**Crew observations show that checklists
and monitoring are not as effective as generally assumed.**

Designing a Better Error Trap

BY BENJAMIN A. BERMAN AND R. KEY DISMUKES

Early morning at the gate, powering up the jet from cold. Flow-scan the overhead panel, as you have done so many times before. Up and down, left to right. All the switches are in their usual positions. Last is the air panel — six switches and two rotary selectors. A quick glance shows they are good. You call for the checklist. The first officer's first challenge is "Pressurization?" Your eyes go to the landing altitude rotary selector on the air panel. "Set," you reply.

It is still dark after takeoff. Climbing through 3,000 ft, the first officer, the flying pilot, calls, "Flaps up, 'After Takeoff' checklist." You run your hands around the overhead panel, turning off the

ignition and auxiliary power. Pressurization check: A peek at the differential gauge shows that it is off the lower peg. Just then the controller instructs you to contact departure. After acknowledging, you pick up the checklist. "Pressurization?" Remembering your earlier glance at the gauge, you reply, "Checked."

Through 15,000 ft now, and an insistent beeping jars your senses. The takeoff warning horn. Why now? While you think about this, the master caution light comes on, indicating an equipment cooling fan failure. As you get out of your seat to check the fan's circuit breakers, you tell the first officer to keep flying. You stand up, turn around and feel a bit woozy. The last thing you remember is deciding, for some reason, to sit down in the narrow aisle behind the pilot seats.

Accident investigators comb through the wreckage for clues and determine you did not notice that the pressurization system selector on the air panel had been left on "MAN" (manual) by the maintenance department. The pressure differential had increased enough in manual mode to let you see the gauge off zero but not enough to maintain a livable atmosphere as the aircraft climbed. It is likely you forgot that the takeoff warning horn, which you had heard during systems tests before every flight, doubles as a cabin altitude warning. The conclusion: Both pilots succumbed to hypoxia because they did not identify, or react to, a lack of pressurization.

A sequence much like this occurred on Aug. 14, 2005, as a Helios Airways Boeing 737 climbed out from Larnaca, Cyprus (*ASW*, 1/07, p. 18). Automation kept the aircraft aloft and on its programmed flight plan until the fuel was exhausted over Grammatiko, Greece.

Although such accidents are extremely rare, they point to the crucial

roles played by checklists and monitoring in helping pilots catch system malfunctions and human error, and manage the challenging situations that sometimes arise on routine flights.

Line Observations

To find out how checklists and monitoring work in actual practice, we observed line operations during 60 flights conducted by three air carriers from two countries.[1] We used a structured technique to observe and record checklist and monitoring performance, and situational factors that might affect performance. Because an important function of checklists and monitoring is to catch, or "trap," operational errors, we also recorded deviations in aircraft control, navigation, communication and planning. When a deviation was observed, we tracked whether crewmembers identified and corrected it, and whether there were any consequences that might affect the outcome of the flight.

During the 60 flights, we recorded 899 deviations, of which 194 were in checklist use, 391 in monitoring and 314 in operating procedures (Table 1, p. 14). The total number of deviations per flight ranged from one to 38.

Many of the deviations we observed were errors. For example, one airline had a mixed 737 fleet, with a few aircraft requiring the first officer to place the pressurization system in flight mode during the flow portion of the "After Start" checklist procedure. On one flight, perhaps reverting to the procedure required for the more common aircraft, the first officer omitted this during the flow check. The pilots then did not notice the incorrect system configuration while conducting two subsequent checklists, both of which included verification of the relevant panel settings.

Some deviations, however, were not necessarily intrinsic errors. For example, several involved a standard operating procedure (SOP) at all three airlines that required the monitoring (nonflying) pilot to make a callout 1,000 ft prior to reaching each assigned altitude during climb and descent. We observed 137 instances of pilots omitting this callout or making it late. Climb and descent are busy periods, and at times a pilot may need to give priority over a callout to other tasks, such as air traffic control (ATC) communications. Consequently, omitting or delaying this callout may sometimes be a strategic workload management choice rather than an error.

This is not to suggest that the 1,000-ft callout is trivial. On the contrary, it ensures that both pilots concur about the altitude target, directs the attention of a flying pilot who might be distracted back to the impending level-off and draws both pilots' attention to what the autopilot is supposed to be doing.

Airlines should examine their SOPs to specifically define the objectives of each procedure and to determine whether it is realistic to assume that pilots can perform the procedure reliably under actual line conditions. Pilots must be aware that in deviating from any procedure, they might be giving up safety margin that is not apparent.

Checklist Deviations

Among the most common deviations in checklist usage was incorrect application of the *flow and check* procedure implemented by the three airlines. The procedure involves using a memory-based flow pattern for setting systems and controls, and then following up with verification using a printed or electronic checklist.

In 48 of the 194 checklist deviations recorded, the flow and check procedure was not performed correctly. One or

J.A. Donoghue

Deviations Observed on 60 Line Flights		
Category	**Deviation**	**Number**
Checklists	Flow-check as read-do	48
	Responded without looking	43
	Item omitted/incomplete/incorrect	42
	Poor timing	31
	Performed from memory	17
	Not initiated	13
	Total	194
Monitoring	Callout late or omitted	211
	Not monitoring aircraft state or position	67
	Verification omitted	113
	Total	391
Primary procedures	Systems configuration	62
	Contingency planning/execution	57
	Crew — crew coordination	56
	Automation — FMS	40
	Crew — ATC coordination	33
	Automation — MCP	18
	Conducting unstabilized approach	10
	Crew — ground personnel coordination	8
	Profile planning/execution	7
	Lateral path control	7
	Crew — flight attendant coordination	6
	Aircraft configuration	4
	Vertical path control	3
	Automation — head-down	2
	Airspeed control	1
	Total	314
Grand total		**899**

ATC = air traffic control; FMS = flight management system; MCP = mode control panel

Source: Benjamin A. Berman and R. Key Dismukes

Table 1

both pilots tasked with the flow procedure did not do it or attended to only some of the flow items. As a result, most items were performed only while using the checklist, eliminating the protective redundancy designed into the flow and check procedure; other items — those that were in the flow procedure but not repeated in the checklist — were not completed.

Many people find it difficult to force themselves to carefully check something twice within a brief period. A pilot may consider it wasteful of limited time and attention, and less efficient than combining the flow and the checklist into a single sequence of actions. If airlines want to maintain the error-trapping value of a redundant flow and check procedure, they must explicitly acknowledge this human tendency and explain to pilots why they are asked to check things twice. Airlines should clearly define which items should be double-checked and which responses can rely on a memory of having performed the item during the flow. Airlines also should review normal checklists to eliminate excessive repetition of items on the flow and the checklist.

Looking Without Seeing

We observed 43 instances in which checklist items were responded to without effective visual verification. In some cases, the responses were incorrect. For example, a first officer challenged, "Doors?" and the captain responded, "Closed," although the aft cargo door was actually open, as indicated on the overhead panel. The captain was looking down at his flight bag when he responded. The first officer caught the error, however.

On another flight, the captain responded, "On," to the challenge "APU [auxiliary power unit] bleed?" but the bleed was off. Because the captain was looking at the bleed switch when he made the incorrect response, this may have been an instance of "looking without seeing," in which we see what we *expect* to see, rather than what is actually there.

We observed a pilot using a nice technique of pointing to each item on the overhead panel as he gave the response. This makes the checklist more reliable by drawing both pilots' attention to the items being verified, and it can also slow the pace of checklist execution just enough to make checking more effective. In general, taking a few extra seconds to perform an error-trapping procedure in a deliberate manner — that is, carefully and thoughtfully — makes it much more effective. The "point and shoot" technique is worth adopting, and airlines should promote and train deliberateness.

Checklist items were omitted or performed incompletely or incorrectly in 42 instances. For

example, the checklist item "hydraulics" had a specified response of "Set and checked," referring to setting the pump switches on the overhead panel to the "ON" position and checking the pressure gauges on the forward instrument panel. Some pilots looked only at the overhead panel before making the specified response, omitting the other item, the gauge indications, that was to be verified. This shows the vulnerability to error of checklist designs that include more than one item on a single challenge-response element, and the subtlety of breakdowns in this area. We suspect that many of the pilots involved in this kind of deviation were not even aware of the omission.

Another common checklist deviation was initiating a checklist at a bad time. We observed this in 31 instances. Some were delayed initiations, with heavy workload a key factor; others involved pilots calling for a checklist when it interfered with other tasks and posed a significant distraction or workload spike. For example, a captain called for the "Taxi" checklist just as the aircraft was approaching a runway intersection, drawing the first officer's attention away from visually clearing

the taxi path from his side of the flight deck. This is an example of an error-trapping procedure that can potentially detract from safety when not handled properly. Pilots can reduce this risk by exercising proactive workload management, deliberately choosing the optimal time to perform a checklist (within the guidelines of the SOP) so as to minimize interference with other tasks. Airlines should train this mode of workload management, and reinforce it in line checks and line observations.

Deviations in Monitoring

Among the 391 monitoring deviations that we observed, 211 involved callout omissions. Callouts are the outward manifestations of monitoring that are scripted into SOPs and are easier to observe than other aspects of monitoring. Some omitted callouts more clearly undermined flight safety than the "1,000 to go" callouts previously discussed. For example, a flight crew was engrossed in increasing the descent gradient to accommodate a "slam dunk" ATC clearance when the monitoring pilot omitted the callout at 1,000 ft above airport elevation. This illustrates the tendency of pilots to

shed monitoring when primary control task workload is high and the corollary that monitoring tends to drop out of the picture just when it is needed most.

Verification omissions occurred in 113 instances. In one case, while descending through Flight Level (FL) 310 (approximately 31,000 ft), the flight crew received clearance to FL 240. The first officer set and called out the new altitude, but the captain was distracted by conversation and did not verify the new altitude on the primary flight display. There was no adverse outcome because the first officer had set the altitude correctly.

Potentially more consequential was an instance in which the first officer transposed the digits of a heading assigned by ATC while the captain was occupied with taxiing the aircraft onto the runway. The captain did not verify the heading selection at this busy time. The error was not trapped. In this case, the observer spoke up about the heading mis-selection to reduce the risk of a traffic conflict after departure.

Another frequent deviation was not monitoring the aircraft, observed in 67 instances. Both the flying pilot and monitoring pilot are required to attend to the aircraft. We observed numerous instances of pilots looking elsewhere as the aircraft began turning or leveling off at an assigned altitude, most often while under autopilot control. Not monitoring the aircraft suggests over-reliance on automation, an understandable reaction to automation's high reliability. But accidents and incidents have happened when the automation was misprogrammed. Automation does fail occasionally, but because it generally is so reliable, pilots likely do not even realize when they may, at least at times, no longer be actively monitoring the aircraft.

© Chris Sorensen Photography

Procedural Deviations

The 314 deviations in primary procedures included 62 involving configuration of equipment/systems. An example was when a captain turned on the engine anti-ice system before the airplane entered the clouds in icing conditions but neglected to turn on the engine ignition.

Deviations in planning for, or responding to, contingencies occurred in 57 instances. For example, an airplane was at 6,000 ft and near the end of a flight when ATC transmitted, "Braking action fair reported by all types." The crew made no comment in response, and they did not recalculate landing distance for the reported braking condition.
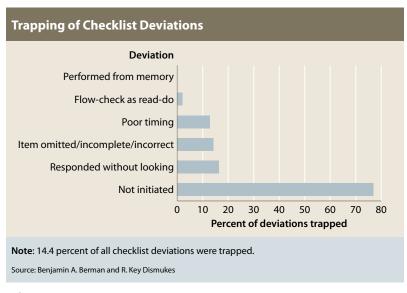
Trapping of Checklist Deviations

Note: 14.4 percent of all checklist deviations were trapped.

Source: Benjamin A. Berman and R. Key Dismukes

Figure 1

Trapping of Monitoring Deviations

Note: 5.6 percent of all monitoring deviations were trapped.

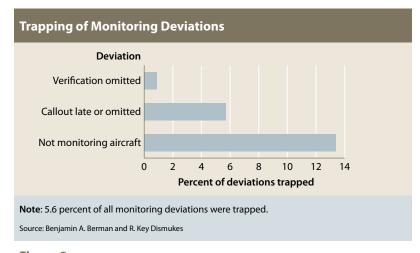Source: Benjamin A. Berman and R. Key Dismukes

Figure 2

We recorded 56 deviations in crew-crew coordination. In one instance, a flight crew was cleared to navigate directly to a fix; the captain entered and executed the route change without waiting for the first officer to confirm the change.

Deviations in data entry or in use of the flight management system or the mode control panel occurred in 40 and 18 instances, respectively. An example was a first officer who did not arm the autopilot to capture the instrument landing system (ILS) localizer as the flight neared the final approach course.

Effectiveness of Trapping

Overall, only 18 percent of the observed deviations were trapped by the crew. However, the efficiency of the trapping varied dramatically among the deviation types. More than 14 percent of the checklist deviations were trapped (Figure 1), while only about 6 percent of the monitoring deviations were caught (Figure 2). The best performance was in primary procedural deviations, with more than 35 percent trapped (Figure 3). However, there were eight instances in which flight crews failed to reject unstabilized approaches before or upon reaching the point at which a go-around was required by SOPs, and there were 10 discrete deviations during these approaches in which crews then did not challenge or trap their continuation of the approach while unstabilized.

Pilots trapped most erroneous mode control panel entries, most system misconfigurations and most failures to call for a checklist. In contrast, they rarely caught deviations in contingency planning, crew-crew coordination, monitoring and most aspects of checklist execution. From the jump seat, we were not able to distinguish whether deviations by one pilot were not noticed by the other pilot or whether the other pilot noticed but chose not to speak up.

One of the key discoveries from our study was that, although primary procedures most often were performed as prescribed, checklists and monitoring currently do not trap all procedural threats and errors to the degree that the aviation industry generally assumes. For example, even though slightly more than half of the 62 instances

of system misconfiguration were trapped, many of these events were not identified or corrected. The industry needs more reliable trapping for this and many other kinds of primary procedural deviations.

Most checklist and monitoring deviations were not trapped either by the flight crewmembers or by others. It appears that pilots are not likely to notice or take corrective action when checklists and monitoring have been weakened and their error-trapping functions cannot be relied upon. This may remain as a latent threat, allowing a primary procedural deviation to slip through.

Captains and first officers, and flying pilots and monitoring pilots, made about the same number of deviations overall. However, we found that first officers were significantly less effective at trapping errors while they were performing the monitoring role; they caught 12.1 percent of the deviations that captains made as the flying pilot, while captains caught 27.9 percent of deviations that first officers made as the flying pilot. Previous studies based on flight simulator observations and on accidents found a similar disparity. The greater difficulty that first officers face in challenging their captains (compared to the reverse) is clearly a stubborn problem for which a solution has not yet been found.

Implications

In our full report, we discuss factors that make even experienced, conscientious pilots vulnerable to the observed deviations. It is naïve to think that any crew can always perform perfectly in real-world conditions; nevertheless, our findings show that checklist and monitoring performance can be improved. In responding to these findings, airlines must not assume that the deviations are the result of laziness. Pilots face interruptions and concurrent task demands
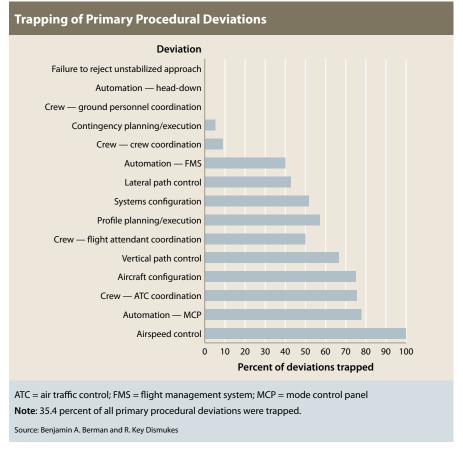

Trapping of Primary Procedural Deviations

ATC = air traffic control; FMS = flight management system; MCP = mode control panel

Note: 35.4 percent of all primary procedural deviations were trapped.

Source: Benjamin A. Berman and R. Key Dismukes

Figure 3

during actual line operations, and idealized SOPs do not take these factors into account. Also, pilots cope with operating procedures and equipment designs that sometimes are poorly matched to the ways the human mind processes information. Finally, pilots may slip into rushing through procedures when they are under the time pressures now common in airline operations; neither pilots nor airlines may recognize just how much rushing undermines reliable performance.

For these reasons, simply admonishing pilots to follow procedures as written is unlikely to improve performance. Rather, we encourage airlines to analyze actual operations thorough line observations, revise procedures and practices as needed, provide training to help pilots understand the cognitive nature of vulnerability to error, and provide specific techniques

to reduce that vulnerability. Pilots, flight managers, procedures designers, equipment designers and scientists should work together in this effort. The full report of our study provides detailed suggestions for reducing vulnerability and improving deviation trapping.

Benjamin A. Berman is a senior research associate at the U.S. National Aeronautics and Space Administration (NASA) Ames Research Center/ San Jose State University and a pilot for a major U.S. air carrier. R. Key Dismukes, Ph.D., recently retired from NASA as chief scientist for aerospace human factors at the Ames Research Center.

Note

1. This article is based on a study funded by NASA and the U.S. Federal Aviation Administration. When published, the full report, *Checklists and Monitoring in the Cockpit: Why Crucial Defenses Sometimes Fail*, will be available from <human-factors.arc.nasa.gov/ihs/flightcognition/>.