the following cases:

First, when we required subjects to change the departure runway, two changes occurred: the hard restriction at PORTE was changed to 9000A and a discontinuity was created after the SID. Indications for each of these were present among the full set of indications on the flight deck interface. The new altitude restriction for 9000A appeared on both the LEGS page (if it was displayed) and on the Nav display. However, only four of 20 subjects noticed this change prior to take-off, even though the indications were present. This is not surprising since there is nothing salient about the indications; they are not changed in form after they revert to 9000A. The same is true with the creation of a discontinuity although indications may not have been present early in the scenario. The discontinuity does appear on the LEGS page just after the plane passes PORTE, but most subjects didn t notice this discontinuity until quite a bit later.

Second, we found that subjects didn t always look at the FMA to confirm the selection of a mode or a transition to a mode tied to level-off. Subjects may rely primarily on monitoring the MCP or looking at other airplane behavior to determine that the mode transition has occurred.

Third, a number of subjects failed to look at the pitch FMA during descent when we had artificially altered it to VNAV SPD. It is possible that subjects were using other indications during this period (e.g., the vertical path indicator) to monitor how well they were maintaining the VNAV descent path (we haven t done this analysis yet).

These cases show that subjects do not always monitor indications that could help them verify the state of the automation. However, a bigger concern, addressed in the next section, is that even when an indication is fixated, the subject doesn t understand the implications of the indication.

4.4.2 Fail to Understand Implications

Monitoring (fixating) is only the first stage of maintaining an awareness of the automation state. It is also critical to understand the implications of the indications. Our data revealed a number of cases in which subjects seemed to fail to understand the importance of an indication.

The most telling example is the presence of VNAV ALT during cruise. Sixteen of the 20 subjects flew the entire cruise segment in VNAV ALT, instead of VNAV PTH. For some of these subjects, it was clear that they did not understand that the airplane would not descend at the FMC T/D point in this mode. The eye fixation data show that, in general, these subjects had fixated the pitch FMA, but they failed to understand its meaning. One subject actually commented, It s in VNAV, it should start down. It seems that, in his mind, VNAV is treated as the primary mode and the modifier (ALT or PTH or SPD) has little meaning.

A second example is the set of artificial modes that we introduced: VNAV PTH on climb, VNAV SPD on descent, and THR on descent. Subjects (with one exception) failed to comment on these incorrect mode annunciations, suggesting that they did not know they were inappropriate. In some cases, these subjects made statements in the mental model interviews that suggested they knew what modes are appropriate for climb and descent, but this knowledge was not applied in this situation. It is possible that their mental model knowledge is in an abstract form (called inert knowledge) and untied to actual knowledge use.

A third example is the use of the APP mode when there is no glideslope signal. In this case, the subject should select the LOC mode. The automation in APP is giving pitch guidance through the FD pitch bar that is not tied to an actual glideslope signal. Apparently, most of our pilots did not understand this element of using APP mode.

A fourth example is the acceleration height. As part of pre-flight, the TP and the subject entered an acceleration height of 3000 ft. This became a significant number when each subject was trying to expedite climb to 4000 at D6. At 3000 ft, the airplane pitched down to accelerate, and its climb rate was reduced. For some subjects, this shift to speed over climb made it more difficult to reach 4000 ft at D6. The subjects knew about the acceleration height value but failed to account for the implications of this value during the maneuver. Note that it is possible that the urgency of the situation may have also influenced the subject s inability to apply this knowledge.

In other cases, subjects behavior suggested that they did not fully understand the automation system behavior (and monitoring was less of a contributor). The prime example is the MCP altitude knob. Not one subject used it correctly every time. Part of the reason is that it is an input device with a fairly complex set of rules behind it.

4.5 Design Implications

These results suggest that future flight deck interface designs need to provide more direct feedback about the authority/control actions of the automation. Pilots would benefit significantly from a graphical depiction of the airplane s path: whether it will begin descending, whether it will change airspeed, whether it will level off, etc. Strategic flight planning is handled by the FMC. Pilots can indicate the set of restrictions on vertical path, lateral path, and airspeed they desire. In addition, however, pilots need a more tactical display that indicates clearly the actual path and performance parameters the airplane will take, and how that relates to what was planned in the FMC.

4.6 Pilot Proficiency and Automation Level

In this study, pilots were asked to use full automation: autopilot coupled to LNAV/VNAV. A number of pilots talked to us about how reliance on automation may be tied to a pilot s age/experience. The stereotype we heard is that older pilots (who were trained in pre-glass days) learned to fly the airplane first and use automation when it made sense. The younger pilots (who were computer-literate before learning to fly) gravitate to the CDU/FMC and aren t as capable with hand-flying. Moreover, there has been an untested assumption that as pilots rely more on automation, their hand-flying skills degrade. Thus, a tension is posed between skilled automation use and skilled manual flying.

Because we did not ask pilots to fly any way other than with the autopilot, we were unable to evaluate this question. However, one method for looking at this issue is to see how quickly (or under what circumstances) a pilot gives up VNAV/LNAV and reverts to MCP modes or to manual flying. If someone is uncomfortable with VNAV/LNAV, he is likely to abandon it more quickly than another pilot might. This analysis will be done at a later time.

4.7 Further Analysis of Eye Fixations

The current study focused on performance data, the mental model data, and certain issues within the eye fixation data. As with all studies, resources were limited and certain analyses were given less priority. We hope to continue to look at more detailed analysis of the eye fixation (monitoring) data. In particular, now that we have established where differences lie in subject performance, we can conduct a more detailed analysis of monitoring patterns to differentiate between good and poor performers. After looking at monitoring in individual events, we can attempt to characterize good monitoring practices more broadly. That is, there may be certain monitoring patterns that characterize better performers. Also, we need to determine how the demographic variables (flight hours, hours in the 747-400, seat, etc.) are correlated with performance, mental model, and monitoring data.

4.8 Conclusions

We found that experienced 747-400 pilots had difficulty with a number of the scenario events involving proficient use of automation or automation monitoring. In some cases, subjects failed to detect meaningful indications that were present (e.g., a change in a waypoint altitude constraint) or to look for those indications (e.g., FMA changes). In other cases, subjects failed to understand the implications of the indications they monitored (e.g., the implications of VNAV ALT during cruise). These findings bolster the literature suggesting that the flight deck interface fails to provide a complete picture of the automation s authority and control actions. Further, these findings reveal the importance of pilots having strong expectations of the automation s behavior to support knowledge-based monitoring. Because much of the burden of monitoring the automation falls on the pilot, pilots need a more complete and accurate mental model of automation behavior.

We used an open-ended approach to exploring subjects mental models. Our results show that when subjects offer statements about automation behavior, they usually articulate correct statements (about 92% of the time). However, subjects either do not know about certain aspects of automation behavior or were unwilling to offer other statements in this context. Further, even when subjects made a correct statement about which mode to expect in a certain situation, they did not always detect inappropriate modes in those situations. Therefore, their knowledge may be poorly linked to the operational context.

The monitoring data showed some predictable trends at a high level of analysis, and these results fit with those from other studies. Interestingly, there was considerable diversity of scan patterns across pilots, and it was difficult to find scanning patterns that were clearly better than others as had been possible to do in the studies of conventional aircraft scanning (Bellenkes et al., 1997).

5. References

Abbott, K., Slotte, S., Stimson, D., Bollin, E., Hecht, S., Imrich, T., Lalley, R., Lyddane, G., Thiel, G., Amalberti, R., Fabre, F., Newman, T., Pearson, R., Tigchelaar, H., Sarter, N., Helmreich, R., and Woods, D. (1996). *The Interfaces Between Flightcrews and Modern Flight Deck Systems.* Federal Aviation Administration Human Factors Team Report. Washington, D.C.: Federal Aviation Administration.

BASI (Bureau of Air Safety Investigation) (1998). *Advanced technology aircraft safety survey report*. Canberra, Aus: BASI.

Bellenkes, A.H., Wickens, C.D., & Kramer, A.F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine*, 68, 569-579.

Carbonnell, J.R., Ward, J.L., & Senders, J.W. (1968). A queuing model of visual sampling: Experimental validation. *IEEE Transactions on Man-Machine Systems*, MMS-9, 82-87.

Eldredge, D., Mangold, S., & Dodd, R.S. (1992). *A review and discussion of flight management system incidents reported to the Aviation Safety Reporting System*. (DOT/FAA/RD-92/2). Washington, DC: FAA.

Feltovich, P.J., Spiro, R.J., and Coulson, R.L. (1991). *Learning, Teaching and Testing for Complex Conceptual Understanding* (Technical Report No. 6). Springfield, IL: Southern Illinois University School of Medicine - Conceptual Knowledge Research Project.

Fitts, P., Jones, R.E., & Milton, E. (1950). Eye movements of aircraft pilots during instrument landing approaches. *Aeronautical Engineering Review*, 9, 24-29.

Harris, R.L., & Christhilf, D.M. (1980). What do pilots see in displays? In G. Corrick, E. Hazeltine, & R. Durst (Eds.), *Proceedings of the 24th Annual Meeting of the Human Factors Society* (pp. 22-26). Santa Monica, CA: Human Factors Society.

Helleberg, J., & Wickens, C.D. (2000). *Datalink modality in general aviation* (Technical Report ARL-00-7/FAA-00-4). Savoy, IL: University of Illinois, Aviation Res. Lab.

Huettig, G., Anders, G., & Tautz, A. (1999). *Mode awareness in a modern glass cockpit attention allocation to mode information*. In R. Jensen (Ed.), Proceedings of the 1999 Ohio State University Aviation Psychology Conference. Dayton, OH: Ohio State University.

Hutchins, E. (1997). *The integrated mode management interface*. NASA Contractor Report NCC 2-591. Moffett Field, CA: NASA-Ames Research Center.

Lyall, B. & Funk, K. (1998). *Flight deck automation issues.* Proceedings of the third automation and human performance conference, March 25-28, 1998. Norfolk, Virginia: Old Dominion University.

Mitchell, C.M., Chappell, A.R., Gray, W.M., Quinn, A., & Thurman, D.A. (2000). *Intelligent tutoring systems for aviation automation*. Paper Presented at the World Aviation Congress, San Diego, CA.

Mumaw, R.J., Boorman, D., Griffin, J., Moodi, M., & Xu, W. (2000). *Training and design approaches for enhancing automation awareness*. (Boeing Document D6-82577). Seattle, WA: Boeing.

Norman, D.A. (1990).˚ The problem of automation: Inappropriate feedback and interaction, not over-automation. *Philosophical Transactions of the Royal Society of London,* B 327:585-593.

Sarter, N.B. (2000). The need for multisensory feedback in support of effective attention allocation in highly dynamic event-driven environments: The case of cockpit automation. *International Journal of Aviation Psychology,* 10(3), 231-245.

Sarter, N.B., & Woods, D.D. (1992). Pilot interaction with cockpit automation: Operational experiences with the flight management system. *The International Journal of Aviation Psychology,* 2, 303-321.

Sarter, N.B., & Woods, D.D. (1994). Pilot interaction with cockpit automation II: An experimental study of pilots model and awareness of the flight management system. *The International Journal of Aviation Psychology,* 4, 1-28.

Sarter, N.B., & Woods, D.D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors,* 37, 5-19.

Sarter, N.B., & Woods, D.D. (1997). Team play with a powerful and independent agent: Operational experiences and automation surprises on the Airbus A-320. *Human Factors,* 39, 553-569.

Sarter, N.B., & Woods, D.D. (in press). Teamplay with a powerful and independent agent: A full-mission simulation study. *Human Factors,* 42(3)

Sarter, N.B., Woods, D.D., & Billings, C.E. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd edition, pp. 1926-1943). New York: Wiley.

Wickens, C.D. (2000). *Pilot expectancy and attentional effects for hazard awareness: Implications for free flight and datalink.* (Final Technical Report ARL-00-6/FAA-00-3). Savoy, IL: University of Illinois, Aviation Res. Lab.

Wickens, C.D., Xu, X., Helleberg, J.R., Carbonari, R., & Marsh, R. (2000). *The allocation of visual attention for aircraft traffic monitoring and avoidance: Baseline measures and implications for free flight*. (Technical Report ARL-00-2/FAA-00-2). Savoy, IL: University of Illinois, Aviation Res. Lab.

Wiener, E.L. (1989). *Human factors and advanced technology (glass cockpit) transport aircraft*. (TR 177528). Moffett Field, CA: NASA Ames Research Center.

Woods, D.D. and Sarter, N.B. (2000). Learning from automation surprises and "going sour" accidents. In N.B. Sarter and R. Amalberti (Eds.), *Cognitive Engineering in the Aviation Domain* (pp. 327-353). Hillsdale, NJ: LEA.

Acknowledgments

The following people at Boeing provided important support for preparing and maintaining the simulator facility and in collecting data: Charlie Bell, Jack Butcher, Nate Chan, Dennis Cleary, Robert Corcoran, Greg Dorland, Harold Edwards, Diane Ford, Jenny Gertley, Marty Hartell, Dean Holt, Art Howell, Malon Hull, Tom Kmitta, Mike Moodi, John Nymon, Sue Orlicky, Dale Pool, Mike Quigley, Bill Shontz, Dave Swan, Steve Sweum, Jim Terhune, Laura Townsend, Ron Wiertella, and Chuck Wilson. Also, thanks to John Helleberg and Bobbie Seppelt at the University of Illinois.

Capts. Jack Bard and Bob Reid (both retired) were critical for the study s success through their insights into operational practices and ability to execute the scenario successfully.

Appendix A
Informed Consent Form

The purpose of this study is to better understand how experienced 747-400 pilots use the aircraft s automation for routine tasks. We are going to have you fly a routine revenue flight from San Francisco to Los Angeles in a fixed-base simulator to observe and measure the way you use the automation tools available to you. The duration of the scenario should be about 1 hour, followed by 30-60 minutes of debriefing.

We are going to measure performance in several ways.
- The simulator will track and record your performance and your ability to meet the requirements of the scenario.
- We will ask you to wear a piece of headgear that measures eye-fixations. That is, it records where you are looking as you fly. It will tell us which instruments you are using and how often you look at them. In brief, the eye-tracking system integrates head position and eye position to come up with point of gaze information. Head position is determined using a magnetic sensor attached to a head band. Eye position is determined using the retinal reflection through the pupil and corneal reflection of a tiny infrared light source. These reflections are input to a camera that is also mounted on the headband and then in turn to a computer for processing. The headband weighs 2 pounds with those components on it.
- We will have an observer who will be recording your response to certain events in the scenario. All of these data will be later combined to give us a clearer picture of how you are using the automation.

This is not a scenario designed to trip you up or make you fail, but there will be a number of events that you will need to respond to so that we can understand how 747-400 pilots respond to common ATC requests, etc. We want all pilots to fly successfully so we can carefully document important pilot skills.

IMPORTANT: Your participation in this study is **completely voluntary**. If at any time during the study you would like to stop and leave, you are free to do so. There will be no hazards or penalties associated with withdrawing.

<u>Potential Benefits of this Research</u>

Boeing is conducting this research in collaboration with NASA to advance our knowledge of how line pilots use current advanced technology flight deck interfaces. We hope to apply the results to improve either future interface designs or the materials that are used in transition training.

<u>Potential Risks to You</u>

In this study, we will be recording your eye movements that is, we will be tracking where your eyes are looking as an indication of which flight data you are using. To record your eye movements, we will ask you to wear a type of head gear that holds a reflector over your eyes. This head gear will be fitted to your head to be as comfortable as possible and to calibrate it properly. We will do what we can to remove any discomfort you experience, but there is the possibility of some mild discomfort associated with wearing the headgear. If you do not want to wear this apparatus, we will stop the study and let you leave.

Also, as mentioned above the eye-tracking system reflects a low-level infrared light off of your eye. This is an established technology that the manufacturer has assured us uses radiation levels several magnitudes below the levels that have potential for any damage, and we have used this system in previous studies with no eye injuries. This system creates no risk for eye injury.

Confidentiality of Results

We will be recording your flight performance in the simulator, your eye movements, and videotaping your performance in this session. We will also ask you to answer some questions after the simulator session regarding the 747-400 automation. These data will NOT be linked to your name, but will be coded with an identifier (e.g., a number). Also, in our report, we will refer only to an American air carrier, and will not mention A/B Airline specifically. These data will be analyzed along with the data of about 20 other 747-400 line pilots. All results will be reported for the group. If an individual s results are reported alone, they will be identified only with the number designation, not by name.

Complete Results of Study

We will be able to answer some questions today about this study and your performance. However, after we have collected data from all subjects and completed our analysis, we will send you by mail a more complete description of our results and how we are likely to use the data.

If you have any questions about this study or your rights as a study participant after you leave here today, you can contact Dr. Randy Mumaw at Boeing: (425) 965-0793.

Please sign below to indicate that you have read and understood this consent form:

I understand that my performance data will be recorded and analyzed for research, and that Boeing and NASA will not link my name to my performance record. No test subject s identity will be released to anyone without that person s consent, unless specifically required by law. I also understand that I am participating as a volunteer in this study and that I may leave at any time without any penalty.

Signature ____

Name (Printed) ____________________________________

Date _____________

Appendix B
Subject Briefing Narrative

Objectives

The purpose of this study is to better understand how experienced 747-400 pilots use the aircraft s automation for routine tasks.

We are going to have you fly a routine revenue flight from San Francisco to Los Angeles to observe and measure the way you use the automation tools available to you. In general, we would like you to minimize manual flying with the Flight Director bars and use the automation in a way that works best for you. That is, use the MCP or CDU/FMC as you deem appropriate.

We are going to measure performance in several ways.
- the simulator will track and record your performance and your ability to meet the requirements of the scenario.
- we will ask you to wear a piece of headgear that measures eye fixations. That is, it records where you are looking as you fly. It will tell us which instruments you are using and how often you look at them.
- we will have an observer who will be recording your response to certain events in the scenario. All of these data will be later combined to give us a clearer picture of how you are using the automation.

This is not a scenario designed to trip you up or make you fail, but there will be a number of events that you will need to respond to so that we can understand how 747-400 pilots respond to common ATC requests, etc. We want all pilots to fly successfully so we can carefully document important pilot skills.

Simulator Notes

The simulator we are using today is a fixed-base simulator. It has a full 747-400 model and cockpit, but there is no motion platform, and you might find the visual displays through the windows to be of lower quality than what you expect in your airline s training simulators.

There are a few other limitations on the simulator: some of the sensors and controls are simulated less realistically for example, the IRS panel is a dummy panel and does not tie into the simulation. TCAS is not working on this system. The fuel flow gauges do not work correctly. These items are not important elements for the scenario you are flying.

Note that every now and then the simulator presents an incorrect indication. For example, there is a CDU message about fuel disagreement that is meaningless. When you see this message or other indications that you think are not valid, please let the Team Pilot know, and in most cases he will clear the indication, or at least make a note of it so we can let others know. These invalid indications will not occur frequently.

Finally, for recording your communications with ATC, we will have a speaker on in the simulator.

Scenario Notes

As I said above, the flight today is a revenue flight from San Francisco to Los Angeles, as you might fly for an extra section. You will get the full set of briefing materials with the necessary procedures, Jeppeson charts, and weight manifest. Note that the Navigation database on the simulator and the Jeppeson charts are frozen to an earlier date so that all pilots work with the same information.

When you get into the simulator, it will be positioned on the taxiway to 28L and 28R at SFO with the engine running. Please take as much time as you need to familiarize yourself with the initial configuration and feel comfortable that everything was set up correctly. The flight plan will already be programmed in the FMC for you, but you should review it to ensure it is correct with regards to your clearance. Please ask questions if you need to and let the Team Pilot know when you are ready to begin.

After landing at LAX, the scenario will stop. We will get out of the simulator to review the scenario during a debriefing session.

We will also be getting the equipment ready for taking the eye-movement data, and some calibration will have to be done when you first get into the simulator. Also, for some pilots we may have to recalibrate during the scenario, and this might take a few minutes.

Finally, because we are studying your use of automation, we would like you to engage the autopilot as soon as possible above 400 feet instead of hand-flying on departure, and you should use VNAV when you can, especially on descent.

Team Pilot Role

In the other seat will be a Team Pilot who will support you. You will be flying with either Capt. Jack Bard or Capt. Bob Reid, both recently retired 747-400 pilots. In general, the Team Pilot is there to ensure that your tasks can get done, but this pilot will NOT be as proactive as is typical. The Team Pilot will take the Pilot Not Flying role as would be done for a training scenario.

Also, note that the Team Pilot is not there to create problems. He will carry out your requests to the best of his ability.

Please review the briefing package now (this package included the relevant Jeppeson charts and the weight manifest). Note that the first waypoint out of SFO is labeled D6 on the Jeppeson charts, and ATC may refer to it as SFO 6 DME, but it shows up as N37W122 on the Nav display.

Your initial clearance is shown on this sheet of paper:
Airline A/B 400,
cleared to LAX airport via PORTE 3 departure,
AVENAL transition,
direct DERBB,
then via Los Angeles SADDE6 arrival,
DERBB transition.
Maintain flight level 350.
Cross PORTE intersection at 9,000, then resume climb.
Contact Bay departure 135.1

Appendix C
Scenario Procedure

Staffing:
Cab Ops — simulator cab operator; ran the simulator (TKmitta)
Mgr — study manager; coordinated activities (R. Mumaw)
ATC — air traffic control; provided ATC support (J. Bard, B. Reid, B. Shontz, or D. Holt)
TP — team pilot; pilot not flying role (J. Bard or B. Reid)
ET — eye tracker; calibrated and monitored the eye tracking equipment (W. Xu)
Obs — observer; sat behind the subject and recorded key behaviors (S. Kimball)

Simulator Pre-Flight List

Cab Ops
- simulator up and Cab Ops checklist complete
- illuminate Simulator in Use sign (#6 button)

Mgr/Obs
- set pilot clocks to local time
- turn down intercom/PA dial on wall
- simulator speakers on (switch on center console); left radio selected
- verify Cab Ops has armed simulator data collection
- check that correct PFD displays are in
    - 30097860 (to be used for ATC station)
    - 30160741 (to be placed in subject s side - default to FO)
    - 30097946 (to be placed in team pilot s side)

TP
- load flight plan according to Flightplan Checklist
- set MDA and DH to 0
- check all breakers (e.g., for aurals)
- get Checklists from Mgr

ET
- completes ET checklist

ATC
- check communications working (except lapel mic)
- check the loudspeakers to ensure that ATC can be heard clearly without headset.
- set Nav display range at 40 nm (through Cab Ops)

****

Subject Arrives

Mgr
- brief subject in Briefing Room
    - briefing notes
    - clearance written on separate sheet
    - charts
    - manifest copy

- procedures
- complete subject contact info and history form
- make sure Subject has read and signed Informed Consent

****

Subject Comes to Cab

TP
- review flight plan
- review EICAS and (many) STATUS messages; in particular
    - oil temperature odd behavior
    - ground proximity message
    - fuel disagree message
    - TCAS inoperable
- get subject comfortable in simulator

Mgr
- install lapel mic
- make sure sliding door closed
- make sure storm lights are on in cab and filtered
- make sure room lights are lowered appropriately (preset A)

ET
- fit head gear and calibrate subject

ATC
- check Mgr s lapel mic connection to ATC station

****

Subject ready; Team ready; Sim ready

Mgr/Obs
- push real time button on sim

ET
- start eye track, start video

TP
- use Before Takeoff Checklist

ATC
- after go-ahead from Mgr, give first clearance:

Airline A/B 400, cleared to LAX airport via PORTE 3 departure, AVENAL transition, direct DERBB, then via Los Angeles SADDE6 arrival, DERBB transition. Maintain flight level 350. Cross PORTE intersection at 9,000, then resume climb. Contact Bay departure 135.1

TP
- give OK to Subject to taxi

****

Subject starts to taxi

ATC
- when evidence of taxiing (ground speed), give runway change clearance:
Airline A/B 400, change departure runway 28R, taxi to and hold short of 28L, advise when ready.

ET
- verify calibration
- eye mark #1 — when runway clearance complete

Obs/ET
- capture data from response to clearance

TP
- use Before Takeoff Checklist

ATC
- after pilots advise they are ready, give clearance to takeoff:
Airline A/B 400, cross runway 28L, cleared for takeoff runway 28R.

****
Subject starts down runway

ET
- eye mark #2 — throttles forward to start takeoff and a/t mode appears
- eye mark #3 — wheels off runway

ATC
- when airborne, say:
Airline A/B 400, contact departure.

****
Plane reaches 1500 ft altitude

ATC
- when altitude reaches 1500, give clearance to expedite to 4000
Airline A/B 400, Bay departure, expedite climb to cross the SFO 6 DME at 4,000, then resume PORTE3 with restrictions.

Obs
- capture data from response to clearance

****
Plane passes through D6 (ideally at 4000)

Obs
- capture data from subject action to restore VNAV and 9000 restriction

ATC
- <u>MAY</u> need to remind pilot to make 9000 ft restriction at PORTE. The plane can stay at 4000 through the turn and doesn t really have to be prompted on the 9000 restriction until it straightens out. At that point, if needed, ATC should use the following:
Airline A/B 400, will you be able to make PORTE at 9000?
- Or, if subject sets 35,000 and starts screaming up prior to PORTE, wait until he crosses 8000 and say:
Airline A/B 400, will you be able to make PORTE at 9000?

ATC
- prior to PORTE, say:
Airline A/B 400, Bay departure, contact Oakland Center 132.95
- after check in, say:
Airline A/B 400, Oakland Center, radar contact, report leaving 9,000

****
Plane passes through PORTE and transitions to VNAV SPD

Mgr/Cab Ops
- within 60 secs: have Cab Ops make the manipulation to force pitch mode to VNAV PTH

ET
- eye mark #4 — when VNAV changes to VNAV PTH

****
Plane reaches FL180

ATC
- when plane reaches FL180, say
Airline A/B 400, contact Oakland Center 133.7

****
Plane reaches FL200

ATC
- when altitude reaches 20,000, give clearance to hold FL210; heading 150:
Airline A/B 400, maintain FL 210 for traffic and turn right heading 150.

Obs
- capture data from response to clearance

****
Plane takes heading 150 and wings level

ATC
- when plane comes level at 210 and on heading 150, say
Airline A/B 400, Eastbound traffic, FL220, at 2 o clock.

Mgr/Cab Ops
- send visual target through

****

Plane comes abeam WAGES

ATC
- when tip of triangle just passes WAGES (and WAGES still active), give clearance to resume flight plan; heading 090:
Airline A/B 400, now turn left heading 090, intercept the AVENAL transition, on course. Continue climb to FL 350.

ET
- eye mark #5 — when this clearance complete

Obs
- capture data from response to clearance

****
Plane reaches 31,500

ATC
- when plane reaches 31,500, say
Airline A/B 400, Oakland Center, maintain FL 330.

****
Plane levels at 33,000

ET
- eye mark #6 — when plane levels at 33,000

****
Subject calls ATC regarding cruise altitude

ATC
- when subject requests, give clearance to use 330 as cruise:
Airline A/B 400, your request has been forwarded to Los Angeles Center. Contact them now on 133.05.
- after check in, say:
Airline A/B 400, Los Angeles Center, radar contact. It looks like FL 330 will be your final altitude today.

- if pilot does NOT call in by the time he is <u>40 nm before AVENAL</u>, say
Airline A/B 400, Oakland Center, contact Los Angeles Center 133.05.
- after check in, say
Airline A/B 400, Los Angeles Center, radar contact, use FL330 as your final altitude today.

ET
- eye mark #7 — when 330 in FMC, and execute

Obs
- capture data from response to clearance

****

Some point early in cruise

TP:
- request LAX ATIS as a prompt to initiate approach planning

ATC:
when ATIS request arrives, say
Los Angeles Information Echo: Los Angeles weather 1955 Zulu; wind 240 variable 280 at 5 to 10 knots, visibility 5 miles, scattered at 4,000, temperature 16, dew point 10, altimeter 29.92, ILS and visual approaches to 25L are in use, inform ATC on initial contact you have information echo.

****
Plane reaches point 40 nm before AVENAL

ATC
- give clearance for restrictions at SYMON:
Airline A/B 400, Los Angeles Center,
(wait for reply, and then continue)
after AVENAL, go direct to DERBB intersection. You are cleared to LAX, SADDE6, DERBB transition. Cross SYMON at 12,000, 280 knots, then resume the STAR. Report leaving FL330.

Obs
- capture data from response to clearance

****
Plane is 15 nm prior to DERBB

ATC
- give clearance to slow to 260:
Airline A/B 400, Los Angeles, slow to 260 and maintain that airspeed.

Obs
- capture data from response to clearance

****
Plane reaches T/D point

ET
- eye mark #8 — when plane reaches T/D and path deviation indicator appears

Obs
- capture data on subject behavior

TP
- encourages subject to regain VNAV PTH

****

After T/D, and 20 dme REYES

ATC
- at 20 dme REYES, give clearance to resume speed:
Airline A/B 400, Los Angeles, resume normal speed

****

Plane achieves VNAV PTH

Mgr/Cab Ops
- wait 1-2 minutes, then have Cab Ops make the manipulation to force pitch mode to VNAV SPD

ET
- eye mark #9 — when pitch mode goes to VNAV SPD

Obs
- capture data from subject s response to change

****

Halfway between REYES and FILLMORE

Mgr/Cab Ops
- have Cab Ops make the manipulation to force a/t mode to THR

ET
- eye mark #10 — when a/t mode goes to THR

Obs
- capture data from subject s response to change

****

Plane reaches FL190 OR 15 dme FILLMORE (whichever first)

ATC
- when passing FL 190, say
Airline A/B 400, Contact Los Angeles Center on 127.9.
- when check in, say
Airline A/B 400, Los Angeles Center, radar contact.

Mgr/Cab Ops
- change visibility at LAX to 5 miles

****

Plane passes FILLMORE

ATC
- at FILLMORE, say
Airline A/B 400, Los Angeles Center, now contact SOCAL approach on 124.5.
- at check in, say

Airline A/B 400, SOCAL approach, radar contact. Cross BAYST at 10,000, then descend to cross Santa Monica between 7 and 8,000, descend to and maintain 3,500.

****

Plane reaches Santa Monica

Mgr/Cab Ops
- fail G/S if not already failed
- change range on ATC station to 20 nm

ATC
- at Santa Monica, say

*Airline A/B 400, SOCAL approach, fly to heading 070, maintain 3,500, contact Approach on 128.*5.
- at check in, say

Airline A/B 400, Approach, radar contact. Maintain 3,500, slow to 170.

ET
- eye mark #11 — when clearance given for vectors is complete

****

Plane comes abeam Hunda

ATC:
- just after abeam Hunda intersection, say

Airline A/B 400, Approach. Turn right heading 160.
- shortly after last clearance, say

Airline A/B 400. Turn right heading 200, intercept the localizer, you are cleared for the ILS approach, runway 25L. Contact tower at Limma, 120.95.

****

Plane turns onto base leg

ET
- eye mark #12 — when plane turns onto base leg

Obs
- capture data from subject response to loss of G/S diamond

****

At some point, call to ATC about G/S

TP
- request information from ATC on status of G/S

ATC
- when responding, say something like

Airline A/B 400, Approach, we ve determined that the G/S has been damaged. You are cleared for a LOC only approach, 25L.

****

Plane reaches Limma

ATC
- at Limma, say
Airline A/B 400, Los Angeles tower. Cleared to runway 25L, wind 270 at 8 knots.

****
Plane turns onto final leg

ET
- eye mark #13 — when plane turns onto final

Mgr/Cab Ops
- if trouble on landing, clear weather

****
Plane wheels touch down

ET
- eye mark #14 — when wheels touch down

Mgr/Cab Ops
- sim data goes to stop

ET
- stop eye tracker, video off
- eye tracker head gear off subject

Appendix D
Simulator Variables Recorded

Time (seconds)

Autothrottle Modes *(FMA)*
THR engaged
THR REF engaged
HOLD engaged
IDLE engaged
SPD engaged

Roll Modes *(FMA)*
HDG HOLD engaged
HDG SEL engaged
LNAV armed
LNAV engaged
LOC armed
LOC engaged
ROLLOUT armed
ROLLOUT engaged
ATT engaged
GS engaged
FLARE armed
FLARE engaged

Pitch Modes *(FMA)*
TOGA engaged
ALT engaged
VS engaged
FLCH engaged
VNAV armed
VNAV PTH engaged
VNAV SPD engaged
VNAV ALT engaged
GS armed

Autothrottle Modes
THR button pushed
SPD button pushed

Roll Modes
LNAV button pushed
HDG HOLD button pushed
LOC button pushed

Pitch Modes
VNAV button pushed
FLCH button pushed
ALT hold button pushed
V/S button pushed
APP button pushed

AFDS Status *(FMA)*
FDs engaged
Autopilot engaged
LAND2 engaged
LAND3 engaged
No Autoland

MCP Window Values
Heading setting (deg; ±180)
IAS/Mach setting (kts)
Vertical speed setting (fpm)
Altitude setting (feet)

Flight Parameters
indicated airspeed (kts)
barometric altitude (ft)
radio altitude (ft)
magnetic hdg (deg; ±180)
lattitude (deg)
longitude (deg)
vertical path deviation (ft)
g-load (g; -1g norm)
LOC error (deg)
GS error (deg)

Flight Control Inputs
control wheel pos (deg)
control column pos (deg)
rudder position (deg)
throttle 1 position (deg)
throttle 2 position (deg)
throttle 3 position (deg)
throttle 4 position (deg)
thrust setting
flap position

Dummy Variables
overriding VNAV (variable = 1 when VNAV override used)
force to VNAV mode (variable = 1 for force to VNAV PATH; =2, for force to VNAV SPD)
force autothrottle to THR
fail GS indication

Appendix E
Mental Model Test

The following lists the questions asked of each subject and then presents the elements of a complete answer to each question:

1. What are VNAV PTH, VNAV SPD, and VNAV ALT?

- VNAV PTH is the VNAV mode that means the FMC is flying to an FMC-calculated geographic fixed path. You will see PTH on climb when there is an FMC altitude constraint tied to a waypoint. VNAV will transition from VNAV SPD to VNAV PTH as it starts to capture the altitude (since the FMC altitude constraint is tied to a geographic location), and then will transition back to VNAV SPD after passing the waypoint and resuming climb. You will also see VNAV PTH when you are at the FMC cruise altitude, and when you are on the FMC-created VNAV descent path. VNAV PTH flies altitude/path on elevator and airspeed on autothrottle (SPD mode).
- VNAV SPD is used when there is no fixed path e.g., climbing to cruise altitude. In this case, VNAV takes an airspeed target from the FMC and flies speed on elevator. VNAV SPD can also occur in descent if you deviate from the FMC path sufficiently, or if you invoke speed intervention by opening the MCP speed window.
- VNAV ALT occurs when the MCP altitude is used to level the airplane. For example, on climb, if you set 9000 ft in the MCP altitude window (and it s not in the FMC), VNAV will transition from VNAV SPD to VNAV ALT as it starts to capture the altitude and then will transition back to VNAV SPD after you have set a new MCP altitude and started climbing. Note that when you are using the MCP altitude window to level off as described here, you need to push the altitude selector to initiate climb after setting the new MCP altitude. VNAV ALT, like VNAV PTH, flies altitude/path on elevator.

2. There are automatic transitions from one mode to another; for example, VNAV PTH can transition to VNAV SPD. What conditions will trigger these automatic transitions?

- VNAV PTH to VNAV SPD — Two primary situations this transition can occur if you leave the VNAV descent path and revert to VNAV SPD. Also, if on climb you have leveled at an FMC altitude constraint, and then resume climbing. There are some other cases as well e.g., speed intervention on descent will transition you to VNAV SPD.
- VNAV SPD to VNAV PTH — Two primary situations: climbing in VNAV SPD, and then leveling at either an FMC intermediate level-off, or at the FMC cruise altitude. Or, recapturing the VNAV descent path on descent will transition you back to VNAV PTH.
- VNAV PTH to VNAV ALT — if you are descending in VNAV PTH and use altitude intervention to hold at an MCP altitude, you will transition to VNAV ALT.
- VNAV ALT to VNAV PTH — the first interesting case is the one from the scenario where you set 33000 in the MCP to level off lower than cruise, and then put 330 in the FMC to redefine CRZ altitude. As described above, pushing the MCP altitude selector transitions you back to VNAV PTH. Also, if you level off on descent in VNAV using the MCP altitude, you will be in VNAV ALT. If you were below the path and then intercept it and start down, you will transition to VNAV PTH.

3. On descent, how does VNAV assign priority to speed and path? That is, which one will it give up first?

VNAV will fly to both path and airspeed targets on descent. Initially, it will give up airspeed (e.g., if there are unanticipated winds) to preserve the descent path. However, when the airspeed deviates by more than 15 kts and you are more than 150 ft from the path, VNAV gives up the path and transitions to VNAV SPD. Automation will protect the path tracking if you are low or slow, by calling for thrust automatically (if the autothrottle is on). If, however, you are high or fast, the FMC will display the scratchpad message DRAG REQUIRED. If drag is not provided by the pilot (speedbrakes) or drag authority is insufficient, VNAV will leave the path (reverting to VNAV SPD) and can only be returned by pilot intervention.

4. Is there any point in time when it is important to have one of these modes active? If so, which one at what time?

Generally, it is good to ensure that you have VNAV PTH on cruise and descent. In this mode, VNAV will start down at T/D (assuming you have dialed down the MCP altitude), and you will meet the speed and altitude constraints on descent.

5. Speed Protection - by speed protection, I am referring to the case where the automation prevents you from exceeding maximum speeds (e.g., Vfe) or falling below a safe airspeed. When is this type of protection available?

Speed protection is available in VNAV and FLCH. It is not available if you are in V/S pitch mode or if the autothrottle is not engaged. Because of the loss of speed protection in V/S, you need to be cautious about using V/S (e.g., in situations where airspeed can bleed off below stall speed, such as climbing to a new cruise altitude).

6. If you set an airspeed in the FMC on the CRZ page, does that airspeed propagate to the DES page?

No, you have to also program descent speeds on the DES page. If you don t, airspeed will transition to Econ speed at the T/D point and will fly the descent path based on the original Econ descent airspeed.

7. Pitch-Autothrottle Combinations: What autothrottle modes occur at what times?

Generally,
on take-off and climb you will see THR REF and HOLD, maybe SPD
on cruise you will see SPD and HOLD
on descent, you will see IDLE, HOLD, and SPD

We were trying to see if you thought THR would occur in descent

8. What are the requirements for LNAV capture?

Generally, the airplane must be pointed at the active leg (inbound leg to the active waypoint in line 1-left on the ACT RTE LEGS page). If the airplane is not pointed at the active leg, the scratchpad message "not on intercept heading" is displayed. LNAV will capture the active leg (roll flight mode annunciation changes to LNAV green) whenever the cross track error is less than 2.5 NM. It will use a greater distance for steep intercept angles, in which case, LNAV engages at a distance from the magenta line where a turn at 22.5 degrees bank angle will result in the airplane rolling out on the magenta line inbound to the active waypoint. The distance at which LNAV will engage depends on a number of factors including wind, groundspeed and course intercept geometry.

9. When you ve put an altitude on the MCP, under what conditions do you need to push the altitude selector?

It seems that most pilots push the MCP altitude selector more often then they need to. One item to remember.

- The airplane will generally* not fly through the MCP altitude. So, if you are climbing at 19,500 and asked to level at 21,000. You can merely set the MCP altitude to 21,000. No button push is required to level at 21,000. In this particular case, a button push causes no other changes (assuming there is no FMC altitude constraint between you and 21,000); it is just an unnecessary action.
(*Note that the airplane can fly AWAY from the MCP altitude e.g., in V/S and that the airplane will fly through the MCP altitude when the glideslope is captured.)

Here are some general rules (for VNAV altitude intervention).

When to press the MCP altitude selector:
When you are using altitude intervention, each time you press the MCP altitude selector you remove a restriction, if the restriction is your current (captured) altitude or it is an FMC restriction between your current altitude and the MCP altitude window target you are flying toward. Sometimes the restriction is an FMC altitude restriction at a waypoint; sometimes it s a cruise altitude restriction; and, sometimes it is an MCP altitude restriction that you have captured.

When not to press the MCP altitude selector:
The simplest way to state this rule is if you get the MCP altitude target out of your way, the FMC will make altitude transitions on its own. Stated more completely, if you are in VNAV, and the MCP altitude target is not your current altitude (i.e., it s above you when climbing or below you when descending), FMC will make altitude transitions on its own (unless you pass the constrained waypoint or FMC T/D point). This last point is important: if you have passed the constrained waypoint or FMC T/D point and maintained that altitude by keeping the MCP altitude set at your current altitude, then you must push the MCP altitude selector after setting the new altitude to command the climb or descent.

Appendix F
Tables from the Results Section

Table 3-1. Subject response to reversion from 9000 to 9000A

	S notices + changes 9000A	alt at PORTE
S4:	N	9000 V-A
S5:	N	8800 V-A
S6:	N	9000 V-A
S8:	N	9000 V-A
S9:	N	8750 ALT
S14:	N	9000 V-A
S17:	N	9000 V-A (TP saved him from dialing FL350 on MCP)
S19:	N	9000 V-A
S20:	N	9000 V-A
S11:	N	8300 V-A
S22:	N	8550 V-A
S15:	notices (after TO); no change	8500 V/S
S13:	notices (Nav-after TO); no chnge	8800 V-A
S16:	notices (before TO); no change	9000 V-A
S23:	notices (after TO); no change	9000 V-A
S7:	Y (Nav display-after TO)	9000 V-P
S10:	Y (anticipated-before TO)	9000 V-P
S12:	Y (LEGS-before TO)	9000 V-P
S18:	Y (LEGS-before TO)	9000 V-P
S21:	Y (Nav display-after TO)	9000 V-A

Legend
The first column shows each subject s response:
- Y (yes) means that the subject both noticed the reversion and then corrected it. For these subjects, we indicate the (likely) source of their noticing and whether it was noticed before or after take-off (TO).
- N (no) means that the reversion was neither noticed nor corrected; these subjects maintained 9000A in the FMC. N is coded as an incorrect response.
- some subjects noticed the reversion but did not change it.

The second column shows the altitude at which the subject crossed PORTE and the pitch mode at the time of crossing. Any altitude that is 300 or more feet below 9000 is coded as an unsafe performance. The pitch modes are:
- V-A for VNAV ALT
- V-P for VNAV PTH
- V/S for vertical speed mode
- ALT for altitude hold mode

Table 3-2. Subject response to AVE-DERBB discontinuity

S notices + closes discntnty		when noticed	likely display used
S4:	Y	before TO	anticipated
S9:	Y	before TO	anticipated (Nav)
S18:	Y	before TO	anticipated
S15:	Y	b/w PORTE +WAGES	LEGS
S19:	Y	b/w PORTE +WAGES	LEGS
S22:	Y	b/w PORTE + WAGES	LEGS
S5:	Y	b/w WAGES + AVE (em5+2:18)	LEGS
S6:	Y	b/w WAGES + AVE (em5+4:04)	Nav
S7:	Y	b/w WAGES + AVE (em5+1:28)	Nav
S8:	Y	b/w WAGES + AVE (em5+1:53)	Nav
S10:	Y	b/w WAGES + AVE (em5-0:36)	LEGS
S11:	Y	b/w WAGES + AVE (em7+4:46)	clearance for DES
S12:	Y	b/w WAGES + AVE (em7+3:31)	clearance for DES(Nav)
S13:	Y	b/w WAGES + AVE (em5+4:29)	Nav
S14:	Y	b/w WAGES + AVE (em5+0:31)	LEGS
S16:	Y	b/w WAGES + AVE (em5+6:03)	LEGS
S17:	Y	b/w WAGES + AVE (em5+4:26)	Nav
S20:	Y	b/w WAGES + AVE (em5+4:15)	LEGS
S23:	Y	b/w WAGES + AVE (em5+1:22)	Nav
S21:	-- (discontinuity got closed but can t tell how)		

Legend

The first column shows each subject s response:
- Y (yes) means that the subject both noticed the discontinuity and then closed it. Data is missing for one subject.

The second column shows when each subject first mentioned the discontinuity. There were three responses
- before take-off
- between PORTE and WAGES waypoints
- between WAGES and AVE waypoints

each of the latter two responses is also related to elapsed time after one of the eye marks (em5 or em7)

The third column shows the interface element that we believe revealed the discontinuity (this judgment was made from viewing the fixation crosshairs on the videotape).

Table 3-3. Subject response to clearance to expedite climb to 4000

	S strategy early	late (usually after 3000)	alt at D6
S4:	nothing		3940
S17:	nothing		**3700**
S5:	nothing	spd intrvene	3800
S11:	nothing	spd intrvene: (250)	3850
S14:	nothing	spd intrvene	4000
S18:	nothing	spd intrvene	4000
S19:	nothing	spd intrvene: (188)	3920
S9:	FLCH-keeps speed low (183)		3940
S23:	FLCH-keeps speed low (180 250)		3960
S15:	V/S	spd intrvene+thrust increase to CLB	4000
S13:	**spd intrvene (270)**		**3550**
S16:	spd intrvene (245 230)		**3600**
S6:	thrust increase to CLB		4000
S7:	thrust increase to CLB		4000
S8:	thrust increase to CLB		4120
S10:	thrust increase to CLB		3900
S12:	thrust increase to CLB		4000
S20:	thrust increase to CLB	spd intrvene: (154)	4000
S21:	thrust increase to CLB		4000
S22:	thrust increase to CLB	spd intrvene: (245)	4000

Legend

The first column shows the control strategy that each subject took when the clearance was first given. Several actions were taken:

- thrust increase to CLB means that the subject selected full climb power instead of the derated climb thrust that was used for take-off.
- spd intrvene (speed intervention) was used to keep speed low (speed target in parentheses). Otherwise, the automation transfers power from climbing to increasing speed to 250 kts. S13 s response is coded as incorrect since he violated the 250 kt maximum airspeed below 10,000 ft.
- some subjects engaged FLCH pitch mode for more direct control of airspeed.
- one subject engaged V/S pitch mode to control climb rate more directly.
- nothing - some subjects took no action at all.

The second column shows that some subjects re-assessed the situation later and took action (or an additional action).

The third column shows the altitude at which the subject crossed D6. Any altitude that is 300 or more feet below 4000 is coded as an unsafe performance.

Table 3-4. Subjects ability to meet restriction at PORTE

	S leaves 4000 promptly	alt at PORTE
S5:	**N** (S; late 1:12)	8800
S16:	**N** (S; late 0:58)	9000
S9:	**N** (ATC; late 1:30)	8750
S11:	**N** (ATC; late 1:43)	**8300**
S13:	**N** (ATC; late 1:27)	8800
S15:	**N** (ATC; late 1:47)	**8500**
S22:	**N** (ATC; late 1:51)	**8550**
S4:	Y	9000
S6:	Y	9000
S7:	Y	9000
S8:	Y	9000
S10:	Y	9000
S12:	Y	9000
S14:	Y	9000
S17:	Y	9000
S18:	Y	9000
S19:	Y	9000
S20:	Y	9000
S21:	Y	9000
S23:	Y	9000

Legend
The first column shows whether each subject started climbing promptly after crossing D6.
- Y (yes) means that the subject did leave 4000 promptly.
- N (no) means that the subject waited to start climbing and is coded as an inappropriate response since the subject was cleared to start up immediately. For those who did not start climbing, we indicate what stimulus prompted them to start climbing (either a call from ATC or the subject remembered on his own), and we indicate the latency from crossing D6.

The second column shows the altitude at which the subject crossed PORTE. Any altitude that is 300 or more feet below 9000 is coded as an unsafe performance.

Table 3-5. Subject response to the artificial VNAV PTH pitch mode

	alt	(em4) V PTH time	reversion	total time	alt	detected?	#fixations	MM
S7:	11,4	12:06:41	12:09:38	2:57	20,7	N	--	Y
S20:	10,6	11:38:33	11:41:54	3:21	20,7	N	--	Y
S5:	10,7	12:21:15	12:24:56	3:41	20,7	N	0	--
S15:	10,3	11:37:51	11:41:20	3:29	20,7	N	0	Y
S19:	13,2	11:01:18	11:03:34	2:16	20,7	N	0	Y
S21:	10,2	13:32:12	13:35:48	3:36	20,7	N	0	Y
S4:	11,1	12:00:54	12:04:10	3:16	20,7	N	5	Y
S6:	10,6	11:04:27	11:07:59	3:32	20,7	N	2	Y
S8:	11,6	12:22:36	12:25:34	2:58	20,7	N	4	--
S10:	11,5	2:08:50	2:11:50	3:00	20,7	N	3	Y
S11:	11,2	10:50:59	10:54:04	3:05	20,7	N	3	--
S12:	10,9	11:33:16	11:36:45	3:29	20,7	N	6	Y
S13:	10,8	11:22:36	11:26:01	3:25	20,7	N	7	Y
S14:	10,5	11:24:42	11:28:20	3:38	20,7	N	1	--
S17:	10,6	3:17:53	3:21:39	3:46	20,7	N	1	--
S18:	10,4	11:11:40	11:15:23	3:43	20,7	N	2	Y
S22:	10,4	14:32:08	14:35:43	3:35	20,7	N	3	Y
S23:	11,0	11:13:14	11:16:28	3:14	20,7	N	1	--
S9:	10,7	11:40:37	11:40:48	--	--	--	--	Y
S16:	10,8	11:42:25	11:42:55	--	--	--	--	--

Legend
The first column shows the altitude (in100s) at which the pitch mode was changed to VNAV PTH.

The second column shows the (videotape) time when the pitch mode was changed to VNAV PTH.

The third column shows the time when the pitch mode was restored to the actual mode. Typically, this occurred when the subject started to level at FL210 for the clearance.

The fourth column shows the total time that the artificial pitch mode was displayed.

The fifth column shows the altitude (in 100s) at which the pitch mode was restored to the actual mode.

The sixth column shows whether the subject noticed and commented on or took action on the artificial pitch mode. An N (no) response was coded as incorrect.

The seventh column shows the number of times the subject fixated the pitch FMA during the time the artificial pitch mode was displayed. (Note: in 4 cases, marked -- , eye fixation data was unavailable)

The last column shows whether the subject, during the mental model interview, indicated that VNAV SPD should be annunciated during climb. Y (yes) means the subject made this statement.

Table 3-6. Subject response to the clearance to resume flight plan after WAGES

	strategy to resume	timing to arm LNAV	active wpt when arm
S4:	hdg and alt	immediate (12:07:00)	**INTC**
S6:	1st - hdg and alt	immediate (11:11:10)	**INTC**
	2nd - 118 int to AVE	after 118 int (11:13:18)	AVE
S7:	hdg and alt	immediate (12:12:53)	**WAGES**
S8:	1st - hdg and alt	soon (12:29:54)	**WAGES**
	2nd - 118 int to AVE	after 118 int	AVE
S9:	hdg and alt	immediate (11:47:09)	**INTC**
S11:	hdg and alt	immediate (10:57:02)	**INTC**
S14:	hdg and alt	immediate (11:31:32)	AVE
S16:	hdg and alt	immediate (11:49:00)	**WAGES**
S17:	hdg and alt	waits (talking) (3:25:37)	AVE
S18:	hdg and alt	immediate (11:18:35)	**INTC**
S19:	hdg and alt	waits (11:06:46)	AVE
S20:	hdg and alt	immediate (11:45:00)	**INTC**
S21:	hdg and alt	waits for turn to 90 (13:39:28)	**INTC**
S10:	118 int to AVE	after 118 int (2:15:11)	AVE
S12:	118 int to AVE	after 118 int (11:40:12)	AVE
S13:	118 int to AVE	--	-- (eye track distraction)
S15:	118 int to AVE	immediate (11:47:21)	AVE
S22:	118 int to AVE (twice)	after 118 int (14:40:19)	AVE
S23:	118 int to AVE	after 118 int (11:20:01)	AVE
S5:	direct to AVE (mistake)	after AVE set up (12:27:51)	AVE

Legend

The first column shows how each subject responded to the clearance

- 118 int to AVE means that the subject set up (through the CDU) a route directly into AVE
- hdg and alt means that the subject simply set the new heading assigned (090) and the altitude (FL350). Note that 2 subjects armed LNAV and then decided to set up a course to AVE and re-arm LNAV.
- in one case, ATC mistakenly allowed a subject to do a direct AVE, which is different from the clearance initially given.

The second column shows each subject s timing on arming LNAV and the (videotape) time LNAV was armed.

The third column shows what waypoint was active when LNAV was armed. The pilot should have AVE as the active waypoint when he arms LNAV. Other waypoints are coded as incorrect.

Table 3-7. Subject response to the clearance to change cruise altitude to FL330

	330 entered	which CDU page	outcome	MM (1,2)	fixation
S5:	after clrnc	CLB	**V-ALT**	-,-	--
S6:	before clrnc	CLB	**V-ALT**	Y,-	Y
S7:	before clrnc	CLB	**V-ALT**	Y,Y	--
S9:	after clrnc	PERF INIT, CLB	**V-ALT**	Y,-	--
S10:	before clrnc	CLB	**V-ALT**	Y,-	Y
S11:	before clrnc	CLB	**V-ALT**	Y,-	N
S12:	after clrnc	CLB	**V-ALT**	Y,-	--
S13:	before clrnc	CLB	**V-ALT** (mentions prior to T/D)	Y,-	Y
S14:	never entered	--	**V-ALT**	**x**,-	Y
S15:	before clrnc	CLB	**V-ALT**	Y,-	Y
S16:	after clrnc	CLB	**V-ALT**	-,-	Y
S17:	after clrnc	CLB	**V-ALT**	Y,-	Y
S18:	after clrnc	CLB	**V-ALT**	-,Y	Y
S19:	after clrnc	CLB	**V-ALT**	Y,-	N
S20:	after clrnc	CLB	**V-ALT**	-,-	--
S22:	before clrnc	CLB	**V-ALT** (mentions but no follow-up)	Y,Y	Y
S4:	after clrnc	CLB	V-PTH (MCP alt knob)	Y,Y	--
S8:	after clrnc	CLB	V-PTH (MCP alt knob)	Y,-	Y
S21:	after clrnc	CLB	V-PTH (cycle VNAV; TP prompt)	Y,Y	Y
S23:	after clrnc	CLB	V-PTH (cycle VNAV)	Y,Y	Y

Legend
The first column shows when the subject put the FL330 into the FMC as the new cruise altitude. Some subjects did this after ATC gave them the clearance that FL330 was the assigned cruise altitude; and some subjects took this action before they got ATC clearance (because it was anticipated).

The second column shows the CDU page that the subject used to enter the new cruise altitude.

The third column shows the pitch mode that resulted and was flown for the cruise phase. VNAV ALT is coded as an incorrect outcome.

The fourth column shows how subjects responded to mental model prompts about VNAV PTH on cruise. There were two opportunities. First, a Y (yes) indicates that when asked about when you will generally see VNAV PTH mode, the subject said during cruise (see Table 3-18). Second, a Y indicates that when asked about when certain VNAV modes are important, the subject said that it is important to have VNAV PTH in cruise (see Table 3-26).

The last column shows whether the subject fixated the pitch FMA during the cruise phase.

Table 3-8. Subject initiation of descent

	mode-T/D	action to initiate descent	DES outcome	V-P alt
S6:	V-ALT	**alt knob-1:05 late**	VNAV SPD (spd int)-260	24,6
S9:	V-ALT	**alt knob-0:11 late**	VNAV SPD (spd int)-260	23,7
S11:	V-ALT	**alt knob-0:21 late**	VNAV SPD (spd int)-260	29,8
S12:	V-ALT	**tries DES NOW early; then alt knob-0:32 late**	VNAV SPD (spd int FMC)-260	23,7
S13:	V-ALT	**alt knob-0:19 late** (TP)	VNAV SPD (spd int)-260	29,4
S14:	V-ALT	**dial 12; alt knob-1:16 late**	VNAV SPD (spd int)-260	--
S15:	V-ALT	**tries DES NOW early; V/S-1:46 late**	V/S -260	27,4
S16:	V-ALT	**V/S-0:24 late**	V/S FLCH -260	21,6
S17:	V-ALT	**FLCH-0:52 late**	FLCH -260	22,7
S18:	V-ALT	**alt knob-0:13 late**	VNAV SPD (spd int)-260	29,4
S22:	V-ALT	**alt knob-0:24 late****	VNAV SPD (spd int)-260	29,7
S5:	V-ALT	early with alt knob	VNAV SPD (spd int)-260	25,1
S7:	V-ALT	early with alt knob	VNAV SPD (spd int)-260	28,9
S10:	V-ALT	early with alt knob	VNAV SPD (below path-FMC)-260	24,8
S19:	V-ALT	early with alt knob	VNAV SPD (spd int)-260	18,0
S20:	V-ALT	none*	VNAV PTH (FMC)-**282**	29,5
S4:	V-PTH	none	VNAV SPD (spd int)-260	19,6
S8:	V-PTH	none	VNAV SPD (spd int)-260	25,1
S21:	V-PTH	none	VNAV PTH (FMC)-**282**	29,1
S23:	V-PTH	early with alt knob	VNAV SPD (spd int)-260	21,9

* for some unknown reason, the simulator started down at T/D even though the subject was in VNAV ALT. Videotape analysis was unable to account for this.
**for this subject, we were doing eye tracker calibration just prior to T/D, and the S was unable to monitor at T/D. We completed calibration just after crossing T/D.

Legend
The first column shows the pitch mode during cruise as the plane approached T/D.
- V-ALT means VNAV ALT
- V-PTH means VNAV PTH

The second column shows the action the subject took to initiate descent.
- none means no action was needed to initiate descent.
- early with alt knob means that the subject initiated descent early by pushing the MCP altitude knob with 12,000 on the MCP altitude window.
- for all other cases, the subject started down late, sometimes using the MCP altitude knob, and sometimes using V/S or FLCH mode. Not starting down at T/D is coded as incorrect.

The third column shows the pitch mode and airspeed at which the plane started descending. The airspeed of 282 is coded as incorrect since the clearance was for 260.

The fourth column shows the altitude at which the subject achieved a stable VNAV PTH mode (after receiving clearance to resume a descent speed of 282).

Table 3-9. Subject management of airspeed restrictions

	spd intrvn?	CRZ pg?	DES pg?	CRZ trgt?	DES trgt?	could ve closed spd wndw	MM
S9:	Y	Y	Y	MCP	MCP	**on CRZ, DES**	?
S4:	Y	N	Y	MCP	MCP	**on DES**	Y
S5:	Y	N	Y	MCP	MCP	**on DES**	Y
S7:	Y	N	Y (DES)	MCP	MCP	**on DES**	Y
S8:	Y	N	Y	MCP	MCP	**on DES**	Y
S12:	Y	N	Y	MCP	MCP, FMC	--	Y
S20:	Y (closes)	Y	N	FMC	MCP	--	Y
S6:	Y	N	N	MCP	MCP	--	x
S11:	Y	N	N	MCP	MCP	--	Y
S13:	Y	N	N	MCP	MCP	--	Y
S14:	Y	N	N	MCP	MCP	--	Y
S16:	Y	N	N	MCP	MCP	--	Y
S17:	Y	N	N	MCP	MCP	--	Y
S18:	Y	N	N	MCP	MCP	--	x
S19:	Y	N	N	MCP	MCP	--	Y
S22:	Y	N	N	MCP	MCP	--	Y
S23:	Y	N	N	MCP	MCP	--	Y
S10:	Y (closes)	Y	Y	FMC	FMC	--	Y
S15:	Y (closes)	Y	Y	FMC	FMC	--	Y
S21:	Y (closes)*	Y	Y (DES)	FMC	MCP	--	Y

* TP helped S see mistake of forgetting to close speed window.

Legend
The first column shows whether the subject used speed intervention to set 260 kts in the MCP airspeed window. If (closes) is shown, then the subject closed the airspeed window after putting 260 kts in the FMC.

The second column shows whether the subject entered 260 kts as the selected speed on the CRZ page.

The third column shows whether the subject entered 260 kts as a selected speed on the DES page. If (DES) is shown, then the subject waited until he got to the DES phase before entering a speed. Otherwise, the entry was made during cruise.

The fourth column shows where the plane took its airspeed target during cruise (either the MCP airspeed window or the FMC CRZ page).

The fifth column shows where the plane took its airspeed target during descent (either the MCP airspeed window or the FMC DES page).

The sixth column shows cases in which the subject had entered a speed in the FMC but kept the MCP airspeed window open. That is, he set up the FMC so that the MCP airspeed window could be closed, but kept the window open anyway. These items are coded as incorrect because it results in VNAV SPD.

The last column shows how subjects responded to a mental model prompt regarding propagation of airspeeds from the CRZ page to the DES page. A Y (yes) indicates that the subject knew that the airspeed does NOT propagate. One subject didn t know (?) and two subjects gave incorrect accounts (x).

Table 3-10. Subject response to the artificial VNAV SPD pitch mode

	alt	(em9) V SPD time	reversion	total time	alt	detected?	#fixations	MM (1,2)
S7:	26,8	12:36:49	12:43:37	6:48	12,7	N	--	Y,Y
S9:	22,2	12:12:58	12:18:50	5:52	10,9	N	--	Y,-
S20:	28,4	12:08:17	12:17:42	9:25	10,6	N	--	Y,Y
S5:	23,2	12:53:10	12:58:37	5:27	12,1	N	0	Y,Y
S6:	22,4	11:36:30	11:41:35	5:05	11,8	N	0	x,-
S17:	21,7	3:50:21	3:54:34	4:13	12,7	N	0	Y,Y
S21:	28,0	14:02:05	14:06:48	4:43	18,0	N	0	-,-
S22:	27,6	15:02:29	15:09:15	6:46	13,1	N	0	Y,-
S23:	21,0	11:46:15	11:49:52	3:37	13,6	N	0	Y,-
S4:	18,0	12:35:23	12:37:44	2:21	13,4	N	1	Y,-
S8:	23,9	13:03:08	13:07:27	4:19	14,5	N	2	-,-
S10:	22,8	2:41:05	2:48:41	7:36	10,0	N	6	Y,Y
S11:	27,3	11:22:01	11:31:55	9:54	10,0	N	4	Y,-
S12:	22,7	12:05:46	12:11:04	5:18	12,1	N	2	-,-
S15:	25,9	12:08:43	12:12:35	3:52	17,6	N	2	Y,-
S16:	20,9	12:15:25	12:22:20	6:55	10,0	N	2	-,-
S18:	27,5	11:42:01	11:48:04	6:03	14,7	N	1	Y,Y
S19:	17,6	11:34:54	11:37:00	2:06	13,3	N	2	Y,Y
S13:	27,7	11:52:46	11:59:35	6:49	13,5	Y	10	Y,-
S14:	-- (we didn t alter mode)						--	Y,-

Legend

The first column shows the altitude (in 100s) at which the pitch mode was changed to VNAV SPD.

The second column shows the (videotape) time when the pitch mode was changed to VNAV SPD.

The third column shows the time when the pitch mode was restored to the actual mode. Often, this occurred when the subject started to level at 12,000 for the restriction at SYMON.

The fourth column shows the total time that the artificial pitch mode was displayed.

The fifth column shows the altitude (in 100s) at which the pitch mode was restored to the actual mode.

The sixth column shows whether the subject noticed and commented on or took action on the artificial pitch mode. An N (no) response was coded as incorrect.

The seventh column shows the number of times the subject fixated the pitch FMA during the time the artificial pitch mode was displayed. (Note: in 4 cases, marked -- , eye fixation data was unavailable)

The last column shows whether the subject, during the mental model interview, indicated that VNAV PTH should be annunciated during descent. There were two opportunities. First, a Y (yes) indicates that when asked about when you will generally see VNAV PTH mode, the subject said during descent (see Table 3-18). Second, a Y indicates that when asked about when certain VNAV modes are important, the subject said that it is important to have VNAV PTH in descent (see Table 3-26).

Table 3-11. Subject response to the artificial THR autothrottle mode

	alt	(em10) THR time	reversion	total time	alt	detected?	#fixations	MM
S7:	19,7	12:40:06	12:43:42	3:36	12,5	N	--	N
S9:	20,8	12:13:39	12:17:47	4:08	12,5	N	--	N
S20:	21,2	12:11:35	12:15:59	4:24	12,3	N	--	N
S4:	17,4	12:35:46	12:37:53	2:07	13,1	N	0	N
S6:	21,1	11:37:03	11:41:14	4:11	12,3	N	0	N
S15:	20,8	12:11:03	12:12:35	1:32	17,6	N	0	N
S16:	19,8	12:15:58	12:19:45	3:47	12,3	N	0	Y
S17:	21,0	3:50:39	3:54:38	3:59	12,5	N	0	N
S19:	16,8	11:35:16	11:37:06	1:50	13,1	N	0	N
S5:	21,5	12:53:58	12:58:20	4:22	12,5	N	3	N
S8:	22,8	13:03:39	13:07:14	3:35	14,5	N	1	N
S10:	19,0	2:42:48	2:46:05	3:17	12,3	N	1	N
S11:	20,4	11:25:12	11:29:12	4:00	12,5	N	5	N
S12:	21,6	12:06:15	12:10:44	4:29	12,5	N	6	N
S13:	20,4	11:56:06	11:59:42	3:36	13,2	N	7	N
S18:	20,4	11:45:14	11:48:15	3:01	14,3	N	1	N
S21:	21,0	14:05:19	14:06:56	1:37	17,6	N	1	N
S22:	20,2	15:05:53	15:09:15	3:22	13,1	N	3	N
S23:	20,4	11:46:29	11:49:57	3:28	13,4	N	7	N
S14:	-- (we didn t alter mode)						--	N

Legend
The first column shows the altitude (in 100s) at which the autothrottle mode was changed to THR.

The second column shows the (videotape) time when the autothrottle mode was changed to THR.

The third column shows the time when the autothrottle mode was restored to the actual mode. Typically, this occurred when the subject started to level at 12,000 for the restriction at SYMON.

The fourth column shows the total time that the artificial autothrottle mode was displayed.

The fifth column shows the altitude (in 100s) at which the autothrottle mode was restored to the actual mode.

The sixth column shows whether the subject noticed and commented on or took action on the artificial autothrottle mode. An N (no) response was coded as incorrect.

The seventh column shows the number of times the subject fixated the autothrottle FMA during the time the artificial mode was displayed.

The last column shows whether the subject, during the mental model interview, indicated that THR could be annunciated during descent. An N indicates that when asked which autothrottle modes can occur, the subject did NOT indicate that THR was a possibility (see Table 3-29). Many subjects indicated that they did not actually know which autothrottle modes were possible. One subject (#16) identified THR as a possible autothrottle mode on descent (although he incorrectly linked it to VNAV PTH).

Table 3-12. Subjects ability to meet altitude constraints on descent

	SYMON 12,000	BAYST 10,000	SMO 7-8,000
S8:	**12,400** (?)	10,000	OK (can t tell exact altitude)
S11:	12,000	10,000	**9100** (started down late; forgets to push alt knob)
S16:	11,950	10,000	**8350** (started down late; forgets to push alt knob)
S17:	12,000	10,000	**8400** (started down late; forgets to push alt knob)
S4:	11,940	10,000	7560
S5:	12,040	10,000	7360
S6:	12,000	10,000	7100
S7:	12,040	10,000	7400
S9:	11,950	10,000	7350
S10:	11,980	10,000	7300
S12:	12,040	10,000	7100
S13:	12,050	10,000	7600
S14:	12,000	10,000	7650
S15:	11,900	10,000	7150
S18:	12,050	10,000	7450
S19:	12,000	10,000	7200
S20:	11,950	10,000	7080
S21:	12,050	10,000	7550
S22:	12,050	10,000	8050
S23:	11,930	10,000	7960

Legend

The first column shows the altitude at which the subject crossed SYMON. The clearance was for 12,000 ft. Any altitude that is 300 or more feet above 12,000 is coded as an unsafe performance.

The second column shows the altitude at which the subject crossed BAYST. The clearance was for 10,000 ft. Any altitude that is 300 or more feet above 10,000 is coded as an unsafe performance.

The third column shows the altitude at which the subject crossed Santa Monica (SMO). The clearance was to cross between 7,000 and 8,000 ft. Any altitude that is 300 or more feet above 8,000 is coded as an unsafe performance.

Table 3-13. Subjects ability to meet altitude constraints on descent

	noticed	armed when cleared for APP?	armed LOC later?
S21:	**never** (ATC)	APP	**N**
S5:	before (TP)	APP	**N**
S6:	before	APP	**N**
S9:	before	APP	**N**
S13:	before	APP	**N**
S16:	before	LOC; then APP	**N**
S22:	before (TP)	APP	**N**
S4:	after (TP)	LOC; then APP	**N**
S7:	after	APP	**N**
S10:	after (TP)	APP	**N**
S15:	after	APP (TP)	**N**
S17:	after	APP	**N**
S12:	before	APP	Y
S14:	before	APP	Y
S19:	before	APP	Y
S20:	before	APP	Y
S11:	after (TP)	APP	Y
S18:	after	APP	Y
S8:	after	LOC	--
S23:	after	LOC	--

Legend
The first column shows when the subject commented that the glideslope was missing. We indicate that it was noticed either before or after intercepting the final leg. (TP) means that the TP prompted the S about the glideslope status. (ATC) means that ATC told the subject about it.

The second column shows what approach mode the subject armed when he was cleared for the approach (on base leg). The choices are to arm LOC only mode or the APP mode.

The third column shows whether the subject armed the LOC only mode when they were told that the glideslope was out of service. An N (no) is coded as an inappropriate response.

Table 3-14. Subjects management of automation on approach

	a/p shut off	a/t shut off	FD shut off
S4:	1200*	150*	--
S5:	780	700	--
S6:	1400	1100	--
S7:	3400	3000	--
S8:	840	800	400*
S9:	600	150	--
S10:	1900	1050	--
S11:	1500	500	--
S12:	500	**never**	--
S13:	1050	td	--
S14:	1200	**never**	--
S15:	950	950	--
S16:	1200	1000	--
S17:	2200	2100	--
S18:	600	350	--
S19:	1050	200	--
S20:	1000	750	550
S21:	2000	600	--
S22:	1200	15	--
S23:	900	250	--

*all values are in feet AGL.

Legend
The first column shows when the subject shut off the autopilot (and reverted to FD mode).

The second column shows if and when the subject shut off the autothrottle. td means touchdown. Not shutting off the autothrottle before touchdown (never) was coded as an inappropriate response.

The third column shows if and when the subject shut off the FD. -- means not shut off by the pilot.

Table 3-15. Subjects use of the MCP altitude knob: leveling on climb

	4000 push?	FL210 push?	FL330 push?
S11:	Y	Y	Y
S7:	N	YY	Y
S10:	N	Y	Y (Y-trying to change CRZ altitude?)
S12:	N	Y	Y (Y-trying to change CRZ altitude?)
S14:	N	Y	Y
S18:	N	Y	Y
S4:	N	N	Y
S6:	N	N	Y
S8:	N	N	Y
S9:	N	Y	N
S13:	N	N	Y
S16:	N	N	Y
S19:	N	N	Y
S22:	N	Y	N
S5:	N	N	N
S15:	N	N	N
S17:	N	N	N
S20:	N	N	N
S21:	N	N	N
S23:	N	N	N

Legend
The first column shows whether the subject pushed the MCP altitude knob after setting the altitude at 4,000. Since there is no need to push the altitude knob, a Y response is coded as incorrect.

The second column shows whether the subject pushed the MCP altitude knob after setting the altitude at FL210. Since there is no need to push the altitude knob, a Y response is coded as incorrect.

The third column shows whether the subject pushed the MCP altitude knob after setting the altitude at FL330. Since there is no need to push the altitude knob, a Y response is coded as incorrect.

Table 3-16. Subjects use of the MCP altitude knob: climbing from level

	4000 9000	9000 FL350	FL210 FL350
S4:	Y	Y	**N** (ATC-2:39 later)
S5:	**N** (S-1:12 later)	Y	Y
S8:	Y	**N** (S-0:07 later)	Y
S9:	**N** (ATC-1:30 later)	N (VNAV)	Y
S10:	Y	**Y** (not needed)	Y
S11:	Y*	Y	**Y** (but in ALT; S engages VNAV-0:01 later)
S12:	Y	Y**Y**	Y
S13:	--*(TP pushes)	**N** (TP-0:07 later)	Y
S22:	Y*	Y	**N** (ATC-2:08 later)
S23:	N (FMC altitude)	Y	**Y** (but in ALT; S engages VNAV-0:44 later)
S6:	Y	Y	Y
S7:	Y	Y	Y
S14:	Y	Y	Y*
S15:	N*(V/S)	N (V/S)	Y
S16:	Y*	Y	Y
S17:	Y	Y	Y
S18:	Y	Y	Y
S19:	Y	Y	Y
S20:	Y	Y	Y
S21:	Y	Y	Y

* in column 1 and 3 are cases where the S didn t understand (or forgot) the clearance and waited before setting the altitude and starting up. They did push the alt knob as soon as they set the altitude, though.

Legend

The first column shows whether the subject pushed the MCP altitude knob after setting the altitude to 9,000. Two responses are coded as incorrect, and both are cases where the S fails to push the MCP altitude knob immediately.

The second column shows whether the subject pushed the MCP altitude knob after setting the altitude to FL350. Four responses are coded as incorrect. Two of them are cases where the S fails to push the MCP altitude knob immediately. And, two of them are cases where the S pushes the altitude knob when it is not needed.

The third column shows whether the subject pushed the MCP altitude knob after setting the altitude to FL350. Four responses are coded as incorrect. Two of them are cases where the S fails to push the MCP altitude knob immediately. And, two of them are cases where the S pushes the altitude knob when it is not needed.

Table 3-17. Subjects use of the MCP altitude knob: descending

	12,000	10,000	7,000	3,500	2,200	700
S6:	**Y**	**YY**	--	**Y**	Y	N (*600)
S11:	**Y**	N	**N** (late 1:04)	**Y**	Y	N
S17:	**FL**	**N** (late :20)	**N** (late :10)	N	Y	N
S18:	**Y**	**Y**	Y	**Y**	Y	Y
S7:	Y	**N** (late :26)	Y	Y	**Y**	--
S9:	**Y**	N	-- (*5000)	**Y**	FL	V/S
S12:	**Y**	**N** (late :17)	FL	FL	FL	FL
S14:	**Y**	N	**N** (late :33)	FL	V/S	--
S16:	**V/S**	N	**N** (late :52)	FL	V/S	V/S
S20:	--	--	--	N	**Y**	**N** (late)
S4:	N	N	N	**Y**	Y	N
S5:	Y	**Y**	Y	N	FL	FL (*500)
S10:	Y	N	FL	**Y**	FL	FL
S13:	**Y**	Y	N	N	V/S	V/S
S15:	**V/S**	FL	FL	FL	FL	FL
S22:	**Y**	Y	Y	FL	V/S	V/S
S8:	N	FL	Y	FL	FL	V/S
S19:	Y	N	Y	Y	V/S	V/S (*500)
S21:	N	Y	Y	Y	--	Y
S23:	Y	N	FL	N	N	V/S

* these values are alternate altitude values that were used instead of the value at the column heading.

Legend

The first column shows whether the subject pushed the MCP altitude knob after setting the altitude to 12,000. The errors in this column are a recapitulation of Table 3-8, where we described how subjects failed to leave cruise altitude. Note that entries in the table are

- Y for yes, pushed the MCP altitude knob.
- N for no, did not push the MCP altitude knob.
- FL for engaged FLCH to descend.
- V/S for engaged V/S to descend.

The second column shows whether the subject pushed the MCP altitude knob after setting the altitude to 10,000. N responses coded as errors are cases where the subject failed to push the MCP altitude knob immediately. Y response coded as errors are cases where the subject pushed the MCP altitude knob when it was unneeded. Missing data (--) indicate that the subject didn t dial that altitude on the MCP altitude window.

The third column shows whether the subject pushed the MCP altitude knob after setting the altitude to between 8,000 and 7,000.

The fourth column shows whether the subject pushed the MCP altitude knob after setting the altitude to 3,500.
The fifth column shows whether the subject pushed the MCP altitude knob after setting the altitude to 2,200.
The sixth column shows whether the subject pushed the MCP altitude knob after setting the altitude to 700.

Table 3-18. Subject statements about VNAV PTH mode

1 — the FMC flies to an FMC-calculated geographically fixed path
2 — the FMC flies path on elevator and airspeed on autothrottle
3 — FMC altitude and airspeed constraints will be met during descent
4 — occurs when in cruise at FMC cruise altitude
5 — occurs when on the FMC-created descent path
6 — occurs during climb when there is an FMC altitude constraint tied to a waypoint

	St #1	St #2	St #3	St #4	St #5	St #6
S4:	-	1	-	1	1	-
S5:	1	-	-	-	1	-
S6:	-	-	-	1	x	-
S7:	-	-	-	1	1	-
S8:	1	-	1	1	-	1
S9:	-	-	-	1	1	-
S10:	1	-	-	1	1	1
S11:	1	-	-	1	1	-
S12:	-	-	-	1	-	-
S13:	1	-	-	1	1	-
S14:	1	-	-	x	1	1
S15:	1	-	-	1	1	-
S16:	1	-	1	-	-	-
S17:	-	-	-	1	1	-
S18:	1	-	-	-	1	x
S19:	-	-	-	1	1	-
S20:	1	-	-	-	1	-
S21:	-	-	-	1	-	-
S22:	1	-	-	1	1	-
S23:	1	-	-	1	1	-
	12	1	2	15	15	3

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement

Table 3-19. Subject statements about VNAV SPD mode

1 — the FMC controls to an airspeed target but is not linked to a geographic fixed path
2 — the FMC flies speed on elevator
3 — FMC altitude and airspeed constraints may be missed on descent
4 — occurs during climb, when not level at an FMC altitude constraint
5 — occurs during descent, when deviating sufficiently from FMC path or when using speed intervention

	St #1	St #2	St #3	St #4	St #5
S4:	-	1	-	1	-
S5:	1	-	-	-	-
S6:	-	-	x	1	-
S7:	-	-	-	1	-
S8:	-	1	1	-	1
S9:	-	-	-	1	-
S10:	1	-	-	1	1
S11:	1	-	-	-	-
S12:	-	-	-	1	1
S13:	-	-	-	1	-
S14:	1	-	-	-	-
S15:	-	-	-	1	1
S16:	-	-	-	-	1
S17:	-	-	-	-	1
S18:	-	-	-	1	1
S19:	-	-	-	1	-
S20:	-	-	-	1	-
S21:	-	-	-	1	-
S22:	-	1	-	1	-
S23:	-	x	-	-	1
	4	3	1	13	8

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement

Table 3-20. Subject statements about VNAV ALT mode

1 — VNAV ALT active only when level at MCP altitude (that is not in the FMC)
2 —flies altitude on elevator and airspeed on autothrottle
3 —will not lead to an automatic start of descent at T/D
4 —to leave VNAV ALT, dial in a new MCP altitude above or below current altitude and push MCP altitude button
5 — becomes active when the aircraft levels off at an MCP altitude

	St #1	St #2	St #3	St #4	St #5
S4:	1	-	-	-	1
S5:	1	-	-	-	1
S6:	x	-	-	-	-
S7:	1	-	-	-	1
S8:	1	-	-	-	-
S9:	1	-	-	-	-
S10:	x	-	-	-	1
S11:	1	-	-	-	1
S12:	1	-	-	-	-
S13:	1	-	-	-	1
S14:	1	-	x	-	-
S15:	1	-	-	-	1
S16:	-	-	-	-	-
S17:	-	-	-	-	-
S18:	1	-	-	-	-
S19:	1	-	-	-	-
S20:	1	-	1	-	-
S21:	-	-	-	-	-
S22:	1	-	-	-	1
S23:	1	-	-	-	1
	15	0	1	0	9

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement

Table 3-21. Subject statements about VNAV PTH to VNAV SPD transitions

1 — when deviating significantly from VNAV PTH FMC-calculated path
2 — when climb is resumed after leveling off at FMC altitude constraint
3 — when speed intervention is used on descent

	St #1	St #2	St #3
S4:	?	?	?
S5:	1	-	-
S6:	-	-	-
S7:	1	-	-
S8:	1	-	-
S9:	?	?	?
S10:	1	-	-
S11:	1	-	-
S12:	1	-	1
S13:	1	-	-
S14:	1	-	-
S15:	1	-	-
S16:	-	-	1
S17:	-	-	1
S18:	1	-	1
S19:	-	-	-
S20:	-	-	-
S21:	-	-	-
S22:	1	-	-
S23:	-	-	-
	11	0	4

Legend:
1 — S made the correct statement
? — S claimed he didn t know the answer

Table 3-22. Subject statements about VNAV SPD to VNAV PTH transitions

1 —leveling off at either an FMC intermediate level-off or at the FMC cruise altitude after climbing in VNAV SPD
2 — recapturing the VNAV descent path on descent (transition back to VNAV PTH)

	St #1	St #2
S4:	1	-
S5:	-	1
S6:	x	x
S7:	1	-
S8:	1	-
S9:	1	-
S10:	-	1
S11:	-	-
S12:	-	1
S13:	1	-
S14:	-	-
S15:	-	1
S16:	-	-
S17:	-	-
S18:	-	1
S19:	-	1
S20:	-	1
S21:	-	-
S22:	-	1
S23:	-	-
	5	8

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement

Table 3-23. Subject statements about VNAV PTH to VNAV ALT transitions

1 — descending in VNAV PTH and using altitude intervention to level at an MCP altitude

	St #1
S4:	1
S5:	1
S6:	**x**
S7:	1
S8:	1
S9:	1
S10:	1
S11:	1
S12:	1
S13:	1
S14:	-
S15:	1
S16:	-
S17:	**x**
S18:	1
S19:	**x**
S20:	**x**
S21:	-
S22:	-
S23:	1
	12

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement

Table 3-24. Subject statements about VNAV ALT to VNAV PTH transitions

1 —entering an MCP altitude lower than the CDU cruise altitude, then changing the CDU cruise altitude accordingly, and pushing the MCP altitude knob
2 —after leveling off at MCP altitude (below FMC path) on descent, then intercepting path and descending

	St #1	St #2
S4:	-	1
S5:	-	-
S6:	-	-
S7:	-	-
S8:	x	x
S9:	-	-
S10:	1	-
S11:	-	-
S12:	-	1
S13:	-	-
S14:	-	-
S15:	1	-
S16:	-	-
S17:	-	-
S18:	-	-
S19:	-	1
S20:	-	-
S21:	-	-
S22:	-	-
S23:	-	-
	2	3

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement

Table 3-25. Subject statements about VNAV priorities in descent

1 —VNAV will give up airspeed initially to maintain the VNAV descent path, and will give up on the path and transition to VNAV SPD if the aircraft speed deviates by more than 15 kts and the aircraft deviates by more than 150ft from the flight path

2 — Automation will maintain the path if you are low and slow, but not if you are high and fast.

	St #1	St #2
S4:	1	-
S5:	-	-
S6:	1	-
S7:	1	-
S8:	1	-
S9:	1	-
S10:	1	-
S11:	1	-
S12:	-	-
S13:	x	-
S14:	1	-
S15:	1	-
S16:	x	-
S17:	1	-
S18:	1	1
S19:	x	-
S20:	1	-
S21:	1	-
S22:	1	-
S23:	1	-
	15	1

Legend:

1 — S made the correct statement

x — S made a statement contrary to the correct statement

Table 3-26. Subject statements about VNAV mode and phase of flight

1 —it is important to verify that VNAV PTH is active at cruise altitude to make sure that the aircraft starts its descent when it reaches the T/D point
2 —it is important to verify that VNAV PTH is active on descent because this is a prerequisite for meeting the FMC speed and altitude constraints

	St #1	St #2
S4:	1	-
S5:	-	1
S6:	-	-
S7:	1	1
S8:	-	-
S9:	-	-
S10:	-	1
S11:	-	-
S12:	-	-
S13:	-	-
S14:	-	-
S15:	-	-
S16:	-	-
S17:	-	1
S18:	1	1
S19:	-	1
S20:	-	1
S21:	1	-
S22:	1	-
S23:	1	-
	6	7

Legend:
1 — S made the correct statement

Table 3-27. Subject statements about availability of speed protection

1 —Speed protection is available in VNAV
2 —Speed protection is available in FLCH
3 —Speed protection is not available in V/S
4 —Speed protection is not available if the autothrottle is engaged

	St #1	St #2	St #3	St #4
S4:	1	1	?	1
S5:	1	1	1	1
S6:	1	1	1	1
S7:	-	-	1	-
S8:	1	1	1	1
S9:	1	1	1	1
S10:	1	1	1	1
S11:	1	1	1	?
S12:	1	1	1	1
S13:	1	1	1	1
S14:	1	1	1	NA
S15:	1	-	1	1
S16:	1	1	1	1
S17:	1	1	1	1
S18:	1	1	**x**	?
S19:	?	?	1	1
S20:	1	1	1	-
S21:	-	-	1	1
S22:	1	1	**x**	1
S23:	1	1	1	-
	17	16	17	14

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement
? — S claimed he didn t know the answer
NA — question was not asked explicitly

Table 3-28. Subject statements about airspeed propagation across VNAV pages

1 —An airspeed target entered on the CRZ page does not propagate to the DES page and is not used in VNAV descent.
2 —If new airspeed is not entered on the DES page, VNAV descent reverts to the Econ descent speed.

	St #1	St #2
S4:	1	1
S5:	1	-
S6:	x	-
S7:	1	-
S8:	1	1
S9:	?	?
S10:	1	1
S11:	1	1
S12:	1	-
S13:	1	1
S14:	1	1
S15:	1	-
S16:	1	1
S17:	1	1
S18:	x	-
S19:	1	1
S20:	1	-
S21:	1	1
S22:	1	1
S23:	1	-
	17	11

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement
? — S claimed he didn t know the answer

Table 3-29. Subject statements about use of autothrottle modes

1 —autothrottle modes during takeoff and climb can be what? (best answer: TR,H)(possible: S)
2 —autothrottle modes during cruise can be what? (best answer: S)(possible: T,H)
3 —autothrottle modes during descent can be what? (best answer: I,H,S)(possible: T)
4 —autothrottle modes with VNAV PTH can be what? (best answer: I,H,S)
5 — autothrottle modes with VNAV SPD can be what? (best answer: T,TR,I,H)
6 — autothrottle modes with VNAV ALT can be what? (best answer: S)

	St #1	St #2	St #3	St #4	St #5	St #6	incorrect statement
S4:	NA	NA	NA	I,H,S	TR	S	
S5:	TR	S	H,S	H,S x	H x	S	(no IDLE in VNAV)
S6:	NA	NA	NA	I,H,S	T,TR	S	
S7:	?	?	?	I,H,S	T,TR	S	
S8:	NA	NA	NA	I,H,S	TR,T,I,H	S	
S9:	H	S	I	I,H,S	NA	NA	
S10:	TR,H	?	I,H	I,H,S	?	S	
S11:	?	?	?	I x	T,TR,H	?	(no H or S in V-PTH)
S12:	NA	S	NA	I,H,S	T	S	
S13:	TR,H	S	I,H,S	NA	TR,I,H x	NA	(S for V-SPD)
S14:	?	?	I,H,S	?	?	?	
S15:	TR, S	?	I,H,S	?	?	?	
S16:	?	?	I x	I,H,S x	?	?	(T,TR for V-PTH)
S17:	?	?	I,H	?	?	?	
S18:	TR	S	I,H,S	I,H,S	NA	NA	
S19:	TR,H	NA	I	I	?	?	
S20:	TR	S	I,H	I,H	NA	NA	
S21:	TR,H	NA	I,H,S	NA	NA	NA	
S22:	TR,H,S	S	I,H,S	I,H	?	?	
S23:	TR x	S	I,H	I,H	NA	NA	(T on climb)
	5	8	6	8	1	7	

Responses coded as

THR REF — TR	IDLE — I
THR — T	SPD — S
HOLD — H	

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement
? — S claimed to generally not know these answers
NA — question was not asked explicitly

Table 3-30. Subject statements about capture criteria for LNAV

1 —the plane must be pointed at the active leg, inbound to the active waypoint
2 —LNAV will capture when the cross track error is 2.5 NM (generally)
3 —the distance at which LNAV will capture depends on several factors, including wind, ground speed, and course intercept geometry

	St #1	St #2	St #3
S4:	-	-	-
S5:	-	-	-
S6:	-	-	-
S7:	-	-	-
S8:	-	-	-
S9:	-	-	-
S10:	-	-	1
S11:	-	-	-
S12:	-	-	1
S13:	1	-	-
S14:	-	-	-
S15:	-	-	-
S16:	-	-	-
S17:	1	-	-
S18:	-	-	-
S19:	-	-	-
S20:	-	-	-
S21:	-	-	-
S22:	-	-	-
S23:	-	1	-
	2	1	2

Legend:
1 — S made the correct statement

Table 3-31. Subject statements about use of the MCP altitude knob

1 —the MCP altitude knob needs to be pushed when you are level at an MCP altitude and have set a higher or lower altitude and want to climb or descend; or when you are at an FMC altitude constraint but failed to reset the MCP altitude above or below you until after you pass the waypoint (or T/D)
2 —the MCP altitude should not be pushed when setting an MCP altitude above you or below you; or when you set the MCP above you or below you and the FMC will initiate the climb or descent

	St #1	St #2
S4:	1	1
S5:	-	-
S6:	-	-
S7:	-	-
S8:	-	-
S9:	1	-
S10:	1	-
S11:	1	1
S12:	1	x
S13:	1	1
S14:	-	-
S15:	1	-
S16:	-	-
S17:	1	-
S18:	-	1
S19:	-	-
S20:	-	-
S21:	-	-
S22:	1	1
S23:	1	-
	10	5

Legend:
1 — S made the correct statement
x — S made a statement contrary to the correct statement

Table 3-32. Subject statements about general rules for use of the MCP altitude knob

1 —pressing the MCP altitude knob, generally speaking, removes a restriction, which can be your current altitude, or an FMC (LEGS page) waypoint restriction, or FMC cruise altitude, or the MCP altitude you have captured.
2 —the FMC will initiate an altitude change without pushing the MCP altitude knob, generally speaking, when you are in VNAV and the MCP altitude is not your current altitude (unless you have passed the constrained waypoint or T/D point)

	St #1	St #2
S4:	NA	NA
S5:	-	-
S6:	-	-
S7:	-	-
S8:	1	-
S9:	-	-
S10:	1	-
S11:	1	-
S12:	-	-
S13:	1	-
S14:	-	-
S15:	-	-
S16:	-	-
S17:	1	-
S18:	1	-
S19:	1	-
S20:	1	-
S21:	1	-
S22:	1	-
S23:	1	-
	11	0

Legend:
1 — S made the correct statement
NA — question was not asked explicitly