UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Modeling Human Error in a Real-World Teamwork Environment

Permalink

Journal

ISSN

Authors

Deutsch, Stephen
Pew, Richard

Publication Date

2002

Peer reviewed

Modeling Human Error in a Real-World Teamwork Environment

Stephen Deutsch (sdeutsch@bbn.com)
BBN Technologies, 10 Moulton Street
Cambridge, MA 02138 USA

Richard Pew (pew@bbn.com)
BBN Technologies, 10 Moulton Street
Cambridge, MA 02138 USA

Abstract

An initial model of human error in a real-world teamwork environment has been developed. The captain and first officer of a commercial aircraft and the air traffic controllers with whom they interact are modeled as the crew executes an approach and landing followed by taxi operations that take them to their assigned gate. Scenario details and human model development were based on the results of a series of full-task simulation experiments using commercial pilots as subjects. The focus of the experiment was on errors committed by the aircrews during taxi operations. The models developed exhibit the robust behaviors typically exhibited by aircrews and identify psychologically grounded windows for error within that robust behavior.

Human Error Modeling Applied to Taxi Operations

NASA Ames Research Center conducted two full-mission studies of airport taxi operations under low visibility and night conditions. The subject of the studies was the Taxiway Navigation and Situation Awareness (T-NASA) system, aircraft flight deck technology designed to improve commercial airport taxi operations in poor weather while maintaining a high degree of safety (Hooey, Foyle, & Andre, 2000). The T-NASA system includes a head-up display, a head-down electronic moving map display, and directional audio alerts. The studies included a series of baseline trials run without the T-NASA system and a series of trials using various configurations of the T-NASA system. The T-NASA system effectively eliminated very nearly all error, hence the focus of the human error modeling effort was on the baseline trials.

The NASA Ames Advanced Concept Flight Simulator (ACFS) used in the studies provided a generic glass cockpit simulator with a 180-degree field of view and a high fidelity rendering of Chicago O'Hare Airport replicating the airport layout including runways, taxiways, signage, painted markings, lights, concourses, and structures (Hooey & Foyle, 2001). In the first study, 16 two-pilot commercial crews each completed six land and taxi-to-gate trials based on current operations using Jeppesen charts for navigation. Half of the trials were under low visibility conditions with runway visual range (RVR) of 700 feet, and half under night visual meteorological conditions (VMC). In the second study, 18 commercial two-pilot crews each completed three land and taxi-to-gate trials based on current operating conditions under 1000 foot RVR conditions. In evaluating these studies, Hooey and Foyle (2001) defined navigation errors as taxiing on a portion of the airport surface on which the aircraft had not been cleared and deviating from their cleared centerline by at least 50 feet. Their analysis revealed 26 navigation errors in 150 current-operation trials—errors were committed on 17.3% of the trials.

Modeling Robust Nominal Performance as a Prelude to Modeling Error

As we set out to identify the sources of error (c.f., Deutsch & Pew, 2001) and then to model error in taxi procedures, we started by refining earlier Distributed Operator Model Architecture (D-OMAR) models (Deutsch, 1998; Deutsch & Adams, 1995) that captured the robustness in aircrew procedures. The models represent the multiple task behaviors of each player as the product of a mix of goals and procedures that operate concurrently to proactively address the player's agenda. Expectations integrate anticipated events while anticipated or unanticipated interruptions must be accommodated. Ongoing tasks determine their own execution times and run to completion unless another procedure defined as a competing procedure with greater priority intervenes. A mix of automatic and thoughtful behaviors are modeled without resorting to a central executive responsible for explicitly scheduling all future actions. A thoughtful cognitive act of decision-making is defined as just that, another procedure that determines the action to follow.

The NASA Information to Modelers package included a Nominal Task Sequence (NASA, 2001a) for the T-NASA baseline conditions. This was used as the basis for the development of the approach-and-landing and taxi procedures that the models of the captain, first officers, and air traffic controllers employed. Approach-and-landing is one of the busiest phases of flight, making high demands on the aircrew. In spite of the high demands of getting the aircraft safely on the ground, it is also the time at which the first steps in the subsequent taxi operations are initiated. The crew is in the process of approaching a given runway and already know the concourse and gate toward which they will be heading. Moreover, as specified in the Nominal Task Sequence, at about eleven miles out they discuss with the air traffic controller and among themselves which runway exit

they will take. As we will argue below, the crewmembers each now have in mind one and perhaps several taxi routes they might take to the gate. Once the runway exit information is in hand, the focus of attention returns to landing the aircraft and rollout.

The information provided in the Nominal Task Sequence was also used as the basis for the modeling of the subsequent landing and rollout sequence. As the rollout sequence is completed and the aircraft approaches the designated runway exit, the taxi sequence is initiated. The first officer provides information to the captain on their position relative to the preferred exit based on notes taken when the preferred exit was agreed on. He/she then informs the controller that the aircraft is clearing the runway, both crewmembers then switch their radio frequency, and the first officer contacts the ground controller. At this point, the ground controller provides the crew with the taxi route to the gate and the first officer writes down the taxi route.

It was at this point that we encountered the first instance of a requirement for a coping strategy. Many of the high-speed exits at O'Hare have a very short run to the first intersection and taxiway routings can be unusually lengthy. We encountered this first when modeling a landing on runway 9R using high-speed exit M6 with an immediate left turn onto taxiway M. The first officer was head-down writing out the taxi directives and was late in providing information to the captain on the upcoming immediate turn. At this point, the captain was also listening to the taxi routing and could go with what he/she heard or slow the aircraft and obtain confirmation on the upcoming turn from the first officer. The coping strategy that we modeled had the captain go ahead with the turn as heard and notify the first-officer of the turn as it was executed.

The process for each subsequent turn in the taxi sequence followed the same pattern. As a turn was completed, the first officer would consult his/her routing notes and the airport diagram, and then prompt the captain on the taxiway and direction for the upcoming turn. As expected, the modeled nominal process proved very robust. By simply changing the routing that the ground controller provided, the captain and first officer were able to execute any desired taxi routing. With these robust aircrew processes in place, the challenge was to model taxi sequences that produced errors consistent with those in the baseline T-NASA experiments.

As the captain and first officer meet their responsibilities during taxi operations, the inherent nature of the tasks that they perform provide them each with a different sense of their immediate location and their location with respect to their assigned taxi routing. They each achieve and maintain different levels of local and global situation awareness (Wickens & Prevett, 1995). If they are working well as a team, they will strive to fill each other's gaps in awareness. In building the aircrew models, we felt that it was essential to reflect this level of teamwork.

The captain was modeled as predominantly head-up during taxi operations. He/she announced each turn as it was executed to keep the first officer informed of their immediate location during such periods as the first officer might have been head-down. Meanwhile, the first officer, working with the airport diagram and written notes on the runway exit and taxiway routing provided the captain with a more global view of their taxiway routing than would have otherwise been available. The teamwork skills of the modeled aircrews were effective in repairing gaps in one another's situation awareness. One effect of providing this level of detail in good crew performance was of course to make the taxiway procedures just that much more robust and error that much less likely.

Making the Wrong Turn at an Intersection

The particular process that produced the errors of interest was the preparation for and execution of the next turn in the taxi sequence as governed by the captain. As modeled, the captain, if left to his/her own resources, must rely on his/her memory of the taxi sequence as conveyed by the ground controller as the aircraft cleared the landing runway. However, the captain gets significant support in this task from the first officer. The first officer takes notes on the taxi sequence as it is received from the ground controller and will, under nominal conditions, prompt the captain with the name of the next taxiway and the direction of the turn required.

Under nominal conditions, the first officer prompts the captain on the upcoming turn and one can reasonable expect that the captain will correctly act in accordance with the prompt. Acting counter to the prompt is an error possibility that we did not pursue. Hence, to find a source for making a turn error at an intersection, we had to construct reasonable scenarios in which the first officer was otherwise occupied and unable to provide the prompt in a timely manner and of course identify an underlying reason for a mistake on the part of the captain. The events that prevented the first officer from providing the prompts are discussed below in the sections providing details on the error scenarios. Here, we examine possible sources for the errors committed by the captain in executing the incorrect turns.

Intention-to-Act

A classical view of the taxiway process might be that, in approaching a turn, the captain has a planning problem whose resolution is then followed by plan execution. Do we in fact need to make a turn at the upcoming intersection and if so, which way? There might be a schema in place for executing the next turn with slots to be filled in for the name of the next taxiway and the direction to turn. In this view of the process, error might come about by incorrectly filling the slot for the next taxiway name, but more probably, the slot for the direction of the turn to make.

We would like to argue in favor of an alternate view in which there are typically several *intentions-to-act* concurrently in process. The intentions may be established at different points in time. One or more of them may lead to a correct turn or to making an error at the intersection. A *winner-take-all* process leads to the execution of one of the intentions-to-act and the correctness of the outcome is the product of the winning intention. At the point of execution, the remaining intentions cease to contend. We label the process intention-to-act to suggest that the process is not the

product of a conscious decision process—it is not a deliberative planning process followed by a plan execution process. There is the immediacy of an automatic, atomic process rather than a sequential process of planning and acting. Each of the intentions-to-act is instantiated with established slot values, rather than with unfilled slot values to be filled by a deliberative process.

Most of the time there is more than one intention-to-act. In the nominal case where the first officer has provided the correct prompt for the turn, the turn is, most likely, simply executed in response to the prompt and most likely in accordance with a pre-existing intention. In lieu of the prompt from the first officer, the captain will act on a pre-existing intention that might lead to the execution of his/her intention to turn or alternatively to pause and query the first officer on the next turn. (We have not pursued the case of the captain's slowing or stopping the aircraft and querying the first officer.) That is, most of the time in the taxi environment, it is reasonable to expect that the captain has an intention-to-act in place and ready to be acted on.

Rather than having a single planning process with slots to be filled from various sources that is followed by a plan execution step, there are multiple intentions-to-act with selection through a non-conscious winner-take-all process. Each of the intentions-to-act has a complete set of immediately filled slots. In the following section, we provide the reasoning supporting this viewpoint.

Intentions-to-Act as Automaticity

At this point, we want to build the case for the idea that in performing relatively simple tasks like correctly executing the next taxiway turn, there may be several competing intentions-to-act. Most may arise as automatic processes that require little or no conscious deliberative thought. They may emerge from different ongoing processes competing in a winner-take-all process to determine the action taken. Occasionally, the winner will determine an action that is in error. During the course of this study, we have attempted to identify some of the sources for these intentions and to provide reasoned explanations on why the errors emerge.

For most of us, there are a broad range of everyday activities that we perform quickly and effortlessly—they appear to be automatic and involve little thought or conscious awareness (Logan, 1988a; James, 1890). Logan (1988a) characterizes this automaticity, the execution of these activities, as fast, effortless, autonomous, stereotypic, and unavailable to conscious awareness. That is, we experience them as fast, effortless, stereotypic, and unavailable to conscious awareness. They are autonomous in the sense that the acquisition of these skills comes about independent of any deliberate intention to learn them.

Logan (1988a) developed the "Instance Theory of Automaticity," a theory for how automatization is constructed. The theory was developed in part through a series of experiments in learning alphabet arithmetic—learning to solve problems of the type "A+2=?" where the answer is "C." Initially, most people solve these problems by explicitly counting out the required steps through the alphabet—they employ an algorithm that they step through when the problem is presented. Through experience they "learn or remember" the answers.

Logan suggests that *each* learned instance is remembered. When presented with a new problem, there is a concurrent attempt to access a remembered instance of a previous solution *and* an explicit algorithm-based problem-solving computation. The memory access is a comparatively fast process, the algorithm-based process comparatively slow. If the memory access is successful in retrieving a solution, there will be a rapid response to the posed problem. If the memory access is not successful, the response will be slower. Through experience, more and more solutions are acquired and at some point, the deliberative process is simply not a contender in the winner-take-all process. For any given problem, there may be several remembered solutions. Due to the remembering of each solution instance, there may potentially be several correct retrievals. It is the one that is first retrieval that determines the time required to solve the problem.

Logan (1988b) further argues that the memory traces that support automaticity may well support declarative as well as procedural knowledge. Logan (1988b) suggests that we "look more broadly for automatic processes. They need not be restricted to procedural knowledge or perceptual-motor skill but may permeate the most intellectual activities in the application environment." Bargh and Chartrand (1999) further suggest that limits on conscious, intentional control requires that non-conscious processes support much of moment-to-moment psychological life. Here we are suggesting that the captain's procedures for addressing the next turn in the taxiway sequence may sometimes be automatic and that while these will often lead to correct behaviors, they may sometimes lead to errors such as those seen in the baseline T-NASA experiments.

Intentions-to-Act as a Source of Error

Our review of the NASA-provided data on the T-NASA experiments pointed to two important factors that we felt deserved particular attention in our modeling effort. NASA (2001c) identified the importance of the location of the destination gate and its relation to the taxi route. Five errors occurred in 48 instances of required turns *away from* the shortest route to the concourse gate while only seven errors occurred in 534 instances of turns *toward* the concourse gate. At any given intersection, the aircrews had a bias to turn toward their destination concourse gate. When the correct turn was one away from the concourse gate, there was a greater tendency toward making an error.

The second observation was the straightforward one that time pressure can lead to error. There was a greater chance of error when a second turn in the taxi sequence closely followed the previous turn. The time pressure of a second turn closely following a first turn was an important factor in each of the errors that we generated in the modeling effort.

To date, four sources of contending intentions-to-act have been identified and modeled. The first is episodic memory—a source for habit-based actions. Similar situations have been encountered in the past and we have a ready source of responses that have worked well. These are

responses that in the past have proven successful and are generally able to carry us through most of the activities of the day. When they fail this is what Reason (1990) refers to as "strong-but-wrong." In our particular case, the aircrews had a history of previous landings at Chicago O'Hare.

A second source of intention-to-act is context-based expectation, driven by partial knowledge. Explicit partial information provided in the current situation prompts a particular intention. Within the taxi-framework, the captain knows the location of the concourse gate and, based on this knowledge, may reasonably have an expectation that the next turn will take them on the shortest route to the gate. These particular situation-specific information points are sufficient to set up an intention for the next turn.

The third source of intention-to-act is the remembrance of the taxi sequence as provided by the ground controller when the aircraft exited the landing runway. As the aircraft approaches a turn, several minutes may have passed since the ground controller provided the taxi directive. The remembrance may or may not be correct, but it can be the source of an intention-to-act.

The fourth source of intention-to-act in the taxi-framework, and the best-grounded source of intention, is the explicit prompt by the first officer based on written notes on the taxi directives provided by the ground controller. In the nominal case, the first officer's prompt will match the captain's intention and will lead to error free performance.

We modeled the contention between these intentions as a winner-take-all process mediated by priority and explored the impact of varying the priorities of the contending intentions. Within the winner-take-all framework, at the winning intention's transition from intention to action, the remaining intentions cease to contend—within the framework of the model, the procedures that would have implemented those intentions fail. The occurrence and timing of the events that drive the intentions determine how they play out, producing successful behaviors or mistakes that lead to an incorrect turn on the taxiway. In particular, to provide a window for error to occur, it was necessary to set up realistic event chains that prevented the first officer from providing the prompt on the next turn to the captain.

As we have suggested, the team-based nature of the taxi procedures makes them very robust and the challenge has been to create situations in which mistakes will lead to error. This effort focused on two error sequences, each requiring two turns in rapid sequence. For case one, there were two instances of the same error as crews took high-speed exit M7 from runway 9R. At the first intersection after the high-speed exit, each captain turned left toward the concourse gate rather right away from the gate as directed by the ground controller. In the second case, there were two scenarios that shared a similar turn sequence: after turning onto taxiway F in the first instance and M2 in the second instance, there was a quick right turn onto taxiway B. In each of the scenarios, one of the captains turned left rather than right. The errors were noteworthy, because in committing the error the captains each turned *away* from their intended concourse gate rather than toward the gate as directed.

Error Driven by Partial Knowledge Our hypothesis is that the incorrect turn following the high-speed exit (see Figure 1) was driven by the captain's expectation that the shortest route to the gate was the route to be taken. (The small arrows that denote the errors in Figure 1 indicate the incorrect left turns taken just after the high-speed runway exit. They are in red when viewed in color—in grayscale, they may be difficult to make out.) The intention-to-act arose at the point of the early discussion of the runway exit with the approach controller and the first officer. At this point, the captain knew the runway exit and the concourse gate, and might reasonably have expected to turn left from the high-speed exit at taxiway M taking him/her toward the gate. It became one intention contending to be executed at the first turn after exiting the active runway. From Reason's (1990) perspective, this is an automatic retrieval process based on similarity-matching and frequency-gambling that opens a window for error.

As the scenario played out in the nominal case, the first officer completed the task of taking notes on the taxiway
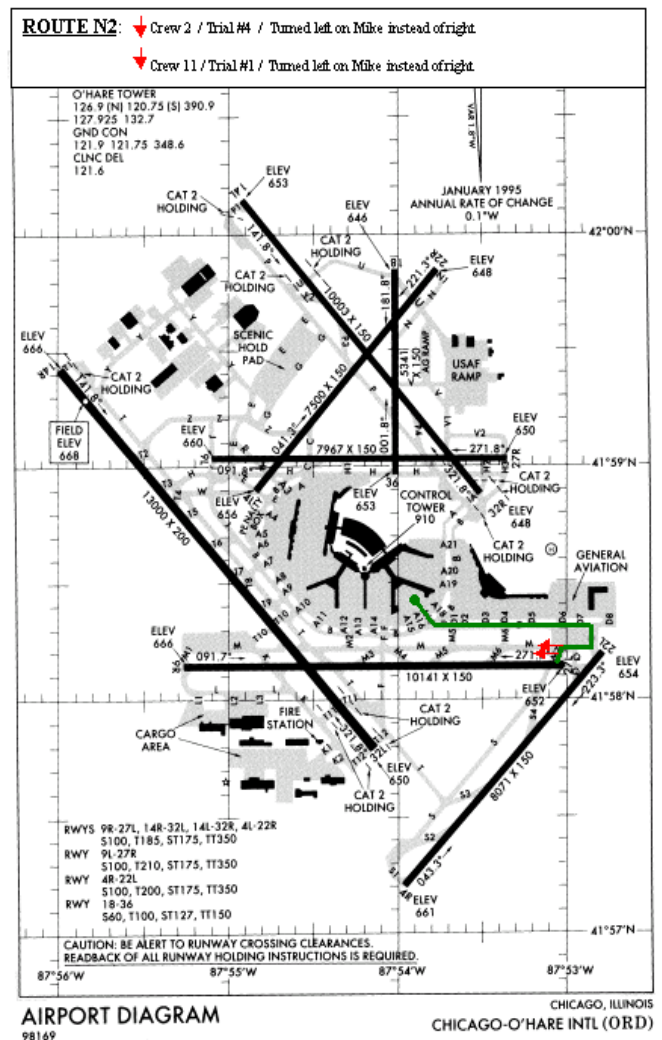

Figure 1: Errors driven by partial knowledge (NASA 2001b).

sequence and then prompted the captain on the first turn following the runway exit. The first officer's prompt triggered a new, contending intention-to-act on the captain's part. The new intention may or may not have been consistent with preexisting contentions. In the nominal case, it dominated and the captain turned right correctly. Given the correct prompt by the first officer, we deemed it highly unlikely that the captain would incorrectly execute the turn.

To open a window for an error to occur, it was necessary to construct a situation that reasonably occupied the first officer, preventing him/her from providing the captain with the explicit prompt on the upcoming turn. The very short run to the first turn after the high-speed exit was the essential factor. The first officer was already busy taking notes on the taxiway routing. Indeed, in some scenarios the taxiway routing was so lengthy that in the nominal case the first officer was still taking notes as the first turn was executed. In this scenario, this was not the case, hence a "mistake" was needed to additionally task the first officer. The failure to preset the radio frequency for the transfer to the ground controller provided the delay. The few seconds necessary to set the new radio frequency provided enough delay to prevent the first officer from prompting the captain before the turn. This was a mistake on the part of the aircrew in the sense that it is always incumbent upon them to complete an action at the earliest available time, rather than risk a situation such as this in which there are contending tasks in process.

Let us recap the captain's intentions-to-act as the aircraft approached the first turn onto taxiway M after the high-speed exit on taxiway M7. The first officer has been otherwise occupied and has not provided the captain with the explicit prompt on the upcoming turn. Based on the coping strategy described earlier, the captain might have a correct intention-to-act based on having attended to the ground controller's taxi directive and an incorrect intention based on the expectation of receiving a shortest route to the concourse gate. Much of the time the coping strategy might be expected to win the winner-take-all competition and lead to a correct turn—some of the time the expectation-based intention-to-act might be acted upon, leading to a taxiway error. Hence, a reasonable, grounded source for an error consistent with the T-NASA experiments has been identified and modeled.

Error Driven by Habit The second scenario examined the surprising cases in which an aircrew incorrectly turned away from the shortest course to the gate (see Figure 2). (In Figure 2, the small arrow denoting the error indicates the incorrect left turn taken just after the short north-bound segments near the center of the airport diagrams. It is in red when viewed in color—in grayscale, it may be difficult to make out.) The basic intention to take the shortest route to the gate would have led to the correct behavior, yet it was not the one acted upon. There were two instances of this error at similar intersections. In the first case (Figure 2), the aircraft was proceeding north on taxiway F and had been instructed to turn right onto taxiway B, but the captain turned left instead. In the second case, the aircraft was

proceeding north on taxiway M2 and had been instructed to turn right onto taxiway B, but the captain turned left. We speculated that a crew whose company gates were on the opposite side of the airport from those required by the scenario might incorrectly turn toward their company gates, exhibiting an error based on long established habit. Requiring an aircrew to proceed to a gate opposite in direction from their company gates might be considered an artifact of the particular scenario, but in a commercial air travel environment that has seen many company failures and mergers, it is not uncommon for aircrews to find themselves working for new companies with new gate locations.

The turn at which the errors occurred closely followed a previous turn, creating a time-pressured situation. Once again, we manipulated the situation such that the first officer was not able to provide a timely prompt to the captain on the upcoming turn. Conflicting taxiway traffic was present on the first officer's side of the aircraft during the approach to the first turn. The first officer informed the captain of the
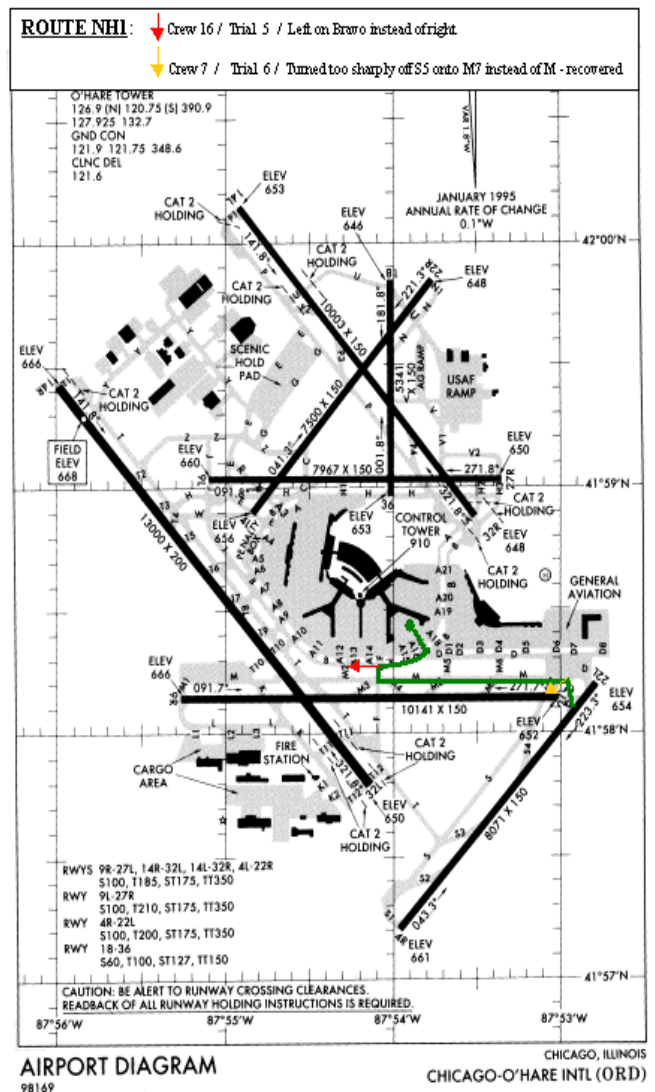

Figure 2: Error driven by habit (NASA, 2001b).

presence of the traffic and continued to monitor the other aircraft. Consequently, the first officer was delayed in going head-down to review his/her notes on the upcoming taxiway turn and checking the airport diagram. Following the delay, the first officer's prompt on the upcoming turn was immediately interrupted by a message from the ground controller directing the other aircraft to hold short of the upcoming intersection, allowing the first aircraft to proceed with the turn. Very slight changes in timing of the interruption would have opened the window for a timely and successful prompt.

In the absence of the prompt, there were still multiple intentions-to-act. As modeled, there were intentions-to-act based on the remembrance of the ground controller's taxi directive and on habit based in episodic memory. When the captain's habit-based intention-to-act won the winner-take-all competition and was acted upon, the error was committed. An informal post hoc analysis of the human subject trial error provided support for the speculation that the model represented (B. Hooey, personal communication).

Heuristically Guided Search of the Error Space

The incidence of error in the current-equipment T-NASA experiments was strikingly high when compared to the typical behaviors of professional aircrews. In general, the low frequency of mistakes and the even lower frequency of mistakes combining to produce errors renders a simple stochastic exploration of the behaviors space impractical. The robustness of aircrew team procedures that employ checking and cross-checking of critical actions means that most mistakes will be caught, further compounding the search task. Estimating error frequency for error types can also be a problem. The frequency of some errors (e.g., discrimination of taxiway signage) might be reasonably estimated; the frequency of others (e.g., the onset of a particular intention-to-act) is more difficult.

Timing is also critical. Very small variations in timing can open or close the window in which an error might occur. Timing was particularly critical in the scenario in which the habit-based error occurred. The combination of the demand on the part of the first officer for head-up time to monitor the approaching traffic and the precise moment of the ground controller's interruption of the first officer's prompt for the upcoming turn was necessary to open the window to error. It might well have been possible to generate many hundreds of runs, slightly varying several of the timings, and never have produced a single habit-based error.

To address this problem, we have employed a heuristically guided search of the space in which forced sequences of mistakes are generated, looking for those that lead to error. The errors produced to date are initial examples of the product of such a process. We have identified several novel potential sources of mistakes and worked to create situations in which they might reasonably be expected to occur. We have taken advantage of the time pressure inherent in the closely spaced turn sequences to manipulate the timing of events to construct sequences of

mistakes that do in fact lead to error. For the present, this heuristically guided exploration of the error space has been manipulated by hand. In the future, we would like to move toward a more automated exploration of the error space.

References

Bargh, J. A, & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462-479.

Deutsch, S. E. (1998). Interdisciplinary foundations for multiple-task human performance modeling in OMAR. *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society,* Madison, WI.

Deutsch, S. E., & Adams, M. J. (1995). The operator model architecture and its psychological framework. *Proceedings of the 6th IFAC Symposium on Man-Machine Systems.* Cambridge, MA.

Deutsch, S. E., & Pew, R. W. (2001). *Modeling Human Error in D-OMAR.* (Tech. Rep. BBN-8328). Cambridge, MA: BBN Technologies.

Hooey, B. L., & Foyle, D. C. (2001). A post-hoc analysis of navigation errors during surface operations: Identification of contributing factors and mitigating strategies. In *Proceedings of the 11th Symposium on Aviation Psychology.* Columbus, OH: Ohio State University.

Hooey, B. L., Foyle, D. C., & Andre, A. D. (2000). Integration of cockpit displays for surface operations: The final stage of a human-centered design approach. In *Proceedings of the AIAA/SAE World Aviation Congress.* Warrendale, PA: SAE International.

James, W. (1890). *Principles of Psychology.* New York: Holt.

Logan, G. D. (1988a). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527.

Logan, G. D. (1988b). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30,* 583-598.

NASA. (2001a). *Nominal task sequence.* NASA Working Paper. Moffett Field, CA: NASA ARC.

NASA. (2001b). *Routes and errors: Locations and types.* NASA Working Paper. Moffett Field, CA: NASA ARC.

NASA. (2001c). *Enhanced descriptions of off-route navigation errors.* NASA Working Paper. Moffett Field, CA: NASA ARC.

Reason, J. (1990). *Human Error.* Cambridge: Cambridge University Press.

Wickens, C. D., & Prevett, T. T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied, 1,* 110-135.