CONCURRENT TASK MANAGEMENT AND PROSPECTIVE MEMORY: PILOT ERROR AS A MODEL FOR THE VULNERABILITY OF EXPERTS

Key Dismukes
NASA Ames Research Center
Moffett Field, CA

In five of the 27 major U.S. airline accidents between 1987 and 2001 in which the NTSB found crew error to be a causal factor, inadvertent omission of a normal procedural step played a pivotal role. Such omissions are a form of *prospective memory* error. My research group is attempting to link real-world prospective memory phenomena with task demands and with underlying cognitive processes. I briefly summarize studies from three quite different but complementary approaches: ethnographic studies, analyses of accident and incident reports, and laboratory studies. Five types of situation presented prospective memory challenges: episodic tasks, habitual tasks, atypical actions substituted for habitual actions, interrupted tasks, and interleaving tasks/monitoring. An experimental study found that inadequate encoding, inadequate cueing, and competing demands for attention make individuals vulnerable to forgetting to resume interrupted tasks.

Between 1987 and 2001 twenty-seven major airline accidents occurred in the United States in which the NTSB found crew error to be a causal factor. In five of these accidents, inadvertent omission of a normal procedural step by pilots played a pivotal role (NTSB 1988, 1989, 1995, 1997, 2001). For example, in 1994 an airliner ran off the runway at LaGuardia after the crew rejected the takeoff at high-speed because they observed anomalous indications on their airspeed indicators. The NTSB determined that the anomalous indications occurred because the crew failed to turn on the pitot heat, a normal procedural step, that keeps the pitot input to the airspeed indicators from freezing in cold, wet weather. Two other accidents involved failing to set flaps and slats to takeoff position. A fourth accident involved failing to set hydraulic boost pumps to the high position before landing, preventing the landing gear from extending on command, and a fifth accident involved failing to arm the spoilers before landing, which combined with other crew errors and a wet runway to prevent the airplane from stopping before the end of the runway.

Obviously, in each of these flights many factors contributed to the outcome, but a central aspect of each was that highly experienced pilots forgot to perform a crucial procedural step that they had successfully executed previously without difficulty on thousands of flights. Why are skilled pilots vulnerable to such errors of omission? The answer, I suggest, lies at the intersection of an emerging field of research called prospective memory with the domain of concurrent task management. Prospective memory is distinguished by three features: (1) an intention to perform an action at some later time when circumstances permit, (2) a delay between forming and executing the intention, typically filled with activities not directly related to the deferred action, and (3) the absence of an explicit prompt indicating that is time to retrieve the intention from memory—the individual must "remember to remember" (Brandimonte, Einstein, & McDaniel, 1996). This third feature distinguishes prospective memory from traditionally studied retrospective memory. Typically, if queried after forgetting to perform an action, individuals can recall what they intended to do. Thus the critical issue in prospective memory is not retention of the content of intentions but retrieval of those intentions at the appropriate moment, which is quite vulnerable to failure.

A complete theoretical account of prospective memory is not yet available, but most theories focus on individuals noticing an environmental cue associated in memory with the deferred intention. If sufficient activation spreads from the noticed cue to the stored intention it is retrieved. Theories differ as to whether the individual must actively monitor for the cue (Smith, 2003), but evidence seems to support involvement of automatic retrieval mechanisms, which may be supplemented in some situations by conscious monitoring for cues (McDaniel & Einstein, 2000).

My research group is attempting to link real-world prospective memory phenomena with the emerging picture of underlying cognitive processes. I briefly describe studies from three quite different but complementary approaches: ethnographic studies, analyses of accident and incident reports, and laboratory studies.

Aviation studies

Airline operations lend themselves to the study of skilled human performance and human error because these operations are highly standardized, with formal written operating procedures that cover almost every aspect. Because most aspects of flight operations are explicitly scripted, we can readily observe deviations from what is prescribed. Also a fair degree of consensus exists among experts over what actions are appropriate or inappropriate in most normal situations.

We conducted three studies that helped us to identify the kinds of tasks involving prospective memory in airline flight operations and the most common forms of associated error. An ethnographic study focused on a particular aircraft type to allow in-depth analysis (the Boeing 737, one of the most commonly used airplanes in the transport industry). We reviewed written operating procedures, participated in classroom and flight simulation training at two major airlines, and observed a large number of flights from the cockpit jumpseat (Loukopoulos, Dismukes, & Barshi, 2003). A second study analyzed National Transportation Safety Board (NTSB) reports for the 19 major U.S. airline accidents attributed to crew error between 1990 and 2001 (Dismukes, Berman, & Loukopoulos, in press), and a third study sampled 20 percent of all air carrier reports submitted to the Aviation Safety Reporting System (ASRS) over a 12 month period to obtain reports involving any type of memory error (Nowinski, Holbrook, & Dismukes, 2003). A startling finding of this ASRS study was that, of the 75 reports from airline pilots that provided sufficient information to clearly identify a memory failure, 74 involved prospective memory rather than retrospective memory (memory for content).

From these studies (which also address topics beyond prospective memory), we concluded that prospective memory demands in cockpit operations emerge in five types of task situations. These five types were identified by analyzing each accident and incident in which a pilot forgot to perform an intended task and determining what other tasks were being performed concurrently, whether the forgotten task was habitual, what cues would normally be present to trigger retrieval, and whether those normal cues were actually present.

1) *Episodic tasks*. In these situations pilots must remember to perform at a later time some task that is not habitually performed at that time. For example, an air traffic controller may instruct a crew to report passing through 10,000 feet while the crew is still at 15,000 feet, requiring the crew to hold this instruction in memory for perhaps five minutes. During that time, the crew typically will perform other tasks, often tasks that divert attention away from the altimeters. We have found that, in many real-world situations, ongoing tasks direct individuals' attention entirely away from cues that might trigger retrieval, and we argue that this is a major source of variance in performance (Holbrook, Dismukes, & Nowinski, 2005). Most laboratory research on prospective memory has focused on episodic tasks.

2) *Habitual tasks*. Crews perform many tasks in the course of a flight, and many tasks involve multiple steps. Most of the tasks and many of the intermediate steps of the tasks are specified by written procedures, and are normally performed in the same sequence. Thus, for experienced pilots, execution presumably becomes largely automatic and does not require deliberate search of memory in order to know what to do next. Pilots do not need to form an episodic intention to perform each task and each action step—rather the intention is implicit in the action schema for the task, stored as procedural memory. It would be uncommon for a pilot to arrive at work thinking "I will lower the landing gear today when I turn onto final approach" (and it would be rather alarming if a pilot found this necessary).

Remembering to perform habitual tasks seems quite reliable normally—individuals in aviation and in everyday life perform enormous numbers of habitual activities with few complaints—but performance is undermined if normally present cues are for some reason removed. For example, at many airlines, crews normally set flaps to takeoff position immediately following completion of the After Engine Start checklist and before starting to taxi. But what happens if the crew must defer setting the flaps until after taxi to prevent freezing slush on the taxiway from being thrown up on the flaps? The cues that normally trigger crews to set the flaps are no longer present—this action is now out of sequence, temporally separated from completion of the After Engine Start checklist and removed from the normal environmental context provided by being at the ramp.

By deferring the task of setting the flaps the crew has essentially changed a habitual task to an episodic task, but they may not realize that this increases their vulnerability to forgetting because the normally present cues and context are removed.

3) *Atypical actions substituted for habitual actions*. Circumstances sometimes require crews to deviate from a well-established procedural sequence. For example, through long experience departing from a certain airport, a crew might come to know that the Standard Instrument Departure procedure (a written instrument procedure) requires them to turn left to 300 degrees upon reaching 2000 feet. Taking this action would become habitual for the crew. If on rare occasion a controller told them to turn to 330 degrees instead of 300, the crew would have to form both an episodic intention to turn to 330 degrees and an intention to inhibit their habitual response of leveling the wings at 300. However, busy with other tasks, crews often revert to the habitual action, a form of error Reason (1990) described as habit intrusion.

4) *Interrupted tasks*. Crews are often interrupted when performing cockpit procedures, especially when at the gate preparing the airplane for departure (Dismukes, Young, & Sumwalt, 1998; Latorella, 1999). Flight attendants, gate agents, mechanics, and jumpseat riders require the pilots' attention as the pilots perform a fairly long sequence of procedural steps before starting the engines. Interruptions are so abrupt, salient, and common that pilots may do little if anything to encode an explicit intention to resume the interrupted task. A common error is to go on to the next task after the interruption, rather than returning to the interrupted task. In many cases the perceptual cues available in the cockpit do not provide a salient indication of the status of the interrupted task, and the perceptually rich environment of the cockpit is associated with many tasks that remain to be done at this point.

5) *Interleaving tasks and monitoring*. While performing ongoing tasks pilots are often required to monitor the status of other tasks. Some tasks, such as the requirement to report passing through an altitude, previously discussed, involve monitoring for an event that is known will occur. In other situations pilots must monitor for events that occur infrequently if at all. For example, when flying in visual meteorological conditions pilots must scan outside the cockpit windows for other airplanes that might be on a conflicting path. Colvin, Dodhia, & Dismukes, (2005) found that, when conflicting traffic was rare, pilots' scanning of the full visual field became infrequent.

We speculate that it is difficult to maintain the monitoring task goal in working memory when the result of each inspection of the monitored scene reveals that no event has occurred. In this sense the monitoring aspect of the pilots' dual tasks somewhat resembles vigilance tasks (Parasuraman, 1986). Apparently humans are wired to allocate attention heavily toward sources of high information content, and thus have difficulty maintaining monitoring for low probability events, even when those events may have high consequences (see Wickens, Goh, Helleberg, Horrey, & Talleur, 2003, for a model of attention allocation among tasks). But this sort of monitoring differs from traditionally-studied vigilance tasks in that the pilot must interrupt an ongoing task and shift attention to the thing being monitored. When the pilot goes too long without shifting attention, the monitoring task may slip from working memory, and then must somehow be retrieved, just as in other types of prospective memory situation.

An Experimental Study of Interruptions

Dodhia and Dismukes (2005) hypothesized that individuals are vulnerable to forgetting to resume interrupted tasks in large part because of three reasons. (1) The salient intrusion of many interruptions quickly diverts attention and discourages encoding explicit intentions and identifying cues to resume the interrupted task. If no explicit intention is encoded, then remembering to resume the interrupted task will depend on noticing happenstance cues that remind the individual of the status of the interrupted task and the implicit intention of completing all tasks. Even if an intention is explicitly encoded, the conditions for resuming the interrupted task are likely to be framed only as "after the end of the interruption"; individuals are often not in a position to identify and encode specific perceptual cues likely to be present at the end of the interruption. (2) Cues indicating the window of opportunity for resuming the interrupted task at the end of the interruption may not closely match the form in which the intention (implicit or explicit) to resume the interrupted task is encoded. The end of the interruption is not a simple perceptual cue but a state of affairs that requires recognizing, interpreting and integrating diverse perceptual cues. If the individual does not consciously monitor for the end of the interruption, the constellation of perceptual cues may fail to trigger recognition that the interruption has ended. (3) The end of interruptions in real-world situations is often followed immediately by other task demands that may not allow the individual sufficient time to fully process and interpret environmental conditions signifying that the interruption is over or to retrieve the associated intention (Holbrook et al., 2005; Loukopoulos et al., 2003). Further, activation from environmental cues associated with these other task demands may support retrieval of the goals associated with these task demands preferentially over retrieval of the goal to resume the interrupted task.

We designed an experimental paradigm to investigate these three themes. Experiment participants were required to answer a series of questions resembling the Scholastic Aptitude Test, arranged in blocks of different question types (e.g., analogies, vocabulary, math). They were instructed that when blocks were interrupted by the sudden onset of a different block of questions they should remember to return to the interrupted block after completing the interrupting block and before continuing to the next block in the series. In the baseline (control) condition, these occasional interruptions were abrupt–the screen with the question participants were currently working on was suddenly replaced, before they could answer the question, with a screen with a different type of question, and the background color of the screen changed.

After the end of the interrupting block, a screen appeared for 2.5 seconds with the message "Loading next section" (this screen also appeared between all blocks that were not interrupted) and then the next block of questions appeared without any reference to the interrupted block. Without receiving any explicit prompt, participants had to remember to return to the interrupted block by pressing a key. Participants in the baseline condition frequently forgot to resume the interrupted task and instead continued with the next block in the series after an interruption—the proportion of successful resumptions of the interrupted task in the baseline condition was 0.48.

To address our first hypothesis, that the intrusion of a sudden interruption discourages adequate encoding of an intention to resume the interrupted task, we implemented an encoding reminder condition in which the interruption began with a four second text message "Please remember to return to the block that was just interrupted." This manipulation increased the proportion of resumptions from the baseline condition of 0.48 to 0.65, which was highly significant statistically (as were the results of all other manipulations, discussed below).

It was not clear whether the encoding reminder manipulation was effective at improving performance because of the explicit reminder or because of the additional four second delay before participants had to start performing the interrupting task. We therefore performed an encoding pause manipulation in which participants saw only a blank screen for four seconds at the beginning of the interruption. This manipulation also improved performance to 0.65. We interpret these results to indicate that a pause before starting to perform an interrupting task allows individuals time to recognize the implications of being interrupted and to encode information that helps them to remember to resume the interrupted task. The explicit reminder to resume the interrupted task apparently did not provide any additional encoding advantage.

We also addressed our second hypothesis that individuals are likely to forget to resume interrupted tasks because they do not encounter explicit cues signaling the end of the interruption. In the retrieval reminder condition, participants received a message "End of interruption" for 2.5 seconds while the next block was loading. This message appeared above the "Loading next section" message that appeared in all conditions. In support of our hypothesis, this manipulation improved performance to 0.90.

Finally, we addressed our third hypothesis, that individuals sometimes forget to return to an interrupted task because the end of interrupting tasks is often quickly followed by other task demands that do not allow the individual time to fully process and interpret environmental conditions and to retrieve the intention to resume the interrupted task. We created a retrieval pause condition in which the delay between the end of the interrupting task and the beginning of the next block was increased to 8-12 seconds and a countdown clock appeared to display the remaining time to the next block. This manipulation was intended to make clear to participants that they had plenty of time before new task demands would begin. Resumption performance increased to 0.88, supporting the idea that people fail to resume interrupted tasks in part because their attention is quickly diverted to new task demands arising after interruptions end.

In addition to theoretical implications, this study suggests practical ways individuals can reduce their vulnerability to forgetting to resume interrupted tasks. Taken together, our studies illustrate that understanding real-world performance requires integrating diverse research approaches, including observing performance in real-world settings and well-controlled laboratory studies. This integrative approach enhances the power of both field and laboratory studies.

Acknowledgements

This research was supported by NASA's Aviation Safety Program. Thanks to Kim Jobe and Jessica Nowinski for reviewing the draft manuscript.

References

Brandimonte, M., Einstein, G. O. & McDaniel, M. A. (1996). *Prospective Memory: Theory and Applications*. Mahwah, N.J.: Lawrence Erlbaum.

Colvin, K., Dodhia, R., & Dismukes, R. K. (2005). Is pilots' visual scanning adequate to avoid mid-air collisions? In *Proceedings of the 13th International Symposium on Aviation Psychology,* (pp. 104-109). Oklahoma City.

Dismukes, R.K., Berman, B.A., & Loukopoulos, L.D. (in press). *The Limits of Expertise: Rethinking Pilot Error and the Causes of Airline Accidents.* Aldershot, U.K.: Ashgate.

Dismukes, R. K., Young, G. E., & Sumwalt, R. L. (1998). Cockpit interruptions and distractions: Effective management requires a careful balancing act. *ASRS Directline, 10*, 4-9.

Dodhia, R. M. & Dismukes, R. K. (2005, January). *A task interrupted becomes a prospective memory task.* Presented at the Biennial Meeting of the Society for Applied Research in Memory and Cognition, Wellington, New Zealand.

Holbrook, J. B., Dismukes, R. K., & Nowinski, J. L. (2005, January). *Identifying sources of variance in everyday prospective memory performance.* Presented at the Biennial Meeting of the Society for Applied Research in Memory and Cognition, Wellington, New Zealand.

Latorella, K. (1999). *Investigating interruptions: Implications for flightdeck performance*. (NASA Technical Memorandum 209707). Hampton, VA: Langley Ames Research Center.

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2003). Concurrent task demands in the cockpit: Challenges and vulnerabilities in routine flight operations. In *Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 737-742). Dayton, OH: The Wright State University.

McDaniel, M. A. & Einstein, G. O. (2000). Strategic and automatic memory processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14*, S127-S144.

Nowinski, J. L., Holbrook, J. B., & Dismukes, R. K. (2003). Human memory and cockpit operations: An ASRS study. In *Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 888-893). Dayton, OH: The Wright State University.

National Transportation Safety Board (1988). *Aircraft accident report: Northwest Airlines, McDonnell Douglas DC-9-82, N312RC, Detroit Metropolitan Wayne Country Airport, Romulus, Michigan, August 16, 1987*. (NTSB/AAR-88/05). Washington, D.C: National Transportation Safety Board.

National Transportation Safety Board (1989). *Aircraft accident report: Delta Air Lines, Boeing 727-232, N473DA, Dallas-Fort Worth International Airport, Texas, August 31, 1988*. (NTSB/AAR-89/04). Washington, D.C: National Transportation Safety Board.

National Transportation Safety Board (1995). *Aircraft accident report. Runway overrun following rejected takeoff. Continental Airlines flight 795, McDonnell-Douglas MD-82, N18835, LaGuardia Airport, Flushing, New York, March 2, 1994*. (NTSB/AAR-95/01). Washington, D.C: National Transportation Safety Board.

National Transportation Safety Board (1997). *Aircraft Accident Report: Wheels-Up Landing, Continental Airlines Flight 193, Douglas DC-9, N10556, Houston, Texas, February 19, 1996*. (NTSB/AAR-97/01). Washington, D.C: National Transportation Safety Board.

National Transportation Safety Board (2001). *Aircraft accident report: Runway overrun during landing, American Airlines flight 1420, McDonnell Douglas MD-82, N215AA, Little Rock, Arkansas, June 1, 1999*. (NTSB/AAR-01/02). Washington, DC: National Transportation Safety Board.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance (Vol. II: Cognitive processes and performance)* (pp. 43.1-43.39). New York: Wiley.

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

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 29*(3), 347-361.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human Factors, 45*(3), 360-380.