

## THE EFFECT OF AUTOMATION ON THE FREQUENCY OF TASK PRIORITIZATION ERRORS ON COMMERCIAL AIRCRAFT FLIGHT DECKS: AN ASRS INCIDENT REPORT STUDY

Jennifer Rae Wilson  
Research Integrations, Inc.

Ken Funk  
Oregon State University

### ABSTRACT

Task Management (TM) refers to the function in which the human operator manages his/her available sensory and mental resources in a dynamic, complex, safety-critical environment in order to accomplish the multiple tasks competing for a limited quantity of attention. There is reason to believe that the level of automation on the commercial aircraft flight deck may affect TM. Two samples of 210 Aviation Safety Reporting System (ASRS) incident reports were compared to determine how automation affects TM. The first sample consisted of reports submitted by pilots flying advanced technology aircraft and the second by pilots flying traditional technology aircraft. TM was explored by looking at Task Prioritization errors. Twenty-eight incident reports from the advanced technology sample and 15 were from the traditional technology sample were classified as containing Task Prioritization errors. This difference was found to be statistically significant.

### INTRODUCTION

In recent years there has been a growing awareness of human factors issues associated with the increased presence of automated systems on modern commercial aircraft flight decks. As flight deck automation becomes more sophisticated, it is able to perform many of the tasks previously performed by the pilots and the pilot's role becomes more like that of a manager. With this shift in the pilot's role, new strategies must be developed for the pilots to successfully perform their tasks. In order to do this, the effect that automation has on pilot performance must be understood.

### BACKGROUND

In a general sense, today's commercial air carrier fleets are composed of two types of aircraft: advanced technology and traditional technology. The advanced technology aircraft incorporate a number of sophisticated automated systems that have the ability to perform tasks that in the past have been performed

exclusively by the human pilots. These systems include such devices as the advanced autopilot, the Flight Management System (FMS), electronic instrument displays, and warning and alerting systems. Traditional technology aircraft are defined as lacking these types of automated systems. The presence of both types of aircraft in commercial fleets gives us a unique opportunity to compare them in present day operations.

#### *Errors enabled by automation*

While there is little doubt that technology has made significant contributions to both the safety and efficiency of operations, there are still concerns about replacing the human functioning with automated systems. With the introduction of automated systems, some flightcrew errors that had been a problem in the past have been significantly reduced (or eliminated). An example of this is the ability of the automated systems to track a precise heading with minimal deviations from the desired path in situations that pilots flying manually may err. On the other hand, this functional change may also create the opportunity for errors that had not been possible in the past, or increase the chance of previously existing errors to occur; Wiener (1989) has referred to this as "enabling" errors. An example of a type of error that has been enabled is gross navigational deviations due to data entry errors. It was a gross navigational error of this type that contributed to the American Airlines flight 965 accident near Cali, Colombia. Aeronautica Civil, the aircraft accident investigating board of the Republic of Colombia, determined that during the approach into the airport in Cali, Colombia, the flightcrew "selected and executed a direct course to the identifier 'R,' in the mistaken belief that R was Rozo as it was identified on the approach chart. The pilots could not know without verification with the EHSI [Electronic Horizontal Situation Indicator] display or considerable calculation that instead of selecting Rozo, they had selected the Romeo beacon, located near Bogota, some 132 miles east-northeast of Cali" (Aeronautica Civil of the Republic of Colombia, 1997,

p. 41). With the Romeo beacon programmed into the FMS, the airplane departed from its inbound course to Cali and flew east toward Bogota. When the flightcrew realized that they were off-course, they turned right to return to the extended centerline of the runway at Cali. At this point however, a direct course to the Cali airport led the aircraft into high mountainous terrain and shortly after their turn the aircraft impacted the side of a mountain. It would be highly unlikely that a gross navigational error such as this could occur without the automated systems.

Work focusing on these errors should not be to determine whether the traditional technology or the advanced technology produces the most errors overall, but rather to understand the errors that will be encountered in the new aircraft and how they may be successfully controlled.

At this time there is no comprehensive listing of what the errors enabled by the advanced technology flight deck are but there are a number of ideas about what they may be. Funk et al. (in review) have compiled a list of 92 issues about automation on the flight deck from a broad range of sources including accident reports, incident report studies, surveys, and scientific experiments. To determine which of these issues should be valid concerns, they compiled a database of both supporting and contradictory evidence that addresses these issues. They found that many of the automation issues require further investigation to determine if they are indeed problems with which the aviation community should be concerned.

### *Task Management*

Task Management (TM) refers to the function in which the human operator manages his/her available sensory and mental resources in a dynamic, complex, safety-critical environment in order to accomplish the multiple tasks competing for a limited quantity of attention. This function includes task initiation, monitoring, task prioritization, resource allocation, and task termination (Funk, 1991). Flightcrews must perform TM on the commercial flight deck because they do not possess the necessary resources to simultaneously execute all the tasks that demand their attention. The flightcrew must therefore prioritize the tasks in the order of most to least important and then allocate their resources according to this prioritization. In a dynamic system, the state of each task demanding attention continuously changes and as this occurs so too may change the relative urgency with which each task must be completed. Thus, the flightcrew must continuously perform the function of TM in order to

maintain awareness of the changes in the state of the system and make the necessary revisions to the task prioritization.

While recently there have been several studies that have begun to look at TM on the commercial flight deck (Latorella, 1996; Rogers, 1996; Chou, Madhavan, & Funk, 1995), none of these studies has specifically addressed the relationship between TM and automation. There has been speculation that the level of automation on the flight deck may affect TM, in fact, one of the 92 automation issues identified by Funk et al. (in review) mentioned earlier concerns TM:

*issue167:* The use of automation may make task management more difficult for flightcrews, possibly leading to unsafe conditions.

The reasons behind this speculation will be covered in a later section of this paper.

### *Task Prioritization Errors*

TM can be investigated by looking at the errors that flightcrews commit in prioritizing the tasks demanding attention. To do this it must be assumed that there is a “right” way and a “wrong” way to prioritize and that the ultimate prioritization of a flightcrew can be (at least partially) determined by observing the choice of tasks performed. Because the flightcrew is limited by the quantity of attention that they have available to distribute across the tasks they perform, they must manage tasks in such a way that higher priority tasks are allocated the available attention before lower priority tasks. If the flightcrew does not allocate his/her attention in this way, it is said that a Task Prioritization error is committed. Specifically, a Task Prioritization error is when the flightcrew gives their attention to a lower priority task to the detriment of a higher priority task. As the TM becomes more difficult, it is expected that the frequency of Task Prioritization errors would increase.

The prioritization strategy taught to every novice pilot is *aviate, navigate, communicate, and manage* systems. The tasks in the *aviate* category are concerned with using the flight systems and controls to fly the plane. The tasks in the *navigate* category are those concerned with planning the route and high-level route changes. The tasks in the *communicate* category are concerned with explicit communication with systems on the ground, between the flightcrew, with the cabin crew, with the company, and with the passengers. The tasks in the *manage systems* category are those concerned with assuring that the systems are operating normally and are

capable of performing the functions necessary for the aviate tasks. Intuitively this rule of thumb makes sense. For example, it is easy to agree with the idea that keeping the plane in the air (i.e., aviate) is more important than making sure it is headed in the desired direction (i.e., navigate).

### *Consequences of Task Prioritization Errors*

Task prioritization errors can have disastrous consequences as evidenced by several accidents at least partially attributed to Task Prioritization errors. The following are two accidents in which the accident investigating board determined that misprioritization played a key role in the accident.

The first example is the often cited L-1011 Florida Everglades accident. On December 29, 1972, an Eastern Air Lines Lockheed L-1011 aircraft crashed approximately 18 miles west-northwest of Miami International Airport destroying the aircraft and killing 99 people on board. The National Transportation Safety Board (NTSB) determined that the probable cause of this accident was the flightcrew's failure to monitor the flight instruments during the final 4 minutes of the flight. Preoccupation with a malfunctioning nose landing gear position indicating system distracted the flightcrew's attention away from the instruments and allowed the descent to go unnoticed (NTSB, 1973). This would be considered a Task Prioritization error because the lower priority task of troubleshooting the malfunctioning landing gear indication (i.e., a manage systems task) was allocated attention while the higher priority task of maintaining the aircraft's altitude (i.e., an aviate task) was not allocated appropriate attention.

The second example is the Indian Airlines A320 accident in Bangalore, India. On April 14, 1990, an Indian Airlines Airbus A320 aircraft crashed just short of the runway at the Bangalore airport destroying the aircraft and killing 90 people on board. The investigators determined that the probable cause of the accident was the failure of the pilots to realize the gravity of the situation and immediately apply thrust. The pilots spent the final seconds of the flight trying to understand why the plane was in idle/open descent mode rather than taking appropriate action to avoid impact with the ground (Ministry of Civil Aviation, India, 1990). Again, the flightcrew committed a Task Prioritization error by allocating attention to the lower priority task of trying to understand the reason the automation was in a particular mode (i.e., a manage systems task) while the higher priority task of correcting the aircraft's descent (i.e., an aviate task) was not allocated appropriate attention.

Both of these accidents illustrate the disastrous consequences Task Prioritization errors can have. The Lockheed L-1011 is a traditional technology aircraft while the Airbus A320 is an advanced technology aircraft. These two accidents were chosen to illustrate two things: one, Task Prioritization errors occur in both advanced and traditional technology aircraft types; and two, the consequences of a Task Prioritization error can be equally fatal regardless of the aircraft type.

### *Why Automation May Affect Task Prioritization*

An increasing number of accidents and incidents can be attributed to TM errors (e.g. Chou, Madhavan, & Funk, 1995). It has been speculated that higher levels of automation may make TM more difficult for flightcrews. There are several reasons behind this speculation. First, there are a greater number of tasks to be performed in the automated aircraft. All the flight control tasks found in the traditional technology aircraft must still be performed in the advanced technology aircraft but, in addition to these tasks, there are now tasks associated with communicating with and managing the automation. Adding tasks to the queue of tasks demanding attention increases the demands on the flightcrew. While the automation provides additional external resources for the flightcrew to utilize, these resources must be managed which increases demands on the function of TM. Second, the same resources may be overloaded in the automated aircraft. Some of the demands added by automation require the cognitive processing resources that are already taxed in the traditional technology aircraft. Because of this, more prioritization may be required because more tasks are demanding the same resources. And third, some of the advanced systems, such as the Flight Management System (FMS), may inappropriately draw the attention of the flightcrew away from more critical tasks. When the FMS fails to behave as expected, the flightcrew's attention can be drawn away from the highest priority tasks required for flying the aircraft. Two factors contribute to the ability of the FMS to draw the flightcrew's attention. First, because of the nature of the FMS the flightcrew cannot proceed with any other tasks until they either satisfy its needs or they turn it off. If pilots have an incentive to keep the FMS on, then they must correct the problem before their attention can be turned elsewhere. Second, when the FMS fails to behave as expected the flightcrew's attention is drawn toward it as suggested by schema theory. As the functioning of the FMS defies explanation within the currently active schema, attention will be directed toward finding a better fitting schema. This

phenomenon is sometimes referred to as ‘novel pop out’ (Johnston, Hawley, Plewe, Elliot, & DeWitt, 1990).

### *ASRS Incident Reports*

Because our interest lies in TM as it occurs in real flight operations, we ideally would like to collect data from real flight operations. However, this method is often impractical. A viable alternative to viewing actual line operations is the use of incident reports submitted by pilots, such as those submitted to the Aviation Safety Reporting System (ASRS).

The ASRS was created as a means to collect reports of situations that compromise safety so that strategies to prevent these situations from becoming accidents could be created (Chappell, 1994). These reports are called “incident reports” and are submitted voluntarily by aviation operations personnel (e.g., pilots, Air Traffic Controllers, flight attendants, ground personnel). The reports contain a description of a situation occurring in flight operations that the reporter believes has safety implications. With each report providing a description of an event that occurred in operations, they can be used as a practical way to view real line operations from a pilot’s perspective.

An example of an ASRS incident report is given in Figure 1(a). The abbreviations used can make the report difficult to understand so Figure 1(b) presents a more readable translation of this example.

---

#### **Accession #92507**

##### **Synopsis**

ACR MLG ALT DEVIATION EXCURSION FROM CLRNC ALT. REPORTER SAYS FMA CHANGED FLT MODE AND ALT SELECT BY ITSELF.

##### **Narrative**

THE F/O WAS FLYING THE ACFT. WE HAD BEEN ISSUED SEVERAL VECTORS AND TURNS BY ATC FOR FLOW CTL INTO CHICAGO O'HARE. I WAS ON THE P/A EXPLAINING THE ENRTE DELAY TO THE PAX WHEN I NOTICED THE FMA HAD CHANGED FROM "PERF CRUISE" TO "PERF DSCNT," AND THE ALT SELECT HAD CHANGED FROM 35000 TO 33000'. I ASKED THE F/O IF WE HAD BEEN CLRED TO FL330. HE SAID NO. THE ACFT ALT WAS 34600' WHEN I NOTICED THE PROB. THE DSCNT WAS STOPPED AT 34500'. I DON'T KNOW WHY THE AUTOPLT ENTERED A DSCNT MODE. AN ALT WARNING DIDN'T OCCUR BECAUSE THE ALT SELECT HAD CHANGED ALSO. I SUSPECT A PWR SURGE IN THE ELECTRICAL SYS MAY HAVE CAUSED THE PROB. I HAVE EXPERIENCED THIS PROB IN THE PAST WITH THE MLG FLT GUIDANCE SYS WHEN A HYD PUMP IS TURNED FROM LOW TO HIGH.

---

**FIGURE 1(a)** Synopsis and Narrative of ASRS Incident Report #92507.

---

#### **Accession #92507**

##### **Synopsis**

A medium-large transport aircraft used by an air carrier committed an altitude deviation. The aircraft made an excursion from the clearance altitude. The reporter says that the Flight Mode Annunciator (FMA) changed flight mode and altitude select by itself.

##### **Narrative**

The first officer was flying the aircraft. We had been issued several vectors and turns by Air Traffic Control to control the flow of traffic into Chicago O’Hare International Airport. I was on the public address explaining the enroute delay to the passengers when I noticed the FMA had changed from "PERF CRUISE" to "PERF DSCNT," and the altitude select had changed from 35000 to 33000 feet. I asked the first officer if we had been cleared to a flight level of 33000 feet. He said no. The aircraft’s altitude was 34600 feet when I noticed the problem. The descent was stopped at 34500 feet. I don’t know why the autopilot entered a descent mode. An altitude warning didn’t occur because the altitude select had changed also. I suspect a power surge in the electrical system may have caused the problem. I have experienced this problem in the past with the medium-large aircraft flight guidance system when a hydraulic pump is turned from low to high.

---

**FIGURE 1(b)** Readable translation of Synopsis and Narrative of ASRS Incident Report #92507.

In the past due to the nature of the data, ASRS incident reports have been used primarily for descriptive analyses. In this study, however, it would be more useful to conduct an inferential analysis. Such an analysis may be conducted by carefully constructing a research question and choosing an appropriate statistical test. Because few researchers have taken this approach, there are not many examples of effective inferential analysis using ASRS incident report data.

#### **RESEARCH OBJECTIVES**

The flightcrew’s function of TM on the commercial flight deck is an important part of flight operations, and committing errors in TM can have severe consequences. There is reason to believe that the level of automation may affect TM, however to date there has been little research that directly addresses this effect. Thus, the primary objective of this study was to begin evaluating the relationship between TM of commercial airline pilots and the level of automation on the flight deck by determining how automation affects the frequency of Task Prioritization errors as reported in Aviation Safety Reporting System (ASRS) incident reports.

Because ASRS incident reports are primarily used for descriptive analyses, a methodology for conducting

a good statistical comparison analysis is lacking. Therefore, the secondary objective of this study was to create a methodology that models an effective way to use ASRS incident report data in an inferential analysis.

## METHODS

The objectives of this study were met by carefully constructing a study to ensure that a fair comparison was made between the advanced and traditional technology populations. To accomplish this, representative data samples were drawn from an ASRS incident report database and analyzed using an analysis tool constructed specifically for this study.

### *Sample Size Determination*

Two samples of ASRS incident reports were compared in this study to determine if level of automation on the commercial aircraft flight deck affected the frequency of Task Prioritization errors. The first sample was composed of 210 incident reports submitted by pilots flying advanced technology aircraft and the second sample was composed of 210 incident reports submitted by pilots flying traditional technology aircraft. In total, 420 incident reports were analyzed.

The possibility exists that the effect of the level of technology of the aircraft could be confounded with differences in experience level because the advanced aircraft are comparatively new to commercial air carriers' fleets. To help avoid this confounding effect, the two samples were divided into three sub-samples each made up of 70 reports submitted during a specified time period: 1988-1989, 1990-1991, and 1992-1993. These submission periods were based on the availability of incident reports with narratives in the CD-ROM database used.

The sample sizes were determined by performing a power analysis using the following values: power = 0.80<sup>1</sup>, significance level of  $\alpha = 0.05$ , and the effect size index of  $w = 0.20$ . It was determined that a sample size of 196 incident reports was required to reject the null hypothesis, or in other words, conclude that there is a significant difference between the frequencies found in the two samples. Because each sample was to be divided into 3 sub-samples ( $196/3 = 65.333$ ), the sample size was rounded up to 210 ( $210/3 = 70$ ).

A second power analysis was performed to determine if the sub-sample size of 70 was adequate. With the power = 0.80, significance level of  $\alpha = 0.05$ , and the effect size index = 0.40, it was determined that a

sub-sample size of 49 incident reports was required to reject the null hypothesis. Because 70 is greater than 49, the sub-sample size of 70 was determined to be adequate.

It should be noted that the two power analyses conducted each used a different effect size index. The effect size index for each of the power analyses was chosen specifically for the effect that was to be detected. For the two aircraft technology type samples, we wanted to detect the smallest effect size without the sample size becoming prohibitively large. If a difference between the frequency rates of Task Prioritization errors between the two technology types existed, we wanted to detect it. The effect size index of 0.20 (loosely referred to as a 'medium-small' effect) was chosen for these samples. For the submission period sub-samples, we were interested only in detecting an effect of submission period that was large enough to significantly confound the effect of aircraft technology. It was not necessary to detect as small an effect for the sub-samples as was required for the aircraft technology type samples. Thus, we chose to use an effect size index of 0.40 (loosely referred to as a 'medium-large' effect) for the submission period sub-samples.

### *Report Selection Criteria*

The ASRS incident reports used in this study were collected using the ASRS Aeroknowledge CD-ROM database (DOS Version Release 96-1). Homogeneity between samples is very important for statistical comparison studies. In an effort to collect homogenous samples, the sample populations<sup>2</sup> were constrained so that the level of automation (i.e., aircraft technology type) and the submission period were the only two differences between the samples. For example, all the reports from both the advanced technology and the traditional technology samples were constrained to reports submitted by a member of the flightcrew flying a two-person commercial air carrier aircraft in which the aircraft was classified as a medium-large transport, large transport or widebody transport aircraft.

Another parameter that was held constant was phase of flight. Based on the fact that over half of all commercial hull loss accidents (Boeing, 1997) and that approximately 50% of incidents reported to ASRS by commercial air carrier pilots occur during the terminal phases of flight (Wilson, 1998), these phases of flight were considered a good place to look for errors. Thus,

<sup>1</sup> When conducting a power analysis, it is a convention to set the power at 0.80 (Cohen, 1988).

<sup>2</sup> The term population is used here to denote the population of reports that meet the parameters defined. This usage of the term should not be confused with the population of all ASRS incident reports, or the population of all errors committed by flightcrews.

all reports analyzed occurred during the descent or approach phase of flight.

### *Report Selection Methodology*

The reports were collected from the database in the following way to ensure that the samples were representative of the population. First, the six populations (i.e., the two aircraft technology populations each divided into three submission periods) were compiled from the database based on the population parameters described above. Second, based on the total number of reports in each of the six populations 70 random numbers for each sample were generated to determine which of the reports would be included in the sample. This allowed the samples to be drawn randomly without replacement. Third, the appropriate reports were then tagged and downloaded into a word processing document. Fourth, all information related to the report except for the ASRS number, the synopsis, and the narrative was removed. This was done so that the analyst would be unable to use this information to identify the report during analysis. Any information in the synopsis or the narrative that identified the report was not removed because the deletions would have left the report incomplete.

### *Analysis Tool*

An incident analysis form was developed specifically for use in this project. This form allowed the analyst to classify the ASRS incident reports as either containing a Task Prioritization error or not based on the description given in the narrative of the report. Using the form, the analyst identified the tasks that were being performed during the incident period reported. Prioritization was evaluated by identifying whether the active tasks were related to the task categories of aviate, navigate, communicate, manage systems, or non-flight related tasks. If a task of lower priority was active while a task of higher priority that required resources was not active, the report was classified as containing a Task Prioritization error.

The incident analysis form contained a listing of all tasks that must be performed during the descent and approach phases of flight. The task listing used was based on a functional analysis of a generic commercial air transport mission (Alter & Regal, 1992). These tasks were organized into four categories and the priority of the task was determined by the category to which it belonged (where 1 is highest and 5 is lowest): 1. Aviate, 2. Navigate, 3. Communicate, 4. Manage Systems, and 5. Non-Flight Related. There was no

priority hierarchy within a category; it was assumed that all tasks that fell in a particular category were of the same priority. Each listed task was defined not only as performing the task itself, but also as maintaining awareness of the task's status. For example, the task '1.5 Control/monitor vertical profile' included controlling the vertical profile either manually or using the autopilot and monitoring the status of the vertical profile.

To illustrate how the analysis form was filled out, consider incident report #92507 shown earlier in Figure 1(a) and its corresponding analysis form in Figure 2.

Associated with each task listed on the form were three sets of boxes that were marked to highlight the parameters that were considered in the analysis. When any of the boxes were marked for a given task, the analyst entered an excerpt or short summary upon which the judgment to mark the box had been based in the column called 'Related Excerpt/Comment.'

Starting on the left, the first set of boxes, 'Reported Tasks,' were used to indicate all of the tasks that were reported as being performed during the block of time described in the incident. This set of boxes was used to give a rough summary of all tasks that the reporter had described. The analyst marked the 'explicitly stated' box if the reporter specifically mentioned the task in the narrative of the report. For example, given the following statement from incident report #92507:

“WE HAD BEEN ISSUED SEVERAL VECTORS AND TURNS BY ATC FOR FLOW CTL INTO CHICAGO O'HARE.”

The analyst would mark the 'explicitly stated' box for the Task 3.1 'Communicate with ATC' and include the excerpt 'ISSUED SEVERAL VECTORS AND TURNS BY ATC.' Reading on from this statement, it is implied, though not explicitly stated, that the flightcrew began to carry out these requests given by the ATC.

“WE HAD BEEN ISSUED SEVERAL VECTORS AND TURNS BY ATC FOR FLOW CTL INTO CHICAGO O'HARE. I WAS ON THE P/A EXPLAINING THE ENRTE DELAY TO THE PAX...”

The analyst would mark the 'strongly implied' box for the Task 1.3 'Control/monitor lateral profile' and again include the excerpt 'ISSUED SEVERAL VECTORS AND TURNS BY ATC.'

The next box, 'ACTIVE TASKS during CRITICAL PERIOD,' was marked when the task was active during the critical period of the incident. The

## Incident Report Analysis Form

Accession #: 92507

Synopsis: ACR MLG ALT DEVIATION EXCURSION FROM CLRNC ALT. REPORTER SAYS FMA CHANGED FLT MODE AND ALT SELECT BY ITSELF

Descent/Approach Tasks: (check appropriate boxes and include explanatory comments)

REPORTED TASKS		TASK LISTING	ACTIVE TASKS during CRITICAL PERIOD	STATUS during CRITICAL PERIOD			RELATED EXCERPTS / COMMENTS
Explicitly Stated	Strongly Implied			Unknown	Satisfactory	Unsatisfactory	
		<b>1. AVIATE TASKS</b>					
<input type="checkbox"/>	<input type="checkbox"/>	1.1 Control/monitor aircraft configuration	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	1.2 Control/monitor attitude	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	1.3 Control/monitor lateral profile	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	"ISSUED SEVERAL VECTORS AND TURNS BY ATC"
<input type="checkbox"/>	<input type="checkbox"/>	1.4 Control/monitor speed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	1.5 Control/monitor vertical profile	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	"FMA HAD CHANGED ... THE ACFT ALT WAS 34600' "
<input type="checkbox"/>	<input type="checkbox"/>	1.6 Maintain clearances and restrictions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	"DSCNT WAS STOPPED AT 34500' "
<input type="checkbox"/>	<input type="checkbox"/>	1.7 Maintain separation with traffic, terrain	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
		<b>2. NAVIGATE TASKS</b>					
<input type="checkbox"/>	<input type="checkbox"/>	2.1 Determine mode of lat/lon navigation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	2.2 Maintain awareness of temporal profile	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	2.3 Modify route for weather, traffic, hazards	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	2.4 Plan approach	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	2.5 Program route in FMS	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	2.6 Set navigational radios	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
		<b>3. COMMUNICATE TASKS</b>					
<input checked="" type="checkbox"/>	<input type="checkbox"/>	3.1 Communicate with ATC	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	"ISSUED SEVERAL VECTORS AND TURNS BY ATC"
<input type="checkbox"/>	<input type="checkbox"/>	3.2 Communicate with cabin crew	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	3.3 Communicate with company	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	<input type="checkbox"/>	3.4 Communicate with flightcrew	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	"I ASKED F/O IF WE HAD BEEN CLEARED"
<input checked="" type="checkbox"/>	<input type="checkbox"/>	3.5 Communicate with passengers	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	"I WAS ON THE P/A"
<input type="checkbox"/>	<input type="checkbox"/>	3.6 Tune communication radios	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	3.7 Uplink/ downlink information	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	3.8 Receive ATIS	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
		<b>4. MANAGE SYSTEM TASKS</b>					
<input type="checkbox"/>	<input type="checkbox"/>	4.1 Manage/correct system faults	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	4.2 Monitor aircraft subsystems	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
		<b>5. NON-FLIGHT RELATED TASKS</b>					
<input type="checkbox"/>	<input type="checkbox"/>	5.1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Critical Period: "GIVEN CLEARANCE ALTITUDE" to "I NOTICED THE PROBLEM"

Task Prioritization: Was a Task Prioritization error committed? (circle one) **YES** NO

If YES, list the tasks involved in the prioritization error: 1.6, 3.4, 3.5

**FIGURE 2** Incident Report Analysis Form completed for ASRS Incident Report #92507.

critical period consisted of all the events that took place between the time that the "desired state" was defined and the time that the flightcrew became aware that the desired state was not or would not be met (i.e., a deviation occurred). The analyst entered the critical period in the appropriate space at the bottom of the form. In incident report #92507, the critical period was "given clearance altitude" to "I noticed the problem." This would indicate that the critical period included all tasks that occurred between the point that the desired state of maintaining the cleared altitude was declared and the point that the flightcrew realized that they had overshot this altitude. In this report, the clearance for their desired altitude had been given before the window of time described in this incident report so all the tasks described up to the point that the captain noticed the problem were considered active tasks.

The last set of boxes, 'STATUS during CRITICAL PERIOD,' were marked if the task was active during the critical period (i.e. had been marked 'ACTIVE TASKS during CRITICAL PERIOD'). The 'Unknown' box was marked when the analyst was unable to discern the task's status from the narrative. For example, it cannot be determined from this narrative if the public address system was working correctly and that the passengers actually heard the captain's announcement. In this case the analyst would mark Task 3.5 'Communicate with passengers' as status 'Unknown.'

The 'Satisfactory' box was marked when the desired state of the task had and/or would be achieved given the current trend of activities. For example, given the following statement:

"...I ASKED THE F/O IF WE HAD BEEN CLRED TO FL330. HE SAID NO..."

The analyst would mark the 'Satisfactory' box for the Task 3.4 'Communicate with flightcrew'. The first officer and the captain effectively communicated this information.

The 'Unsatisfactory' box was marked when the reporter stated in the narrative that the desired state of the task had not and/or would not be achieved given the current trend of activities. For example, given the following statement:

"...THE ALT SELECT HAD CHANGED FROM 35000 TO 33000'. I ASKED THE F/O IF WE HAD BEEN CLRED TO FL330. HE SAID NO. THE ACFT ALT WAS 34600' WHEN I NOTICED THE PROB..."

The analyst would mark the 'Unsatisfactory' box for the Task 1.6 'Maintain clearances and restrictions.' In this

example, the desired altitude was 35,000 feet yet the altitude of the aircraft was 34,600 feet, a discrepancy of 400 feet.

Once all the appropriate boxes were marked on the analysis form, the incident report was classified as to whether a Task Prioritization error was committed by circling 'yes' or 'no.' The classification was determined by using the following rule:

*If the status of a higher priority task is unsatisfactory and it is not active AND a lower priority task is active, then the incident report is classified as "TP error occurred" (otherwise it is classified as "no TP error occurred").*

When a report was classified as containing a Task Prioritization error then the tasks involved in this error were listed in the space provided at the bottom of the analysis form. In incident report #92507, Task 1.6 'Maintain clearances and restrictions' was not active and unsatisfactory while the lower priority tasks 3.4 'Communicate with flightcrew' and 3.5 'Communicate with passengers' were active, thus this incident report was classified as containing a Task Prioritization error.

#### *Application of the Analysis Tool*

Each incident report was analyzed using the incident report analysis form described above. To minimize bias during the analysis, the two samples (including the three sub-samples within each) were randomly mixed and the sample to which each incident report belonged was not specified until all analyses were complete. After all reports had been analyzed, the reports were sorted and the data summarized.

## RESULTS

### *Overall Effect of Technology*

Of the 420 incidents reports analyzed, 43 (10.2%) were classified as containing Task Prioritization errors. Of these, 28 were from the advanced technology sample and 15 were from the traditional technology sample (see Table 1).

The Chi Square ( $\chi^2$ ) test was used to determine if the difference between the 28 Task Prioritization errors found in advanced technology incident reports and the 15 Task Prioritization errors in traditional technology aircraft was statistically significant. The  $\chi^2$  value calculated was 4.379 at 1 degree of freedom with a p value of 0.036. Using a significance level of  $\alpha = 0.05$ ,

it was concluded that this difference was statistically significant.

**TABLE 1** Summary of the frequencies of Task Prioritization errors.

Submission Period	Task Prioritization Error Frequency		Total Errors by Submission Period
	Advanced Technology	Traditional Technology	
1988-1989	13	7	<b>20</b>
1990-1991	11	5	<b>16</b>
1992-1993	4	3	<b>7</b>
<b>Total Errors by Aircraft Technology</b>	<b>28</b>	<b>15</b>	

#### *Effect of Technology by Submission Period*

The  $\chi^2$  test was used next to compare the frequency difference between advanced and traditional technology aircraft by submission period. For each of the three submission periods, the difference between the technology types was not statistically significant (p-value > 0.10).

#### *Overall Effect of Submission Period*

The two samples were divided into three sub-samples each made up of 70 reports submitted during a specified time period: 1988-1989, 1990-1991, and 1992-1993. The data for each submission period from both the advanced technology and the traditional technology aircraft were combined. The  $\chi^2$  test was used to determine if the differences between the submission periods were significant. The  $\chi^2$  value was 6.891 at 2 degrees of freedom with a p-value of 0.032. Using a significance level of  $\alpha = 0.05$ , it was concluded that this difference was statistically significant.

#### *Effect of Submission Period on Advanced Technology Sample*

The data from the advanced aircraft only was used, and the  $\chi^2$  value was calculated to compare the three submission periods. The  $\chi^2$  value was 5.522 at 2 degrees of freedom with a p-value of 0.063. This would be significant at  $\alpha = 0.10$ .

#### *Incident Report Submission Period - Effect of Submission Period on Traditional Technology Sample*

The same approach taken in analyzing the advanced technology sample frequency data by submission period was used to analyze the traditional technology data. The result was not statistically significant (p-value = 0.423).

## DISCUSSION

### *Primary Objective*

The primary objective of this study was to begin evaluating the relationship between TM of commercial airline pilots and the level of automation on the flight deck by determining how automation affects the frequency of Task Prioritization errors as reported in ASRS incident reports. We found that Task Prioritization errors occurred in both advanced technology and traditional technology aircraft, and that overall there was a statistically significant difference between the number of reports classified as containing a Task Prioritization error in the advanced and traditional technology aircraft. This difference in the frequency of Task Prioritization errors suggests that Task Management may be more difficult in the advanced technology aircraft.

We cannot unequivocally state that the difference was caused by the nature of the design of the automation because this is confounded by the novelty of the advanced aircraft in air carrier fleets. In an attempt to better understand the effect of aircraft technology type, we looked more closely at the difference by submission period between the advanced and traditional technology samples. However, we found that the difference by submission period between aircraft technology was not statistically significant. Why would this be the case? The answer is in the power of the statistical test. For the overall test in which the three submission periods' frequency data were combined for the two technology types, the power of the test was such that a medium-small effect could be detected. For the tests conducted by submission period, however, the power of the test was such that a medium-large effect could be detected. This difference in effect size detection was due to the difference in sample size. In the population of ASRS incident reports, the actual effect that we were trying to detect was smaller than medium-large and therefore the test by submission period lacked the appropriate power to detect it. To determine if there was a significant difference between aircraft technology in each submission period, the sub-sample size would have needed to be increased.

We also looked at the effect of submission period on Task Prioritization errors. By separating the two samples into three equal sub-samples based on submission period, a decrease in the frequency of Task Prioritization errors in both the advanced technology sample and the traditional technology sample over time became apparent. This difference was statistically significant for the advanced technology sample; however, it was not statistically significant for the traditional technology sample. These data are consistent with the idea that industry experience with the advanced technology aircraft played a role in the differences in the frequency of Task Prioritization errors, but this cannot be stated conclusively. It may be the case that improved pilot training programs, or any number of other factors could have contributed to this reduction in Task Prioritization errors and that this reduction may have occurred in all aircraft, regardless of their level of technology. Further research is required to determine if the novelty of the advanced aircraft indeed played the critical role in creating the difference of frequency of Task Prioritization errors between the two aircraft types.

When evaluating the results of this study, one must bear in mind the limitations of ASRS incident report data. The samples used in this study were drawn from a nonrandom sample of events occurring in aviation operations and the ASRS incident reports themselves reflect reporting biases. What can be said with confidence, however, is that Task Prioritization errors do exist in actual line operations and their existence warrants thoughtful consideration. This study shed some light on one factor, automation on the commercial flight deck, which may effect the frequency of these errors.

#### *Secondary Objective*

The secondary objective of this study was to create an effective methodology for using ASRS incident reports for inferential analysis. By carefully constructing a research question and choosing an appropriate statistical test, an inferential analysis was conducted on the data collected. In this study statistically significant results were derived, supporting the notion that ASRS incident reports can be effectively used both for descriptive analyses and for inferential analyses.

By using ASRS data, we took advantage of several of the strengths of this type of data. First, the reports were able to provide a practical alternative to collecting data from the jumpseat of a commercial aircraft. The situations described in the narratives of the reports represented situations that had occurred in line operations that gave this study ecological validity and

avoided the possibility that the effect found was an artifact of a laboratory experiment. Second, the large number of incident reports available made it possible to construct a study with a large enough power to detect a medium-small effect.

#### CONCLUSION AND RECOMMENDATIONS

While Task Prioritization errors occur in both advanced technology aircraft and traditional technology aircraft, these errors occur more frequently in the advanced technology aircraft. The increased frequency of Task Prioritization errors suggests that Task Management may be more difficult in advanced technology aircraft. The submission period effect suggests that there is a downward trend in Task Prioritization errors in advanced technology aircraft.

Based on these conclusions, there are two recommendations that we would like to make. First, we recommend that further research be conducted to differentiate the effect of automation due to the nature of its design and the effect of automation based on its novelty in air carrier fleets. One way this could be accomplished is by analyzing additional submission periods and adding these data to the results presented here. The results of such a study could also be used to determine if the overall downward trend of Task Prioritization errors that appeared in this study continues.

Second, we recommend that when designing a training program for pilots of advanced aircraft that this information be disseminated to the pilots. The information could raise the awareness of pilot's susceptibility to Task Prioritization errors in advanced technology aircraft. It is possible that a heightened awareness could counteract this susceptibility.

#### ACKNOWLEDGEMENTS

This work was supported in part by the Federal Aviation Administration, Office of the Chief Scientific and Technical Advisor for Human Factors (AAR-100) under grant 93-G-039 titled Comparative Analysis of Flight Deck Automation. During the course of this project, the three technical monitors for the grant project have been John Zalenchak, Tom McCloy, and Eleana Edens. Also, we would like to thank Beth Lyall for her valuable contributions during this study and for her review of this paper.

## REFERENCES

- Aeronautica Civil Of The Republic Of Colombia (1997). Aircraft Accident Report: Controlled flight into terrain, American Airlines flight 965, B757-223, near Cali, Colombia, December 20, 1995.
- Alter, K.W., & Regal, D.M. (1992). *Definition of the 2005 flight deck environment*. NASA Contract Report 4479. Washington, DC: National Aeronautics and Space Administration.
- Boeing Commercial Airplane Group (1997). Statistical summary of commercial jet airplane accidents, worldwide operations, 1959-1996. Seattle, WA: Boeing Commercial Airplane Group, Airplane Safety Engineering.
- Chappell, S.L. (1994). Using voluntary incident reports for human factors evaluations. In N. Johnston, N. McDonald, & R. Fuller (Eds.), *Aviation Psychology in Practice*. Brookfield, VT: Avebury Technical. pp. 149-169.
- Chou, C., Madhavan, D., & Funk, K. (1995). Studies of cockpit task management errors. *International Journal of Aviation Psychology* 6(4), pp. 307-320.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Funk, K., Lyall, B., Wilson, J., Niemczyk, M., Vint, R., Suroteguh, C., & Owen, G. (in review). Flight deck automation issues. *International Journal of Aviation Psychology*.
- Johnston, W.A., Hawley, K.J., Plewe, S.H., Elliot, J.M.G., & DeWitt, M.J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General*, 119, pp. 397-411.
- Latorella, K.A. (1996). Investigating interruptions: An example from the flightdeck. In Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting held in Philadelphia, pp. 249-253.
- NTSB. (1973). *Aircraft accident report. Eastern Air Lines, Incorporated, L-1011, N310EA, Miami, Florida, December 29, 1972*. Report No. NTSB-AAR-73-14. Washington, DC: National Transportation Safety Board.
- NTSB. (1975). *Aircraft accident report. Trans World Airlines, Incorporated, Boeing 727-231, N54328, Berryville, Virginia, December 1, 1974*. (Report No. NTSB-AAR-75-16). Washington, DC: National Transportation Safety Board.
- Ministry of Civil Aviation, India. (1990). *Report on accident to Indian Airlines Airbus A-320 Aircraft VT-EPN at Bangalore on 14<sup>th</sup> February 1990*. Government of India, Ministry of Civil Aviation.
- Rogers, W.H. (1996). Flight deck task management: A cognitive engineering analysis. In Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting held in Philadelphia, pp. 239-243.
- Wiener, E.L. (1989). Human factors of advanced technology ("glass cockpit") transport aircraft. NASA Contract Report 177528. Moffett Field, California: NASA-Ames Research Center.
- Wilson, J.R. (1998). *The Effect of Automation on the Frequency of Task Prioritization Errors on Commercial Aircraft Flight Decks: An ASRS Incident Report Study*. Unpublished thesis, Oregon State University.