Automation, Task, and Context Features: Impacts on Pilots' Judgments of Human–Automation Interaction

Kathleen L. Mosier, San Francisco State University, **Ute Fischer**, Georgia Institute of Technology, **Daniel Morrow**, University of Illinois at Urbana-Champaign, **Karen M. Feigh**, Georgia Institute of Technology, **Francis T. Durso**, Georgia Institute of Technology, **Katlyn Sullivan**, Emerson Corporation, and **Vlad Pop**, Georgia Institute of Technology

Human–automation interaction (HAI) takes place in virtually every high-technology domain under a variety of operational conditions. Because operators make HAI decisions such as which mode to use, and when to engage, disengage, monitor, or cross-check automation, it is important to understand their perceptions of how system and situational characteristics affect their interaction with automation. The objective of this study was to examine how systematic variations of automation interface, task and context features influence professional pilots' judgments of HAI situations. Pilots received descriptions of crews interacting with flight deck automation in specific situations and were asked to rate cognitive demands and predict behaviors. Results reflect a complex interplay among automation features, task, and context. Automation features influenced judgments of workload, task management, and potential for automation-related errors; however, the impact of automation on situation awareness seems to be moderated by task features. Unanticipated tasks had broader effects on pilots' judgments than operational stressors. Results suggest that although changes to automated systems may be small in technical terms, their cognitive and behavioral impact on operators may be significant. Performance effects of automation changes in aviation as well as other domains need to be addressed with reference to task characteristics and situational demands.

Keywords: HAI, human–automation interaction, automation, context features

Recent advances in computing have made it possible for many of the tasks once performed by human operators to be performed completely or partially by machine. Nowhere is this more true than in the modern aircraft cockpit. Human–automation interaction (HAI) takes place in virtually every phase of flight under a variety of environmental conditions and involves more than the moment when pilots physically touch an input key or turn a selection dial (Sheridan & Parasuraman, 2006). Pilots plan ahead, instruct the flight management system, monitor flight progress, correct and modify flight paths, comply with ATC (air traffic control) instructions, and learn input/output relationships as they gain experience with an aircraft. Pilots decide when to engage and disengage automation, which flight mode or level of automation to use, the extent to which they should monitor and cross-check the automation, and so on.

Flight systems are likely to become more automated as the national airspace system modernizes to accommodate anticipated increases in air traffic (Joint Planning and Development Office, 2007). In fact, acquisition of the most sophisticated systems will be a competitive requirement, as the most fuel-efficient paths will be assigned to those aircraft with required navigational performance capabilities. Participation in procedures such as reduced vertical separation minimum, tailored arrivals, and continuous descent approaches will entail precise autopilot flight.

Designers and implementers have long touted automation's potential to reduce workload and support situation awareness while improving efficiency and maintaining high levels of safety.

However, human factors professionals have recognized issues associated with the use of complex automated systems for some time (e.g., Billings, 1996; Sarter, Woods, & Billings, 1997). Early implementations of aviation automation, for example, were "clumsy" as they reduced pilots' workload in situations when it was already low and increased it when it was high (Wiener, 1989). The same phenomena were found in medicine, for instance as experienced anesthesiologists tried to overcome clumsiness in new physiological monitoring systems by adapting the systems to minimize their cognitive work, often disabling some of the features the system was intended to provide (Cook & Woods, 1996). In a similar vein, Wein (1996) noted that examples of poor technological design abound, "from the familiar, such as VCRs that no one can program and vending machines whose interfaces frustrate and confuse, to the more exotic, such as the devices used to operate airplanes and nuclear reactors, the poor design of which contributes regularly to catastrophes of 'human error'" (p. 376). Designing and planning new automation for any high-technology operational domain requires recognition and revisiting of issues related to HAI, as well as the ability to predict behavioral and cognitive consequences of new interfaces and more sophisticated automated systems.

Because operators make many decisions concerning automation use, it is important to understand their perceptions of how system and situational characteristics will affect their interaction with automation. Experienced pilots, for instance, develop an understanding of how difficult it is and how long it takes to implement actions using different systems and levels of automation, and whether or not automation will enhance efficiency. HAI benefits as well as problems have often been attributed to design features that facilitate or hinder performance, but they may also stem from characteristics of operators, task features, and constraints of the operational environment. For example, the presence of traffic may affect the extent to which pilots interact with automation and the level of automation they choose. Operational features such as time pressure, weather, and terrain may also change pilots' automation strategies, as may

individual variables such as experience or fatigue. Pilots' strategies for automation use will to a great extent be dependent on their perceptions of how a particular automated system in a particular situation will affect their cognitive and behavioral performance. An understanding of the way automation, task, and context features interact to influence pilots' strategies of automation use is therefore critical to system design as well as automation training.

The current research was undertaken to explore pilots' perspectives of the relationship between situations involving pilot–automation interactions and important cognitive and behavioral consequences as defined by the literature (HART Group, 2011a). The study was part of a broader modeling effort intended to describe a function relating situations containing pilot–automation interaction to cognitive and behavioral consequences of that interaction (see Sullivan et al., 2013). HAI situations were conceptualized as specific configurations of automation features (clumsy vs. efficient interface), operator characteristics (professional as opposed to novice pilots), and operational and task constraints (time pressure and task disruptions). Consequences examined were cognitive and psychological states such as workload, awareness of automation functioning and situation awareness, as well as predicted behaviors, for instance, cross-checking of automation, task management, and the likelihood of automation-related errors. Although the situations in this study were aviation specific, the variables examined are common to other dynamic and increasingly high-technology domains (e.g., medicine; Millitello, Patterson, Saleem, Anders, & Asch, 2008; Morrow, North, & Wickens, 2006).

CLUMSY VERSUS EFFICIENT AUTOMATION

The introduction of the flight management system (FMS) in the cockpit was a radical innovation and was key to the execution of precise navigational paths. However, features of the design and interface of the FMS render it "clumsy" and have resulted in difficulties ever since it was adopted (e.g., Wiener, 1989). Programming the FMS entails a variety of cognitive and behavioral activities, such as selection

of the appropriate flight mode and multiple keystrokes for input, often on various pages (e.g., Casner, 1994; Sarter & Woods, 1994). Although newer systems sought to mitigate past problems, current design characteristics are still deficient, most notably with respect to system *observability*—the extent to which the automation enables the operator to understand its logic and see its current and future activities and targets—and system *directability*—that is, the ease with which the system's design enables the user to intervene when required (Christoffersen & Woods, 2002; Ferris, Sarter, & Wickens, 2010).

Observability is hindered when system displays involve many layers, or when subsystems are tightly coupled and feedback on system states is missing or obscure. Low observability has been linked to poor automation and situation awareness, increased workload, and mode errors (e.g., Sarter, 2008; Sarter et al., 1997; Sarter & Woods, 2000). The impact of poor observability is exacerbated when operators do not have a good mental model of how automation works and the system does not provide sufficient information about its functioning (Javaux, 1998, 2002). Moreover, when more than one automated system can be used to accomplish the same task, operators may be unsure of the best way to manage the task and may select a method that is not the most effective given the situation and their goals (Ferris et al., 2010).

Directability of automated systems may be impaired by a number of physical control features. For instance, the spatial configuration of input devices may limit their accessibility and thus may delay critical input or lead to input errors, as with multifunction knobs that use push versus pull inputs to trigger different actions. A case in point is an incident submitted to the Aviation Safety Reporting System (ASRS) that involved the pilot accidentally and unknowingly triggering the TOGA (takeoff/go-around) switch on the engine power lever while attempting to disconnect autothrust (ASRS Report No. 888422; also see Aircraft Accident Investigation Commission, 1996, in Ferris et al., 2010). Once an error is made, poor directability may also hinder recovery—especially when combined with inadequate system mental models. This was a factor in a 1994 accident in Nagoya, Japan (see Billings, 1996). The crew inadvertently engaged go-around mode on their A300. They were not able to override the autopilot via manual control,

> . . . which in all other aircraft—and in this aircraft in all modes except the approach mode—would normally disconnect the autopilot. However, in this particular aircraft and in this particular mode, the autopilot had to be manually deselected and could not be overridden by control column inputs. (Sheridan & Parasuraman, 2006, p. 98)

Complicated input requirements and coupled automated systems have been implicated in other aviation incidents, presumably because these design features increased pilots' workload and distracted them from other pertinent flight tasks (Dodd, Eldredge, & Mangold, 1992; Lyall, Niemczyk, Lyall, & Funk, 1997). The FMS often requires multiple inputs to make changes to the flight path, increasing workload, head-down time, and the potential for error while changes are entered. In an ASRS incident report modified for this study, for example, the pilots were required to perform several steps to delete a speed restriction while on approach for landing into Phoenix (PHX):

> Arriving into PHX on the GEELA 2 Arrival, we were cleared to descend via the arrival and then told to delete speed restrictions. You can't just delete the speed only on the CDU, so I deleted the line and then retyped the altitude. I believe I might have put in 11,000' instead of 10,000'. Five miles later, as we were reaching 10,000 feet, the Controller asked our altitude. She mentioned we were supposed to be there a few miles ago. I think if ATC is going to change any of the Descend Via [clearance] and not allow you to just fly it as published, they should revert to altitude and airspeed directives. I also believe I won't be retyping altitudes on arrivals just using the airspeed override button. (ASRS Report No. 892363)

Newer models of the FMS have simplified the input process to a certain extent, and implications of some of these changes are also examined in

this study. For instance, some FMS models permit the deletion of a speed restriction without deleting other parameters. In addition, datalink technology in aviation will change the FMS interface significantly by enabling one-step entry of ATC text-based clearances. This capability is predicted to be important to facilitate four-dimensional procedures such as continuous descent approaches (e.g., Coppenbarger, Mead, & Sweet, 2007) and required time of arrival (e.g., Wichman, Carlsson, & Lindberg, 2002), as well as to negotiate complex clearances required for trajectory-based operations with route and/or altitude changes specified by latitude and longitude coordinates (Mueller & Lozito, 2008).

There has been some debate related to the elimination of "extra" steps. An issue arising from datalink concerns the trade-off between the easy-entry feature for complex clearances and its impact on pilots' awareness of potential deficiencies in a clearance, as well as their understanding of what they are accepting and implementing. Riley (2001) has suggested that the functional logic and interface of the FMS should mirror the language of ATC. He contends that pilots using a new interaction design and interface based on his Cockpit Control Language could easily implement clearances using familiar terms and without extra cognitive work. Boorman and Mumaw (2004), however, posit that extra steps may serve a purpose in that they enable pilots to "reformulate," that is, to interpret an ATC clearance and the desired aircraft behavior before entering it into the autoflight interface (Sherry, Polson, Feary, & Palmer, 2002). The trade-off between ease of input and negative outcomes such as poor automation and situation awareness is still under investigation and is explored in this study.

Operator Variables

A number of individual variables may shape operators' interaction with automation, and different paths to the development of expertise may result in different HAI patterns and strategies. Many senior professional pilots, for example, began their flying careers in relatively low-technology aircraft and developed scanning habits with which they are comfortable as they worked in that environment. For those pilots,

monitoring automation may be perceived as an extra task—and one that is not especially critical given the high reliability of automated systems (e.g., Parasuraman & Manzey, 2010; Parasuraman, Molloy, & Singh, 1993). Some current pilot training programs, however, bypass the traditional progression from low-tech to high-tech aircraft. Multicrew pilot license training, for example, focuses on "preparing the co-pilot candidate for the right seat of an advanced airliner, using a competency-based approach to training developed with an emphasis on improving flight deck safety" (Schroeder & Harms, 2007). Pilots who attain this license are likely to exhibit different HAI patterns and preferences than those trained in more traditional programs.

Other individual pilot variables have been cited as factors in aviation incidents and accidents. In particular, fatigue was found to slow pilots' system monitoring and impair their understanding of system indications and may lead to overreliance (Dismukes, Berman, & Loukopoulos, 2007). In fact, recognition of the potentially adverse effects of fatigue led the FAA to release a final rule amending duty and rest requirements for pilots in commercial Part 121 operations (Federal Aviation Administration, 2011). However, in some cases fatigue may cause operators to be more vigilant and may ameliorate problems, such as automation bias (Mosier & Skitka, 1996). Manzey and his colleagues, for example, used a simulated life support system task to examine automation bias and complacency as a function of fatigue and found that fatigued operators were less susceptible to automation bias and complacency and displayed more careful information checking with a decision aid than alert operators (Manzey, Reichenbach, & Onnasch, 2009).

Operational/Task Features

Features of the operational environment will also affect HAI as well as preferred HAI strategies. *Time pressure*, a common feature of many dynamic domains, has been linked to pilots' truncated information search when dealing with automation anomalies, to diagnosis errors, and to a focus on salient cues to the exclusion of other, less salient sources of information (Khoo & Mosier, 2008; Mosier, Sethi, McCauley,

Khoo, & Orasanu, 2007; Wickens, Stokes, Barnett, & Hyman, 1993). Time pressure coupled with poor automation design can severely hamper pilots' ability to deal with HAI events. For example, in their analysis of the TWA 843 hard landing accident at John F. Kennedy airport, Dismukes et al. (2007) discussed in detail the characteristics of the Lockheed-1011 stall warning system—which malfunctioned on takeoff. The system suppressed stall warnings on the ground, with

> . . . the unintended consequence of allowing a false stall warning indication to . . . first be annunciated to the crew just after the airplane lifts off the ground. This design forces pilots to evaluate and attempt to deal with a stall warning under extreme time pressure, when workload is high and they have little time to ascertain whether the warning is false. (p. 27)

The influence of time factors and time pressure on automation use has also been documented during observations in the aviation operational environment. Casner (1994) observed professional flight crews flying between San Francisco and Los Angeles and recorded crews' automation strategies. He found that pilots' choice of automation varied depending on the characteristics of ATC clearances, in particular their predictability, frequency, complexity (i.e., number of clearances issued at the same time), and time constraints. Crews tended to engage the FMS for more predictable clearances and those with a more relaxed time frame, whereas autopilot or manual control were used for unexpected clearances or those involving some time pressure. Casner noted that although pilots preferred to use the FMS for navigation, being required to program the computer for unpredictable time constrained clearances had consequences in terms of cognitive and perceptual-motor workload.

Tasks in dynamic operational environments are also commonly characterized by *distractions* and *interruptions* from internal or external sources, which can derail the flow of HAI. When pilots are interrupted during the performance of a checklist, for example, they may forget where they stopped and which items have been accomplished (Dismukes et al., 2007; Latorella, 1999). New clearances that involve reprogramming of automated systems can also increase workload, especially when they occur during already busy periods. A last-minute runway change for take-off or landing, phases that are already high in workload, may require entry of the new runway and departure or arrival paths into the FMS, putting additional time pressure on the crew (Lindeis, 2010).

The effects of interruptions on stress and workload are well documented, and because of their negative impact, interruption management is a feature of some new automated systems. Dorneich and colleagues, for example, designed the Communications Scheduler, a computer system for soldiers that detects physiological signs of high workload (using EEG and ECG measures) and adjusts the format and alert levels of messages so that low-priority messages are delivered as text messages, enabling attention to high-priority radio messages (Dorneich, Ververs, Mathan, Whitlow, & Hayes, 2012). They found that the Communications Scheduler did improve primary task performance—but at the cost of reduced situation awareness on other tasks. Results such as these demonstrate the need to examine HAI as the interaction among automation, operator, task, and context to illuminate unintended and potentially negative consequences.

Behavioral and Psychological Consequences in HAI

In the current study, we examined how the interaction of automation features and task and context variables discussed above affected pilots' predictions of six behavioral and psychological consequences identified as critical HAI outcomes: workload demands, effort involved in task management, required cross-checking of automation, level of automation awareness, level of situation awareness, and likelihood of an automation-related error.

Task management and workload. When possible, operators are likely to use automation in ways that make their jobs easier and may bypass automation when it is inefficient or increases workload. Kirlik (1993) adopted a modeling

approach to explore how task-context and automation design parameters affected strategies in automation use. Specifically, a Markov decision process was employed to model automation use as it related to three autopilot design features—the ability of the autopilot relative to the crew's ability to control the helicopter, and the time it took to either engage or disengage the autopilot—and three task context features—the duration, time criticality, and frequency of secondary editing tasks. This modeling effort was prompted by an earlier lab study finding that crews did not use the autopilot for its intended purpose—as a task-offload aid (Kirlik, Plamondon, Lytton, Jagacinski, & Miller, 1993); that is, pilots did not consistently delegate the flying task to the autopilot when engaged in a secondary task. The model suggested that the reason for pilots' behavior was a mismatch between the intended strategy for autopilot use and characteristics of the autopilot (long engagement time and constraints on its ability to fly as fast as with manual control) and the task context (frequent and time critical editing tasks). Kirlik concluded that crews did not rely on the automated aid as the designers intended because it was simply not efficient to do so. Likewise, research on automation use in medicine has identified design features that limit the effectiveness and efficiency of automated aids, thus reducing the likelihood that they will be used as intended. Morrow et al. (2006) discussed decision support systems in medicine and noted limitations with even well-designed tools. Many medical diagnostic tools, for instance, are unable to integrate information about symptoms with other more general medical knowledge (Gawande & Bates, 2000) or to account for context-specific patient information (Patterson, Nguyen, Halloran, & Asch, 2004).

Olson and Sarter (2000) surveyed pilots' preferred automation task management strategies in the context of a datalink system and found that pilots generally disliked autonomous automation that required them to override system actions and instead preferred automation that required them to consent to proposed actions by automation, either by pressing a single button or by accepting different elements in the FMS, mode control panel (MCP), or both. However,

pilots' preferences shifted in the presence of certain task and situation variables. In situations that were high in time pressure and task complexity, or involved low task criticality, pilots favored simple automation management strategies—that is, strategies where they could accept or override automation with one button press—apparently in an effort to ensure a swift response while minimizing demands on their attentional and cognitive resources.

Strategies that reduce workload by, for example, facilitating directability may inadvertently reduce automation and situation awareness. Olson and Sarter (2001) compared two management-by-consent automation options during simulated approaches with datalink communication: a manual clearance entry versus a "gating" option of datalink by which the pilot's button-press approval of a datalink clearance routes it directly to the FMS. The gating option was hypothesized to reduce data entry errors and workload; however, performance data showed that pilots in both conditions were generally poor in detecting clearances that conflicted with their goals or current aircraft status. Moreover, when an acceptable clearance was incorrectly implemented by the automation, pilots in the gated condition were least likely to detect this problem. Proposed procedures to ensure comprehension of datalink clearances include some sharing of communication responsibilities so that both pilot flying and pilot monitoring must examine the clearances, review and discuss clearance data, or print out clearances prior to acceptance (Mueller & Lozito, 2008). Although these procedures may decrease the potential for accepting erroneous clearances and improve automation awareness, they will likely affect other cognitive and behavioral variables, such as the crew's workload and task management requirements. As skilled operators adapt their automation management strategies according to automation features and situational and task demands (Durso & Alexander, 2010), their behavior may compound inadequacies of system design and lead to undesirable consequences, most notably increased workload or inaccurate or incomplete automation and situation awareness (Dekker & Orasanu, 1999; Dismukes et al., 2007; Durso & Alexander, 2010; Orasanu, Martin, & Davison, 2001).

Automation and situation awareness. Task management for operators of automated systems involves monitoring the situation, including the instruments that reflect progress status changes and the automation that has been delegated to aid progress. Monitoring of automated systems has been a topic of much research and concern since the introduction of these systems (e.g., Parasuraman et al., 1993). When performing multiple tasks, operators have been found to pay scant attention to those tasks "delegated" to automated systems, rarely cross-checking automation against other system indicators, and as a consequence to have low awareness of the functioning and/or malfunctions of the automation (Parasuraman & Manzey, 2010). The tendency toward complacency with respect to automation functioning and the accompanying loss of awareness may be exacerbated by the complexity of multiple-mode systems.

In work looking specifically at mode monitoring, Sarter, Mumaw, and Wickens (2007) recorded eye movements during a 1-hour B747-400 simulation of a flight from takeoff to landing. Pilots showed a strong preference to look at the "raw data," that is, the data from heading indicator, altimeter, and airspeed indicator, rather than attending to the automation-interpreted data from the flight mode annunciators (FMAs). This emphasis on raw data suggested that pilots might miss changes associated with mode transitions. Consistent with this suggestion, Sarter and colleagues found that only about half of the time did pilots fixate on the FMA within 10 seconds following a mode transition (i.e., when the green alert box was on; also see Björklund, Alfredson, & Dekker, 2006). Even when pilots did fixate, they usually failed to recognize potential problems associated with the change, suggesting that their information processing was not of sufficient depth. These monitoring and awareness issues are complicated by the fact that many of the pilot participants had faulty or incomplete mental models of the vertical navigation modes. The Sarter et al. (2007) results are consistent with those of Javaux (1998), who demonstrated that a combination of overly complicated mode transitions and frequency bias created by uneven usage of the modes over time leaves pilots with only a partial understanding of the overall system.

Likelihood of automation-related errors. Systems with multiple modes are particularly problematic in terms of observability and directability, thus increasing the probability of automation-related errors. In medicine, for example, hidden modes of operation combined with inconsistent mappings between signals and actions and misleading displays were found to undermine anesthesiologists' performance (Cook, Woods, Howie, Horrow, & Gaba, 1992). Mode awareness and mode management have also caused problems on flight decks, as the proliferation and layering of system modes increases complexity, hinders the development of accurate mental models, and increases the chances of programming mistakes (Javaux, 1998; Sarter, 2008).

Vertical flight navigation has often been cited as an example of complexity combined with low observability, and was thus highlighted in this study. Vertical flight control may entail use of the VNAV (vertical navigation) mode of the FMS, or the flight level change or vertical speed modes on the MCP. Sherry and his colleagues (Sherry, Feary, Polson, Mumaw, & Palmer, 2001; Sherry, Feary, Polson, & Palmer, 2001) decomposed the functions and displays of the VNAV system and found that the VNAV button is "overloaded" in the descent and approach phases of flight, in that its selection results in the engagement of one of six possible trajectories. Moreover, these trajectories will change autonomously as the situation evolves. This complexity is particularly problematic as the approach and landing phases of flight already involve high workload and afford pilots with the least time to deal with an ambiguous interface.

Confusion with vertical navigation extends to vertical speed or flight level change modes. Degani and Heymann (2002) demonstrated that the vertical flight control system can trip pilots up as they change altitudes. The system behaves differently according to the relationship between the input altitude and a specific "critical" altitude, and the current mode of the aircraft. For example, if the autopilot has already entered capture mode for a previously set altitude, it may or may not acquire a newly set altitude depending on the new value relative to the capture start altitude. The interface, however, does not signal to the pilot whether or not altitude capture will

occur. This confusing behavior makes it difficult for pilots to form an accurate mental model of the aircraft's vertical flight control and has led to many so-called "altitude busts," that is, situations in which aircraft fly through a target altitude. Altitude busts can increase stress and workload, as crews worry about traffic and regulatory issues associated with being at the wrong altitude.

In an ASRS narrative that was used in this study, the reporting pilot thought he had set up the vertical navigation automation correctly for approach to Los Angeles (LAX), but was actually in what Degani and Heymann (2002) call an *error state*:

> We were on the SEAVU arrival into LAX. ATC cleared us to cross KONZL at 17,000 FT. I was using vertical speed because VNAV had been inconsistent earlier during the flight. Approximately 5 NM from KONZL, as we were approaching 17,000 FT but not yet in altitude hold, ATC simultaneously changed our runway assignment and cleared us for the descent via the arrival. I entered the next altitude in the window as the Captain reprogrammed the FMS for the new runway. A few seconds later, as we reviewed the new runway information, I realized we were descending through 16,700 FT and still 2.4 NM from KONZL. I quickly disengaged the autopilot and corrected to 17,000 FT as we crossed KONZL. I believe the altimeter touched 16,600 during my correction. We checked the TCAS and saw no aircraft within 5 NM at any altitude. ATC didn't mention our altitude deviation. My use of vertical speed combined with the altitude change just prior to level off (and therefore prior to ALT CAP) allowed vertical speed to briefly fly through our assigned altitude. (ASRS Report No. 883679)

Even when information is theoretically observable, a separate issue is whether the operator will actually know what information to look for and when to check it and will comprehend and understand its significance in the current flight mode. A small or nonsalient feature in a display, although technically available, may be difficult or confusing to process. Pritchett (2009) discusses the design of an MCP in which the "same" numeric information results in vastly different aircraft behavior dependent on the selected mode (flight path angle vs. vertical speed). The critical mode information is provided in two- or three-letter acronyms, which pilots may overlook, especially during high-workload phases of flight or when pilots expect a particular mode setting (Johnson & Pritchett, 1995). A tragic example of mode error is an accident that occurred in Strasbourg, France, when pilots of an Airbus A320 set a vertical descent rate of −33 (3,300 ft/min) rather than a path angle of −3.3 degrees and crashed in mountainous terrain several miles short of the airport (in Billings, 1996).

The automation features, operator characteristics, and task and context variables discussed above interact with each other to produce performance enhancements—such as reduced workload, higher situation awareness, and increased efficiency and preciseness—but may also lead to unintended consequences—such as higher workload during some tasks or phases, reduced automation awareness, and increased likelihood of automation-related errors. In this study, we examined professional pilots' perceptions of how these interactions would likely play out in specific operational situations and with respect to specific automation, task, and context variables.

Automation features manipulated in the current study concern some of the most complex and prevalent flight mode issues—vertical flight management via flight level change versus the more comprehensive VNAV—and methods for input to the FMS for flight path changes—manual or multistep versus autoload or one-step. We varied the level of automation available—VNAV versus FLCH, or flight level change—ease of input to the FMS, time pressure, and task disruptions, and asked professional pilots to make judgments about resultant outcomes. Tasks in the current study included unexpected commanded changes (by ATC) to runway or speed, as well as communications concerning other aircraft (traffic) during high-workload periods. Time pressure was manipulated by varying the altitude and distance to the airport at which these interruptions and unexpected commands occurred.

METHOD

This scenario-based paper-and-pencil study was conducted at a national meeting of the Airline Pilots Association (ALPA). Scenario packets were distributed and collected during an annual ALPA Air Safety Forum in Washington, D.C.

Participants

All participants were professional pilots. In all, 173 surveys were handed out, of which 62 (36%) were returned. The majority of respondents had experience on Boeing ($n = 36$) or Airbus ($n = 13$) aircraft; the remaining respondents were CRJ or Embraer pilots ($n = 12$) or did not provide this information ($n = 1$).

Materials

Scenarios. Two incident descriptions were selected from a set of 12 ASRS reports that had been used in a previous study (HART Group, 2011b). Scenarios focused on automation-related incidents during approach because this is a challenging phase of flight as a result of traffic density in terminal areas and likely task disruptions. One incident concerned an approach into Las Vegas (LAS); the other one occurred during an approach to Los Angeles (LAX). Selection of these two incidents was based on the following considerations: Each incident revolved around an unexpected change—an unexpected cancellation of speed restrictions in the former case, and an unplanned change in runway assignment in the latter—that required extensive pilot–automation interaction. Moreover, these unexpected changes were not specific to a particular aircraft type or make, concerned tasks that pilots frequently encounter during approach, and could be achieved using different automated systems or aspects of automation. We edited each report to create a "parent" scenario that described a normal approach to LAS or LAX, that is, the approach followed the flight plan as filed and thus did not include the unexpected change that prompted the original ASRS report. Then, in consultation with two subject matter experts (commercial pilots involved in crew training), four scenario variants ("family" members) around each unexpected change were created by systematically varying automation characteristics and features of the operational context.

Automation features pertained to aspects of the flight control system and varied how a crew implemented changes to the flight plan, specifically how many steps were required to input a new clearance to the FMS. Each scenario family contained the parent scenario and four variants with context and automation features crossed. To limit the number of variables and thus keep the required number of scenario variants at a manageable level, we did not vary pilot characteristics in this study.

Context variables concerned the presence or absence of time pressure and task disruptions. These features were taken from research examining the impact of task and context variables on human performance in general (esp. Hollnagel, Kaarstad, & Lee, 1999; Kim & Jung, 2003) and in past research had been identified as critical factors in aviation incidents and accidents (e.g., Loukopoulos, Dismukes, & Barshi, 2001; Orasanu et al., 2001). Moreover, these features either were present in the original ASRS reports or could be incorporated into a core situation to create realistic scenario variants.

The LAS family (see Table 1) comprised scenarios with two types of FMSs, differing with respect to the number of steps needed to delete a speed restriction that ATC lifted unexpectedly during approach (unexpected task). In situations involving the "clumsy" system, the removal of speed restrictions on the approach page required pilots to delete the line specifying both speed and altitude constraints and then to reenter the altitude information. (This system was the central issue in the original ASRS report.) The alternative "easy" system called for one step to complete the targeted deletion of the speed restriction. Both types of systems are currently in use in commercial operations. Scenarios also varied concerning the presence or absence of two context variables—ATC was late in waiving the speed restriction (= time pressure) and a traffic advisory occurred as the pilot monitoring reprogrammed the FMS (= task disruption). See Appendix A for the LAS family of scenarios.

The LAX family consisted of two family subsets that involved different ways of managing a descent. In one family VNAV was inoperative and the crew used flight level change to implement altitude changes (as in the original ASRS report). In the other subset VNAV was working

TABLE 1: Characterization of LAS Scenarios

	LAS_1 (Parent)	LAS_2	LAS_3	LAS_4	LAS_5
Unexpected task (*new speed issued*)	No	Yes	Yes	Yes	Yes
Automation (*FMS update needs multi- vs. one step*)	N/A	Clumsy FMS	Easy FMS	Clumsy FMS	Easy FMS
Time pressure (*ATC is late*)	No	No	No	Yes	Yes
Task disruption (*traffic advisory*)	No	No	No	Yes	Yes

TABLE 2: Characterization of LAX Scenarios

	LAX_1 (Parent)	LAX_2	LAX_3	LAX_4	LAX_5
Unexpected task (*runway change*)	No	Yes	Yes	Yes	Yes
Automation (*current FMS vs. FMS with datalink autoload*)	N/A	Current FMS	Future FMS with datalink	Current FMS	Future FMS with datalink
Time pressure (*ATC is late*)	No	No	No	Yes	Yes
Task disruption (*parallel traffic*)	No	No	No	Yes	Yes

Note. The variants were the same in both the VNAV working/inoperative subsets.

and was used for altitude changes. Participants responded to only one subset. Variants in both subset families (see Table 2) revolved around an unexpected task (runway change) that crews handled either by reprogramming the FMS according to current procedures or by automatically uploading the clearance via datalink (a "gated" future system similar to the one described by Olson & Sarter, 2001). As in the LAS family, contextual factors were crossed with type of automation and included ATC issuing the runway change at a relatively low (vs. high) altitude (= time pressure) and the presence/absence of parallel traffic (= task disruption). See Appendix B for the LAX families of scenarios.

HAI consequences. Based on a literature review (HART Group, 2011a), a focus group of human subject specialists, and interviews with pilots, 11 critical HAI consequences or outcome variables were identified. Pilots' responses during pretesting revealed that these 11 outcomes should be reduced to 6 to avoid high intercorrelations

(HART Group, 2011b). Pilots during pretesting also reported familiarity with HAI outcomes as these concepts are integral elements of pilot training. The final list of HAI consequences concerned workload demands, effort involved in task management, required cross-checking of automation, level of automation awareness, amount of situation awareness, and likelihood of an automation-related error.

Procedure

Two scenario packets were compiled. Each packet consisted of the LAS family and one subset of LAX scenarios, with VNAV either working or inoperative as described above. For each family, the parent scenario was presented first, followed by the four variants presented in one of five random orders. To mitigate any order effects, pilots were encouraged to "refer to and revise your previous ratings as you respond to subsequent scenarios in the set." Each scenario and associated HAI rating scales were printed on one page. Scenario packets provided the

appropriate Jeppesen approach plates for the scenarios and included demographic question-naires inquiring about respondents' position and time in current aircraft, the aircraft types/makes they were currently flying and had flown in the past, and their familiarity with the LAX and LAS airports. Pretesting had shown that scenario packets could be completed in about 30 minutes.

Instructions. Each family of scenarios was preceded by task instructions that emphasized the following points: First, all five scenarios in a set involve the same core event; differences between scenarios concern the piece of automa-tion a crew uses and the presence or absence of contextual factors. Participants should focus on these differences when making their judgments. Second, ratings should be made only on the basis of the information provided and should consider the crew as a whole rather than focus-ing exclusively on one crew position. Third, rat-ings may be revised in response to subsequent scenarios in a set. Fourth, participants should assume present-day equipment unless specified otherwise. Fifth, reference to present-day sys-tems for which Boeing and Airbus use different names will state both labels (such as FMS/FMGS), and participants should assume that the system of their current aircraft applies in the scenario.

Ratings. Participants rated each scenario version with respect to six automation interac-tion variables: *workload* ("How much mental workload would the crew experience?"), *task management* ("How much time or effort would be required to manage the tasks?"), *automation awareness* ("What level of awareness of what the automation is and will be doing would the crew have?"), *cross-checking automation* ("How much cross-checking of the automation would the crew do?"), *situation awareness* ("How much situational awareness would the crew have?"), and *probability of automation related error* ("How likely would it be for an automation related error to occur?"). Answers were given along 7-point scales with the extremes labeled as *very low* or *very little* and *very high* or *a great deal* according to the ques-tions being asked.

RESULTS

Separate analyses (doubly multivariate analy-sis of variance; Tabachnick & Fidell, 1996) were conducted for each scenario family because of sample size limitations and the fact that families tapped different automation characteristics. For the LAS family, scenario versions and their associated ratings of HAI variables were treated as a within-participants multivariate design. The LAX family, in addition, included a between-participants variable as pilots responded either to scenarios in which VNAV was working or scenario variants in which it was inoperative. An alpha level of .05 was used for all statistical tests. Wilks's lambda and Greenhouse–Geisser values are reported as appropriate. Bonferroni adjustments were used for multiple compari-sons. Means and standard deviations are shown in Table 3.

LAS Scenarios

A multivariate effect of scenario was observed on the combined dependent variables (pilots' ratings of workload, task management, auto-mation cross-checking, situation and system awareness, and likelihood of automation-related error), $F(24, 37) = 8.765$, $p = .0001$, $\eta^2 = .85$. This finding indicates that the different com-binations of automation features and task and context variables influenced pilots' judgments, as the combination of pilots' ratings differed significantly across scenario variants. Univari-ate tests further showed that pilots' ratings of all HAI variables varied by scenario (see Table 3), supporting pairwise comparisons of pilot ratings by scenario. Pairwise comparisons focused on the impact of task characteristics, automation features, and context variables and are discussed below. All reported differences were significant at the .05 level.

Effects of task characteristics. The compari-son of the normal approach situation ("parent scenario," LAS_1) with scenarios involving an unplanned speed change only (LAS_2, LAS_3) revealed that the unanticipated task increased ratings of workload, task management, and the possibility of automation-related errors, irre-spective of features of the automation involved in implementing the speed change. Effects on

TABLE 3: LAS Family—Means, Standard Deviations, and Univariate Test Results for Human–Automation Interaction (HAI) Ratings

HAI Variable	Scenario										
	LAS_1		LAS_2		LAS_3		LAS_4		LAS_5		
	M	SD	M	SD	M	SD	M	SD	M	SD	
Workload	3.51	1.4	5.36	1.02	4.52	1.41	6.13	1.23	5.72	1.23	$F(3.16; 189.41) = 66.01^{***}$, $\eta^2 = .52$
Task management	3.15	1.34	5.36	1.07	4.18	1.5	5.87	1.20	5.11	1.39	$F(3.42; 205.36) = 62.83^{***}$, $\eta^2 = .51$
Automation cross-checking	4.82	1.18	5.67	1.08	5.11	1.28	5.15	1.85	5.13	1.57	$F(2.55; 152.83) = 4.48^{**}$, $\eta^2 = .07$
Situation awareness	5.64	0.98	4.69	1.37	5.16	1.29	4.56	1.78	4.85	1.50	$F(2.61; 156.44) = 8.86^{***}$, $\eta^2 = .13$
Automation awareness	5.43	1.02	4.97	1.22	4.93	1.25	4.52	1.93	4.89	1.44	$F(2.31; 138.33) = 4.86^{**}$, $\eta^2 = .08$
Likelihood of error	3.13	1.34	4.97	1.43	3.82	1.57	5.59	1.48	4.87	1.53	$F(3.38; 202.95) = 45.08^{***}$, $\eta^2 = .43$

Note. N = 61. Greenhouse–Geisser adjustment used for violation of homogeneity of covariance.
**Significant at the .01 level. ***Significant at the .0001 level.

situational awareness and automation cross-checking ratings, on the other hand, were related to automation features. When the FMS required only one step to remove the speed restriction, raters judged that the unanticipated task would not reduce a crew's situational awareness or demand more automation cross-checking compared to the normal approach situation (LAS_3 vs. LAS_1). However, the FMS that involved multiple steps to implement the speed change was associated with more cross-checking of the automation and less situational awareness compared to the normal approach scenario (LAS_2 vs. LAS_1).

Effects of automation features. While responding to an unanticipated task, type of automation influenced pilots' ratings of all HAI consequences, with the exception of situation and automation awareness. The scenario involving the use of a clumsy system (multiple steps required for removing the speed restriction) as opposed to an easy system (one step process) was associated with more workload, task management, and automation cross-checking and was judged to increase a crew's chances of committing an automation-related error (LAS_2 vs. LAS_3). The presence of additional environmental stressors (i.e., ATC clearance was issued late and crew was disrupted while implementing the speed change) was found to exacerbate the effects of automation characteristics on task management and error potential (LAS_4 vs. LAS_5).

Effects of context factors. Context variables influenced how pilots perceived the HAI scenarios. Scenarios that involved time pressure resulting from a late clearance as well as a traffic advisory disrupting its implementation were rated higher in workload, task management and error potential, irrespective of the type of FMS used, than scenarios without these contextual constraints (LAS_4 vs. LAS_2; LAS_5 vs. LAS_3).

LAX Scenarios

MANOVA results did not show a significant effect of the between factor (VNAV operative vs. inoperative), $F(6, 51) = .897$, *ns*, and therefore we collapsed across this variable in subsequent analyses. As with the LAS family, the scenario variable had a significant effect on

the combined dependent variables, $F(24, 33) = 5.904$, $p = .0001$, $\eta^2 = .88$. Univariate tests further indicated that pilots' ratings of all HAI consequences with the exception of automation cross-checking significantly varied across scenarios (see Table 4). Pairwise comparisons explored these differences further, focusing on the impact of task, automation, and context factors on pilots' ratings of workload, task management, situation and automation awareness, and error potential. All reported differences were significant at the .01 level.

The effects of task characteristics. Situations that involved an unanticipated runway change during approach without additional complications were judged to increase a crew's workload compared to the level crewmembers would experience during a normal approach (LAX_2; LAX_3 vs. LAX_1). This effect was observed irrespective of the FMS used to update the runway assignment. Additional effects were associated with the current FMS (manual FMS) but not with the future system that allowed the crew to automatically upload the new clearance via datalink. Reliance on current procedures (i.e., reprogramming the FMS) to cope with the unanticipated runway change was thought to require more task management, decrease a crew's situational awareness, and increase the likelihood of an automation-related error compared to the normal approach scenario (LAX_2 vs. LAX_1).

The effects of automation features. Changes in automation (manual FMS vs. datalink) affected ratings of workload, task management, and error potential but not judgments pertaining to the crew's situational and automation awareness (LAX_2 vs. LAX_3). Raters thought that compared to the current FMS, a future system with datalink would reduce crewmembers' workload, require less task management, and mitigate against automation-related errors as they comply with an unanticipated runway change, even if additional stressors (time pressure and task disruption) were present in the operational context (LAX_4 vs. LAX_5).

The effects of context factors. Pilots' judgments of HAI variables were not significantly influenced by contextual factors (LAX_2 vs.

TABLE 4: LAX Family—Means, Standard Deviations, and Univariate Test Results for Human–Automation Interaction (HAI) Ratings

HAI Variable	Scenario										
	LAX_1		LAX_2		LAX_3		LAX_4		LAX_5		
	M	SD	M	SD	M	SD	M	SD	M	SD	
Workload	3.79	1.31	5.16	1.28	4.52	1.44	5.57	1.30	4.84	1.41	$F(3.26; 182.32) = 25.42^{***}$, $\eta^2 = .31$
Task management	3.62	1.52	5.17	1.19	3.98	1.63	5.26	1.42	4.29	1.50	$F(3.17; 177.69) = 23.78^{***}$, $\eta^2 = .30$
Automation cross-checking	5.33	1.34	5.60	1.18	5.50	1.11	5.41	1.38	5.43	1.26	$F(3.34; 187.03) = 0.54$, ns
Situation awareness	5.59	0.92	4.98	1.22	5.21	1.14	5.12	1.46	5.12	1.31	$F(3.44; 192.74) = 3.68^{**}$, $\eta^2 = .06$
Automation awareness	5.31	1.33	4.95	1.33	4.95	1.07	5.28	1.42	4.88	1.40	$F(3.18; 177.94) = 3.12^{*}$, $\eta^2 = .05$
Likelihood of error	3.70	1.55	4.77	1.42	4.00	1.61	4.97	1.61	4.10	1.73	$F(3.50; 195.86) = 13.66^{***}$, $\eta^2 = .20$

Note. N = 58. Greenhouse–Geisser adjustment used for violation of homogeneity of covariance.
*Significant at the .05 level. **Significant at the .01 level. ***Significant at the .0001 level.

LAX_4; LAX_3 vs. LAX_5). That is, the presence of time pressure (clearance issued at lower altitude) and a task disruption (traffic alert) did not change pilots' judgments of workload, task management, situation and automation awareness, and error potential regardless of whether they were managing a new runway assignment by using the current manual FMS or a new datalink system.

DISCUSSION AND CONCLUSIONS

Results of the study lend support to the established notion that HAI comprises a complex interplay between features of the automation and elements of the task and context. Of particular interest is the correspondence between our results and those of other research, as our findings reflect the perceptions of pilots who are likely users of advanced automation.

Automation features had the most wide-ranging effect, influencing judgments of a crew's workload, task management, and potential for automation-related errors. Consistent with other research, clumsy automation increased work (task management and automation cross-checking) and workload as well as the potential for making errors. Surprisingly, we did not find any effect of automation features on automation awareness. This result may reflect characteristics of the tasks described in the scenarios as they both revolved around data entry into the FMS. Our raters may have assumed that in these situations a crew would consistently have a good awareness of the automated system regardless of what else was going on. This finding, however, may also reflect pilots' somewhat unrealistic evaluations of the consistency of their automation monitoring processes and knowledge of system status (e.g., Parasuraman & Manzey, 2010).

The impact of automation on situation awareness seems to be moderated by task features, as reflected by pilots' judgments. In both scenario families the more laborious FMS was judged to impair a crew's situation awareness as pilots performed an unanticipated task. No such effect was noted when the crew used the easier systems, presumably because data entry into these systems would require less head-down time than interactions with the former systems. This result is consistent with the Olson and Sarter (2000, 2001)

findings that pilots liked one-button "gated" management-by-consent datalink systems.

The demands associated with unanticipated tasks (runway change or deletion of speed restriction) had broader effects on pilots' judgments of HAI variables than stressors in the operational context (time pressure or ATC disruptions). This suggests that cognitive and behavioral consequences of HAI situations may be primarily the result of task characteristics and automation features, whereas variables such as time pressure and distraction may exacerbate their effects. This interpretation is consistent with aviation accident analyses that identified context variables as contributing factors in the chain of events (see Dismukes et al., 2007; Orasanu et al., 2001).

It should be noted that pilots judging vignettes may fail to account for the dynamic nature of the flight situations in which HAI takes place. Participants in our study rated one-button data entry as a better option in terms of situation awareness compared to multistep alternatives, as did pilots in research by Olson and Sarter (2000, 2001). However, methods that allow for quick system input and implementation may foster shallow processing of the information being entered, and thus pilots may fail to appreciate the implications of actions being approved. Olson and Sarter (2001) found evidence of this, as the pilots who used the datalink gating option were less apt to notice conflicting clearances or automation implementation failures.

Our study's reliance on self-reported judgments rather than on actual performance constitutes a limitation as well as a contribution to knowledge of HAI. Paper-and-pencil scenario judgments have long provided complementary data in multimethod research programs (e.g., Mosier, Keyes, & Bernhard, 2003). Human in the loop studies—in aviation as well as in other dynamic and increasingly high-technology domains—are needed to confirm the complex links between automation, task, and context variables and resultant HAI interactions.

The challenge of future research will be to examine these issues in other domains such as medicine and in operationally more realistic conditions than the present study to tease apart the individual and joint effects of automation characteristics, task

features, and context variables on operators' actual interactions with automation and to see whether or not the outcome predictions from our sample hold true. In either case, the results of this study contribute information important to automation design and training, as pilot perceptions of HAI consequences and outcomes are likely to influence their HAI decisions in practice.

Our results indicate that even though changes to automated systems may be small in technical terms, their cognitive and behavioral impact on operators—in this case flight crews—may be significant and should be considered in the design process. It is likely that the same caveat applies in other dynamic domains as well. Moreover, the performance effects of automation changes need to be addressed with reference to task characteristics and situational demands. In aviation, training at many airlines promotes the position that it is the pilot's decision when and how much automation to use. However, future flight operations will require pilots to interact with more and increasingly complex automation. It is conceivable that there also will be increased pressure on pilots to rely on automation to enable high levels of efficiency in operations. Crew training therefore must provide pilots with the opportunity to explore automation use in diverse situations, highlighting different task and operational constraints.

APPENDIX A

Scenarios Describing an Approach to Las Vegas (LAS)

Variable	Scenario Description (based on ASRS Report No. 892363)
LAS_1 (Parent)—ate*Approach as programmed*	An aircraft is on the TYSSN THREE arrival via the Drake transition into LAS. The arrival is entered and briefed, and runway 25L is programmed into the flight management computer. VNAV and autopilot are engaged. As the aircraft descends through FL 140, ATC clears the flight for descent via the arrival and confirms runway 25L. The pilot monitoring repeats the clearance, and the pilot flying selects 8,000 ft. in the altitude window.
LAS_2 Unexpected task—*new speed issued* **Clumsy automation—***FMS update requires multiple steps*	An aircraft is cleared to cross KADDY at 12,000 ft. and then to descend via the TYSSN THREE arrival into LAS. The arrival is entered and briefed, and runway 25L is programmed into the flight management computer. VNAV and autopilot are engaged, and 8,000 ft. is set in the altitude window. As the aircraft descends through FL 160, ATC instructs the crew to **delete the KADDY speed restriction and to expedite descent**. The aircraft is equipped with a flight management system in which **changing speed entries requires the pilot to delete the whole line and then to reenter both the speed and altitude constraints**. The pilot monitoring goes "head down," deletes the speed/altitude entry for KADDY, and reenters the altitude constraint.

(continued)

APPENDIX A: (continued)

Variable	Scenario Description (based on ASRS Report No. 892363)
LAS_3 Unexpected task—*new speed issued* **Easy automation**—*FMS update requires single step*	An aircraft is cleared to cross KADDY at 12,000 ft. and then to descend via the TYSSN THREE arrival into LAS. The arrival is entered and briefed, and runway 25L is programmed into the flight management computer. VNAV and autopilot are engaged, and 8,000 ft. is set in the altitude window. As the aircraft descends through FL 160, ATC instructs the crew to **delete the KADDY speed restriction and to expedite descent**. The aircraft is equipped with a flight management system in which **the pilot can change the speed entry separate from the altitude constraint.** The pilot monitoring goes "head down," and deletes the KADDY speed restriction only.
LAS_4 Unexpected task, Clumsy automation **Time pressure**—*ATC late in lifting speed restriction* **Task disruption**—*traffic advisory*	An aircraft is cleared to cross KADDY at 12,000 ft. and then to descend via the TYSSN THREE arrival into LAS. The arrival has been entered and briefed, and runway 25L programmed into the flight management computer. VNAV and autopilot are engaged, and 8,000 ft. is set in the altitude window. Approximately 5 NM from KADDY, ATC instructs the crew to **delete the KADDY speed restriction and to expedite descent**. The aircraft is equipped with a flight management system in which **changing speed entries requires the pilot to delete the whole line and then to reenter both the speed and altitude constraints.** The pilot monitoring goes "head down," deletes the speed/altitude entry for KADDY, and reenters the altitude constraint. While the pilot monitoring reprograms the flight management system, **TCAS generates a traffic advisory for an aircraft climbing to level at 11,000 ft. Shortly thereafter, ATC calls to advise the crew of traffic at 10 o'clock, crossing left to right, leveling at 11,000 ft**.
LAS_5 Unexpected task, Easy automation **Time pressure**—*ATC late in lifting speed restriction* **Task disruption**—*traffic advisory*	An aircraft is cleared to cross KADDY at 12,000 ft. and then to descend via the TYSSN THREE arrival into LAS. The arrival has been entered and briefed, and runway 25L programmed into the flight management computer. VNAV and autopilot are engaged, and 8,000 ft. is set in the altitude window. Approximately 5 NM from KADDY, ATC instructs crew to **delete the KADDY speed restriction and to expedite descent**. The aircraft is equipped with a flight management system in which **the pilot can change the speed entry separate from the altitude constraint.** The pilot monitoring goes "head down," and deletes the KADDY speed restriction only. While the pilot monitoring reprograms the flight management system, **TCAS generates a traffic advisory for an aircraft climbing to level at 11,000 ft. Shortly thereafter, ATC calls to advise the crew of traffic at 10 o'clock, crossing left to right, leveling at 11,000 ft.**

APPENDIX B

Scenarios Describing an Approach to Los Angeles (LAX)—VNAV (vertical navigation) Is Inoperative

Variable	Scenario Description (based on ASRS Report No. 883679)
LAX_1 (Parent)— *Approach as programmed*	An aircraft is on the SEAVU TWO arrival into LAX. The arrival has been entered and briefed, and runway 25L programmed into the FMS/FMGS. The aircraft is in FLCH/Open Descent because VNAV had been performing inconsistently earlier during descent and the pilot flying decided to switch it off. The aircraft is at 17,000 ft crossing KONZL when ATC clears the flight for descent via the arrival and confirms runway 25L. The pilot monitoring reads back the clearance, and the pilot flying enters the next altitude (ENGLI @ 16,000 ft) in the MCP/FCU window.
LAX_2 Unexpected task— *runway change* **Current system—**reprogram FMS/FMGS	An aircraft is on the SEAVU TWO arrival into LAX. The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. The aircraft is in FLCH/Open Descent because VNAV had been performing inconsistently earlier during descent and the pilot flying decided to switch it off. **The aircraft is at 17,000 ft crossing KONZL, when ATC changes the runway assignment to 25L** and at the same time clears the flight for descent via the arrival. The pilot monitoring reads back the clearance, and goes "head down" to reprogram the FMS/FMGS for the new runway, and the pilot flying enters the next altitude (ENGLI @ 16,000 ft) in the MCP/FCU window.
LAX_3 Unexpected task— *runway change* **Future FMS/FMGS—**single button press to upload information*	An aircraft is on the SEAVU TWO arrival into LAX. **This aircraft has new datalink capabilities allowing pilots to automatically accept a clearance and enter it into the FMS/FMGS with a single button press.** The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. The aircraft is in FLCH/Open Descent because VNAV had been performing inconsistently earlier during descent and the pilot flying decided to switch it off. **The aircraft is at 17,000 ft crossing KONZL when ATC sends a datalink message changing the runway assignment to 25L** and clearing the flight for descent via the arrival. The pilot monitoring accepts and enters the new runway clearance with a single button press on the FMS/FMGS, and the pilot flying enters the next altitude (ENGLI @ 16,000 ft) in the MCP/FCU window.
LAX_4 Unexpected task, Current FMS/FMGS **Time pressure—**runway change issued at lower altitude* **Task disruption—**parallel traffic*	An aircraft is on the SEAVU TWO arrival into LAX. The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. The aircraft is in FLCH/Open Descent because VNAV had been performing inconsistently earlier during descent and the pilot flying decided to switch it off. The aircraft is at 17,000 ft crossing KONZL when ATC clears the flight for descent via the arrival. **As the aircraft is crossing PECOX @ 14,000 ft, ATC changes the runway to 25L.** The pilot monitoring reads back the clearance and goes "head down" to reprogram the FMS/FMGS for the new runway while the pilot flying enters the next altitude (SEAVU @ 12,000 ft) in the MCP/FCU window. **At the same time ATC calls back and advises the crew of traffic on descent for parallel approach to 24R.**

(continued)

APPENDIX B: (continued)

Variable	Scenario Description (based on ASRS Report No. 883679)
LAX_5 Unexpected task, Future FMS/FMGS **Time pressure**—*runway change issued at lower altitude* **Task disruption**—*parallel traffic*	An aircraft is on the SEAVU TWO arrival into LAX. **This aircraft has new datalink capabilities allowing pilots to automatically accept a clearance and enter it into the FMS/FMGS with a single button press.** The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. The aircraft is in FLCH/Open Descent because VNAV had been performing inconsistently earlier during descent and the pilot flying decided to switch it off. The aircraft is at 17,000 ft crossing KONZL when ATC clears the flight for descent via the arrival. **As the aircraft is crossing PECOX @ 14,000 ft**, ATC sends a datalink message changing the runway assignment to 25L.** The pilot monitoring accepts and enters the new runway clearance with a single button press on the FMS/FMGS, while the pilot flying enters the next altitude (SEAVU @ 12,000 ft) in the MCP/FCU window. **At the same time, ATC calls back and advises the crew of traffic on descent for parallel approach to 24R**.

Scenarios Describing an Approach to Los Angeles (LAX)—VNAV (vertical navigation) Is Working

Variable	Scenario Description (based on ASRS Report No. 883679)
LAX_1 (Parent)—*Approach as programmed*	An aircraft is on the SEAVU TWO arrival into LAX. The arrival has been entered and briefed, and runway 25L programmed into the FMS/FMGS. VNAV and autopilot are engaged. The aircraft is at 17,000 ft crossing KONZL when ATC clears the flight for descent via the arrival and confirms runway 25L. The pilot monitoring reads back the clearance, and the pilot flying enters 12,000 ft in the MCP/FCU window.
LAX_2 Unexpected task— *runway change* **Current system—**reprogram FMS/FMGS	An aircraft is on the SEAVU TWO arrival into LAX. The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. VNAV and autopilot are engaged, and 12,000 ft. is set in the MCP/FCU window. **The aircraft is at 17,000 ft crossing KONZL when ATC changes the runway assignment to 25L** and at the same time clears the flight for descent via the arrival. The pilot monitoring reads back the clearance and goes "head down" to reprogram the FMS/FMGS for the new runway.
LAX_3 Unexpected task— *runway change* **Future FMS/FMGS—***single button press to upload information*	An aircraft is on the SEAVU TWO arrival into LAX. **This aircraft has new datalink capabilities allowing pilots to automatically accept a clearance and enter it into the FMS/FMGS with a single button press.** The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. VNAV and autopilot are engaged, and 12,000 ft. is set in the MCP/FCU window. **The aircraft is at 17,000 ft crossing KONZL when ATC sends a datalink message changing the runway assignment to 25L** and clearing the flight for descent via the arrival. The pilot monitoring accepts and enters the new runway clearance with a single button press on the FMS/FMGS.

(continued)

APPENDIX B: (continued)

Variable	Scenario Description (based on ASRS Report No. 883679)
LAX_4 Unexpected task, Current FMS/FMGS **Time pressure**—*runway change issued at lower altitude* **Task disruption**—*parallel traffic*	An aircraft is on the SEAVU TWO arrival into LAX. The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. VNAV and autopilot are engaged, and 12,000 ft. is set in the MCP/FCU window. The aircraft is at 17,000 ft crossing KONZL when ATC clears the flight for descent via the arrival. **As the aircraft is crossing PECOX @ 14,000 ft, ATC changes the runway to 25L**. The pilot monitoring reads back the clearance and goes "head down" to reprogram the FMS/FMGS for the new runway. **At the same time ATC calls back to advise the crew of traffic on descent for parallel approach to 24R**.
LAX_5 Unexpected task, Future FMS/FMGS **Time pressure**—*runway change issued at lower altitude* **Task disruption**—*parallel traffic*	An aircraft is on the SEAVU TWO arrival into LAX. **This aircraft has new datalink capabilities allowing pilots to automatically accept a clearance and enter it into the FMS/FMGS with a single button press.** The arrival has been entered and briefed, and runway 24R programmed into the FMS/FMGS. VNAV and autopilot are engaged, and 12,000 ft. is set in the MCP/FCU window. The aircraft is at 17,000 ft crossing KONZL, when ATC clears the flight for descent via the arrival. **As the aircraft is crossing PECOX @ 14,000 ft, ATC changes the runway to 25L**. The pilot monitoring accepts and enters the new runway clearance with a single button press on the FMS/FMGS. **At the same time ATC calls back to advise the crew of traffic on descent for parallel approach to 24R.**

REFERENCES

Aircraft Accident Investigation Commission, Ministry of Transport—Japan. (1996). *China Airlines Airbus Industrie A300B4-622R, B1816, Nagoya Airport, April 26, 1994* (Rep. No. 96-5). Tokyo: Author.

Billings, C. E. (1996). Human-centered aviation automation: Principles and guidelines. (*NASA Technical Memorandum 110381*). Moffett Field, CA: National Aeronautics and Space Administration.

Björklund, C. M., Alfredson, J., & Dekker, S. (2006). Mode monitoring and call-outs: An eye-tracking study of two-crew automated flight deck operations. *International Journal of Aviation Psychology*, *16*, 263–275.

Boorman, D. J., & Mumaw, R. J. (2004). A new autoflight/FMS interface: Guiding design principles. In *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence.

Casner, S. M. (1994). Understanding the determinants of problem-solving behavior in a complex environment. *Human Factors*, *36*, 580–596.

Christoffersen, K., & Woods, D. D. (2002). How to make automated systems team players. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research, vol. 2: Automation* (pp. 1–12). Oxford, UK: Oxford University Press.

Cook, R. I., & Woods, D. D. (1996). Adapting to new technology in the operating room. *Human Factors*, *381*, 593–613.

Cook, R. I., Woods, D. D., Howie, M. B., Horrow, J. C., & Gaba, D. M. (1992). Unintentional delivery of vasoactive drugs with an electromechanical infusion device. *Journal of Cardiothoracic & Vascular Anesthesia*, *6*, 238–244.

Coppenbarger, R. A., Mead, R. W., & Sweet, D. N. (2007, September). *Field evaluation of the tailored arrivals concept for datalink-enabled continuous descent approach*. Presentation at the 7th AIAA Integrations and Operations Conference, Belfast, Northern Ireland.

Degani, A., & Heymann, M. (2002). Formal verification of human–automation interaction. *Human Factors*, *44*, 28–43.

Dekker, S., & Orasanu, J. (1999). Automation and SA—pushing the research frontier. In S. Dekker & E. Hollnagel (Eds.), *Coping with computers in the cockpit* (pp. 69–85). Aldershot, UK: Ashgate.

Dismukes, R. K., Berman, B. A., & Loukopoulos, L. D. (2007). *The limits of expertise: Rethinking pilot error and the causes of airlines accidents*. Burlington, VT: Ashgate.

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

Dorneich, M. C., Ververs, P. M., Mathan, S., Whitlow, S., & Hayes, C. C. (2012). Considering etiquette in the design of an adaptive system. *Journal of Cognitive Engineering and Decision Making*, *6*, 243–265.

Durso, F. T., & Alexander, A. L. (2010). Managing workload, performance, and situation awareness in aviation systems. In E. Salas & D. Maurino (Eds.), *Human factors in aviation* (2nd ed., pp. 217–247). Burlington, MA: Elsevier.

Federal Aviation Administration. (2011, December). *14CFR Parts 117, 119, and 121. Flightcrew member duty and rest requirements*. Washington, DC: Author.

Ferris, T., Sarter, N., & Wickens, C. D. (2010). Cockpit automation: Still struggling to catch up. In E. Salas, T. Allard, & D. Maurino (Eds.), *Human factors in aviation* (2nd ed., pp. 479–504). Burlington, MA: Elsevier.

Gawande, A., & Bates, D. (2000, February). The use of information technology in improving medical performance: Part I. Information systems for medical transactions. *Medscape General Medicine*, *2*, 1–6.

HART Group. (2011a). *Automation in the cockpit: Toward a human–automation relationship taxonomy* (Interim report to the FAA). Atlanta: Georgia Institute of Technology.

HART Group. (2011b). *Deconstructing human automation interaction using aviation scenarios* (Interim report to the FAA). Atlanta: Georgia Institute of Technology.

Hollnagel, E., Kaarstad, M., & Lee, H. C. (1999). Error mode prediction. *Ergonomics*, *42*(11), 1457–1471.

Javaux, D. (1998). Explaining Sarter and Woods' classical results. The cognitive complexity of pilot-autopilot interaction on the Boeing 737-EIS. In *Proceedings of Human Error, Safety, and System Development (HESSD '98)* (pp. 62–77). Seattle, WA: International Federation for Information Processing.

Javaux, D. (2002). A method for predicting errors when interacting with finite state systems. How implicit learning shapes the user's knowledge of a system. *Reliability Engineering and System Safety*, *75*, 147–165.

Johnson, E. N., & Pritchett, A. R. (1995). *Experimental study of vertical flight path mode awareness* (ASL-95-3). Cambridge, MA: International Center for Air Transportation. Retrieved from http://web.mit.edu/aeroastro/labs/icat/research.shtml

Joint Planning and Development Office. (2007). *Concept of operations for the next generation air transportation system*. Washington, DC: Government Printing Office.

Khoo, Y.-L., & Mosier, K. L. (2009). The impact of time pressure and experience on information search and decision making processes. *Journal of Cognitive Engineering and Decision Making*, *2*, 275–294.

Kim, J. W., & Jung, W. (2003). A taxonomy of performance influencing factors for human reliability analysis in emergency tasks. *Journal of Loss Prevention in the Process Industries*, *16*, 479–495.

Kirlik, A. (1993). Modeling strategic behavior in human–automation interaction: Why an "aid" can (and should) go unused. *Human Factors*, *35*, 221–242.

Kirlik, A., Plamondon, B. D., Lytton, L., Jagacinski, R. J., & Miller, R. A. (1993). Supervisory control in a dynamic and uncertain environment: Laboratory task and crew performance. *IEEE Transactions on Systems, Man and Cybernetics*, *23*, 1130–1138.

Latorella, K. (1999). *Investigating interruptions: Implications for flightdeck performance* (NASA TM-1000-209707). Retrieved from http://techreports.larc.nasa.gov/ltrs/PDF/1999/tm/NASA-99-tm209707.pdf

Lindeis, A.-E. (2010). Air traffic management. In E. Salas & D. Maurino (Eds.), *Human factors in aviation* (2nd ed., pp. 629–657). Burlington, MA: Elsevier.

Loukopoulos, L., Dismukes, K., & Barshi, I. (2001). Cockpit interruptions and distractions: A line observation study. In R. Jensen (Ed.), *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus: Ohio State University.

Lyall, B., Niemczyk, M., Lyall, R., & Funk, K. (1997). Flightdeck automation: Evidence for existing problems. In *Proceedings of the 9th International Symposium on Aviation Psychology*. Columbus: Ohio State University. Retrieved from http://www.researchintegrations.com/publications/lyall-niemczyk-lyall-funk-1997.pdf

Manzey, D., Reichenbach, J., & Onnasch, L. (2009). Human performance consequences of automated decision aids in states of fatigue. In *Proceedings of the 53rd Meeting of the Human Factors and Ergonomics Society* (pp. 329–333). Santa Monica, CA: Human Factors and Ergonomics Society.

Millitello, L. G., Patterson, E. S., Saleem, J. J., Anders, S., & Asch, S. (2008). Supporting macrocognition in health care: Improving clinical reminders. In J. M. Schraagen, L. G. Millitello, T. Ormerod, & R. Lipshitz (Eds.), *Naturalistic decision making and macrocognition* (pp. 203–220). Aldershot, UK: Ashgate.

Morrow, D., North, R., & Wickens, C. D. (2006). Reducing and mitigating human error in medicine. In R. S. Nickerson (Ed.), *Reviews of human factors and ergonomics* (Vol. 1, pp. 254–296). Santa Monica, CA: Human Factors and Ergonomics Society.

Mosier, K. L., Keyes, J., & Bernhard, R. (2003). Dealing with conflicting information—Will crews rely on automation? In G. Edkins & P. Pfister (Eds.), *Innovation and consolidation in aviation: Selected contributions to the Australian Aviation Psychology Symposium 2000* (pp. 243–253). Aldershot, UK: Ashgate.

Mosier, K., Sethi, N., McCauley, S., Khoo, L., & Orasanu, J. (2007). What you don't know CAN hurt you: Factors impacting diagnosis in the automated cockpit. *Human Factors*, *49*, 300–310.

Mosier, K. L., & Skitka, L. J. (1996). Humans and automation: Made for each other? In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 201–220). Mahwah, NJ: Lawrence Erlbaum.

Mueller, E., & Lozito, S. (2008, September). *Flight deck procedural guidelines for datalink trajectory negotiation*. Paper presented at the American Institute of Aeronautics and Astronautics Aviation Technology, Integration, and Operations Conference, Anchorage, AK.

Olson, W. A., & Sarter, N. B. (2000). Automation management strategies: Pilot preferences and operational experiences. *International Journal of Aviation Psychology*, *10*, 327–341.

Olson, W. A., & Sarter, N. B. (2001). Management by consent in human-machine systems: When and why it breaks down. *Human Factors*, *43*, 255–266.

Orasanu, J., Martin, L., & Davison, J. (2001). Cognitive and contextual factors in aviation accidents: Decision errors. In E. Salas & G. Klein (Eds.), *Linking expertise and naturalistic decision making* (pp. 209–225). Mahwah, NJ: Lawrence Erlbaum.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*, 381–410.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced "complacency." *International Journal of Aviation Psychology*, *3*, 1–23.

Patterson, E., Nguyen, A. D., Halloran, J. M., & Asch, S. M. (2004). Human factors barriers to the effective use of ten HIV clinical reminders. *Journal of the American Medical Informatics Association*, *11*, 50–59.

Pritchett, A. R. (2009). Aviation automation: General perspectives and specific guidance for the design of modes and alerts. In F. T. Durso (Ed.), *Reviews of Human Factors and Ergonomics* (Vol. 5, pp. 82–113). Santa Monica, CA: Human Factors and Ergonomics Society.

Riley, V. (2001). A new language for pilot interfaces. *Ergonomics in Design*, *9*, 21–26.

Sarter, N. (2008). Mode errors on automated flight decks: Illustrating the problem-driven, cumulative, and interdisciplinary nature of human factors research. *Human Factors*, *50*, 506–510.

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors*, *49*, 347–357.

Sarter, N., & Woods, D. D. (1994). Decomposing automation: Autonomy, authority, observability and perceived animacy. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 22–27). Mahwah, NJ: Lawrence Erlbaum.

Sarter, N. B., & Woods, D. D. (2000). Team play with a powerful and independent agent: A full-mission simulation study. *Human Factors*, *42*, 390–402.

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

Schroeder, C. C., & Harms, C. D. (2007). MPL represents a state-of-the-art ab initio airline pilot training program. *ICAO Journal*, *62*(3), 15–16, 31–32.

Sheridan, T. B., & Parasuraman, R. (2006). Human–automation interaction. In R. S. Nickerson (Ed.), *Reviews of human factors and ergonomics* (Vol. 1, pp. 89–129). Santa Monica, CA: Human Factors and Ergonomics Society.

Sherry, L., Feary, M., Polson, P., Mumaw, R., & Palmer, E. (2001). A cognitive engineering analysis of the vertical navigation (VNAV) function. *Journal of Human Factors and Aerospace Safety*, *1*(3), 223–245.

Sherry, L., Feary, M., Polson, P., & Palmer, E. (2001). What's it doing now? Taking the covers off autopilot behavior. In *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus: Ohio State University.

Sherry, L., Polson, P., Feary, M., & Palmer, E. (2002). When does the MCDU interface work well? Lessons learned for the design of new flightdeck user-interfaces. In S. Chatty, J. Hansman, & G. Boy (Eds.), *Proceedings of HCI-Aero-2002, International Conference on Human Computer Interaction in Aeronautics* (pp. 180–185). Menlo Park, CA: AAAI Press.

Sullivan, K. B., Feigh, K. M., Mappus, R., IV, Durso, F. T., Fischer, U., Pop, V., . . . Morrow, D. G. (2013). Using neural networks to assess flight deck human–automation interaction. *Reliability Engineering & System Safety*, *114*, 26–35.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York, NY: HarperCollins.

Wein, L. E. (1996). Maladjusted contrivances and clumsy automation: A jurisprudential investigation. *Harvard Journal of Law & Technology*, *9*, 375–427.

Wichman, K., Carlsson, G., & Lindberg, L. (2002, August). *Flight trials: "Runway-to-runway" required time of arrival evaluations for time-based ATM environment—Final results*. Paper presented at the AIAA Guidance, Navigation and Control Conference, Washington, DC.

Wickens, C. D., Stokes, A., Barnett, B., & Hyman, F. (1993). The effects of stress on pilot judgment in a MIDIS simulator. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 271–292). New York, NY: Plenum.

Wiener, E. L. (1989). *Human factors of advanced technology ("glass cockpit") transport aircraft*. (NASA Contractor Report No. 177528). Moffett Field, CA: NASA Ames Research Center.

Kathleen L. Mosier is a professor at San Francisco State University and director of the Applied Psychology Research Laboratory. She received her PhD in industrial/organizational psychology from the University of California, Berkeley, and spent 7 postdoctoral years at NASA Ames Research Center as a senior research scientist. Current research interests include the impact of operational variables such as fatigue, technology, and time pressure on decision-making processes, as well as the impact of time delay on remote team communication and decision making.

Ute Fischer is a research scientist in the School of Media and Communication at the Georgia Institute of Technology. After receiving her PhD in cognitive psychology from Princeton University, she was a postdoctoral fellow and then a senior research scientist at NASA Ames Research Center. Her research addresses aspects of individual and team decision making in complex, high-technology environments, in particular factors affecting practitioners' risk assessment, communication, and decision strategies.

Dan Morrow is a professor in the Department of Educational Psychology, with appointments in the Beckman Institute of Advanced Science & Technology and Department of Industrial and Enterprise Engineering, University of Illinois at Urbana-Champaign. He received a PhD in cognitive psychology from the University of California, Berkeley. His research

interests include the impact of aging on cognition and communication in health care, aviation, and other domains in which people interact with technology.

Karen M. Feigh is an assistant professor at the Georgia Institute of Technology's School of Aerospace Engineering and a faculty member in the Cognitive Engineering Center. She earned a PhD in industrial and systems engineering from the Georgia Institute of Technology in 2008.

Francis T. (Frank) Durso received his PhD from SUNY at Stony Brook and his B.S. from Carnegie-Mellon University. He is currently a professor of engineering psychology at the Georgia Institute of Technology.

Katlyn Sullivan received her BS in mechanical engineering from Washington University in St. Louis in 2010 and her MS in aerospace engineering from Georgia Tech in 2012. During her master's program, she focused on cognitive engineering and human factors and studied the methods by which pilots interact with cockpit automation. She currently works as a human factors engineer for Emerson Electric at the Emerson Innovation Center–Human Centered Design in Pittsburgh, Pennsylvania.

Vlad Pop has a BA in psychology from North Carolina State University and an MS in engineering psychology from the Georgia Institute of Technology. His research interests included vigilance decrements in NextGen air traffic control and trust in automation. He is currently at Georgia Tech pursuing a PhD in engineering psychology with a focus on human error.