THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, *11*(1), 1–14

FORMAL PAPERS

Aircrews and Automation Bias: The Advantages of Teamwork?

Kathleen L. Mosier
Department of Psychology
San Francisco State University

Linda J. Skitka
Department of Psychology
University of Illinois at Chicago

Melisa Dunbar and Lori McDonnell
San Jose State University Foundation
NASA Ames Research Center
Moffett Field, CA

A series of recent studies on *automation bias,* the use of automation as a heuristic replacement for vigilant information seeking and processing, has investigated omission and commission errors in highly automated decision environments. Most of the research on this phenomenon has been conducted in a single-person performance configuration. This study was designed to follow up on that research to investigate whether the error rates found with single pilots and with teams of students would hold in the context of an aircraft cockpit, with a professional aircrew. In addition, this study also investigated the efficacy of possible interventions involving explicit automation bias training and display prompts to verify automated information. Results demonstrated the persistence of automation bias in crews compared with solo performers. No effects were found for either training or display prompts. Pilot performance during

Requests for reprints should be sent to Kathleen L. Mosier, Department of Psychology, San Francisco State University, 1600 Holloway, San Francisco, CA 94132. E-mail: kmosier@sfsu.edu

the experimental legs was most highly predicted by performance on the control leg and by event importance. The previously found phantom memory phenomenon associated with a false engine fire event persisted in crews.

The availability of increasingly sophisticated technology, as well as changes in the National Airspace System, have encouraged and expanded the use of automated aids in all aircraft operations—commercial, commuter, corporate, and general aviation. Although the impact of automation in terms of numbers of decision errors has been generally positive, reduction of some types of errors has been accompanied by the introduction of new varieties of automation-related errors. In particular, the availability of automation and automated decision aids feeds into a general human tendency to travel the road of least cognitive effort (Fiske & Taylor, 1994). Decision makers have displayed a tendency to use these aids as a replacement for vigilant information seeking and processing, a phenomenon referred to as *automation bias* (Mosier & Skitka, 1996). Two types of errors are associated with this tendency: (a) omission errors, defined as failures to take necessary action or to respond to system irregularities when not prompted to do so by an automated device; and (b) commission errors, which occur when operators inappropriately follow an automated directive or recommendation without verifying it against other available information or despite contradictions from other sources of information.

Automation bias has been investigated in a series of studies using students and professional pilots as participants (e.g., Mosier, Skitka, Heers, & Burdick, 1997, 1998; Skitka, Mosier, & Burdick, 2000; Skitka, Mosier, Burdick, & Rosenblatt, 2000). In scenarios in which correct information was available to cross check and detect automation anomalies, error rates approximating 55% were documented across both populations. In a student study, participants in a nonautomated condition outperformed those using an automated aid during equivalent failure events (Skitka, Mosier, & Burdick, 1999). The imposition of predecisional accountability, which has been shown to sensitize decision makers to the need to construct compelling justifications for their choices and how they make them, revealed some interesting differences between the student and pilot participants. Students who were instructed that they would be accountable for their accuracy or for their overall performance, and would be required to justify their performance, were more likely to verify automated information against other indicators, and they committed significantly fewer omission and commission errors than those who were told that their performance would not be monitored and could not be tracked (Skitka, Mosier, Burdick, 2000). Within a small professional pilot sample, however, experimentally induced accountability did not produce significant differences among pilots (Mosier et al., 1998). For these pilots, regardless of experimental condition, self-reported perceptions of being evaluated on their performance and strategies in the use of automation and a stronger need to justify their interaction with the automation were associated with fewer automation related errors. These results suggest that pilots possess, to a greater or lesser degree, an internalized sense of accountability for their interaction with automation, and that this sense is associated with vigilant, proactive strategies that stimulate verification of the functioning of automated systems.

Professional pilots were also sensitive to the importance of correctness for critical flight tasks and made fewer errors on events involving altitude and heading errors than frequency discrepancies (Mosier et al., 1998). In addition, pilots who followed automated directives to shut down a supposedly damaged engine, despite the absence of confirmatory indicators, tended to erroneously “remember” the presence of expected cues, a phenomenon we dubbed *phantom memory*.

All of the studies described previously were conducted in a single-person performance configuration; that is, students and pilots performed all required tasks without the aid of a second crewmember. One could reasonably argue that the situation was not representative of real-world commercial flight, which typically involves at least two pilots. What impact would the presence of a second crewmember have on automation bias and automation related errors?

A SECOND TEAMMEMBER: POTENTIAL IMPACT ON BEHAVIOR

When two people are monitoring system events, it would seem to double the chances that they would detect an anomaly, even if it were not detected by an automated decision aid, or recognize an inconsistent or inappropriate automated recommendation. Many cockpit procedures, in fact, are designed on the premise that crewmembers will cross check system indicators as well as each others’ actions. Moreover, research on effective teams has shown that monitoring and giving feedback on each others’ performance are important teamwork behaviors (McIntyre & Salas, 1995).

Research that has specifically addressed team performance, however, has shown mixed effects when comparing teams with individuals. Although the interaction among teammembers has been proposed as a deterrent to errors that may occur when an individual performs a task alone (e.g., Hackman & Morris, 1975), an individual may also alter his or her behavior according to expectations of another teammember’s performance and increase or decrease vigilance depending on the competency of a teammate (Cannon-Bowers, Salas, & Grossman, 1991). Vigilance would be at its highest when confidence in teammates’ capabilities was lowest. The presence of a competent second crewmember as well as a highly reliable automated system, then, might actually discourage rather than enhance vigilance.

Even if vigilant performance is the dominant response, and the presence of two teammembers increases this vigilance, it does not guarantee that errors will be corrected. The use of a team rather than an individual adds a measure of redundancy, but it also introduces more complexity. Research on situation awareness (SA), for example, has indicated that it is a much more complex task to maintain SA in a team context than when individuals are performing alone (Jentsch, Barnett, Bowers, & Salas, 1999).

In addition, many factors, including status, rank, and tenure, may impede the free flow of necessary information and feedback between cockpit crewmembers (e.g., McIntyre & Salas, 1995). Early in aviation history, Torrance (1954) illustrated that aircrews were much more susceptible to error in a laboratory decision-making task if the navigator rather than a higher ranking crewmember held the right answer. In real-world operations, early analyses of air transport accidents revealed many instances in which important information was supplied by one crewmember but disregarded by another, or in which crewmembers who possessed adequate information did not supply it to others (Foushee, 1984).

AUTOMATION AND TEAM DECISION MAKING

According to the multilevel theory of team decision making, several core team-level constructs will influence decision-making accuracy in hierarchical teams such as airline cockpit crews (Hollenbeck et al., 1995). These constructs are *team informity,* the degree to which the team, as a whole, is aware of all of the relevant cues or information; *staff validity,* the degree to which each member of the team can produce valid judgments on the decision object; and *hierarchical sensitivity,* the degree to which the team leader effectively weights teammember judgments in arriving at the team's decision. Thus, the most accurate decisions will be made by teams in which all members have sufficient expertise and are aware of all of the relevant information for a decision and in which the leader can accurately evaluate the relative value of inputs.

The place of automated systems in this model is not as straightforward as it might seem. At first glance, it would appear that automated systems serve only as sources of cues and information, contributing to team informity. The level of autonomy and authority that has been incorporated into modern automated systems, however, may also give them the character of independent agents, or teammembers (Sarter & Woods, 1997), whose expertise may encourage team leaders to weight their input above that of other members, impacting hierarchical sensitivity. Or, as we have suggested in previous automation bias research, teammembers may use automated input as a shortcut in decision making, whether or not its staff validity is high, given the characteristics of a particular situation. The salience of automated cues, as well as their reliability in terms of internal consistency, may contribute toward this tendency. Time constraints, also, especially in critical events, may exacerbate the tendency to rely on a few cues (or even just one).

Even if crews do try to utilize the cues provided by automated systems as part of a thorough search for relevant information, the nature of the systems does not always make this easy to do. In their study of Airbus (A–320) crews, for example, Sarter and Woods (1997) found that many pilots focus solely on the primary flight display (PFD) and that the fragmentation of the feedback provided in areas scattered around the cockpit made it difficult to verify automation behavior. They described the information gathering done by these crews as guided by specific questions that they asked themselves in specific contexts, and their monitoring behavior, as *expectation driven.* Pilots typically looked only as far as necessary to "verify expected changes in the status and behavior of the automation and to deal with uncertainties about the effects of input on aircraft behavior" (Sarter & Woods, 1997, p. 564). Important relevant information, or evidence that an error has occurred, could easily be missed even with two pilots, because both crewmembers performed the same limited information seeking and verifying behaviors.

It should be noted that these kinds of errors in the use of information are not limited to automated aircraft. The Aviation Safety Reporting System and National Transportation Safety Board (1982, 1993)[1] reports contain many accidents and incidents in which faulty use of cues resulted in omission or commission errors. The particular characteristics of automated cues described previously, however, make them likely culprits in a less-than-complete information search and situation assessment.

AUTOMATION BIAS AND TEAMS

Given these sorts of mixed results with respect to teams, automation, and decision making, it is not at all certain what impact the presence of a second crewmember will have on automation bias and associated errors. As mentioned, an earlier study utilizing professional pilots in a single-person configuration resulted in an omission error rate of approximately 55% and demonstrated pilot sensitivity to the critical nature of flight tasks (e.g., fewer errors were committed during altitude events than during frequency events). During the one event that involved a discrepancy between actions suggested by automated system information and traditional indicators (an "ENGINE FIRE" message on the Engine Indicating and Crew Alerting System [EICAS] that was contradicted by completely normal engine parameters and the absence of any other fire indication), all of the participants acted in accordance with automated directives (Mosier et al., 1998). Will these behavioral tendencies and error rates persist in two-person crews?

One potential method to increase the chances of automation errors being detected by crews is to incorporate verification procedures into systems training. This would involve the instruction of pilots not only on the location of status and

---

[1] We thank an anonymous reviewer for pointing out these examples.

function information, but would also require them to physically check these indicators at given points. Airlines have already introduced this procedure for some functions, taking advantage of the natural redundancy inherent in two-person crews. For example, when one pilot enters a new altitude into the altitude indicator, the other pilot is usually required to verify the correctness of the entry. The assumption is that having pilots crosscheck each other's interaction with systems will increase the chances that errors will be caught. With respect to automation, incorporating procedures that involve checking automated systems and functions against other indicators should enhance the probability that anomalies will be detected and corrected. Verification requirements would also be made more explicit and imperative if a visual prompt to verify automated functioning appeared at appropriate points, such as after an altitude or heading change.

It is also possible that making pilots aware of the tendency toward automation bias and potential errors ahead of time can be effective in forestalling these errors, an intervention that could be referred to as *debiasing* (e.g., Fischoff, 1982). For example, before their participants performed an experimental decision-making task, Ross, Lepper, and Hubbard (1975) gave them *process prebriefings,* informing them of a decision-making phenomenon, biased assimilation of information, that was likely to lead them astray. Forewarned of this bias, individuals did not succumb to it. Informing pilots of the existence of automation bias and training them on how to avoid it (e.g., by verifying automated information against other indicators) might be sufficient to mitigate the bias and decrease associated errors.

A recent study utilizing student teams compared the performance of one- and two-person crews on a low-fidelity, flight analog task (Skitka, Mosier, Burdick, & Rosenblatt, 2000). The study explored the extent to which automation bias remains a pervasive problem or is mitigated (a) in the context of a two-person crew, (b) when participants were given explicit training about automation bias as a potential problem, (c) when participants were explicitly trained to verify automated directives, and (d) when participants were prompted visually to verify automated directives. Results revealed that, for students, having a second crewmember did not significantly impact the number of automation-related errors committed. In fact, none of these independent variables had an impact on the number of omission errors made. Training participants on the phenomenon of automation bias, however, produced a significant reduction in commission errors.

The study reported here was designed to investigate whether the error rates found with single pilots and with teams of students will hold in the context of an aircraft cockpit, with a professional aircrew. In addition to exploring the impact of number of professional flight crewmembers on automation bias and resultant errors, the study reported here also investigated the efficacy of interventions involving verification training, explicit automation bias training, and display enhancements.

METHOD

Participants

Participants included 48 commercial glass cockpit pilots (24 captains, 21 first officers, and 3 second officers) from three major U.S. carriers. Second officers were either former captains who had reached the age of 60 years and moved back to the second officer position ($n = 2$), or they were pilots who had recently moved from a first officer to second officer position in a larger aircraft ($n = 1$). All pilots had flown a glass cockpit aircraft within the previous 6 months and, for this study, flew in their current or most recent flying position. Pilots ranged in age from 25 to 61 years ($M = 48.25$, $SD = 9.74$) and had total flight experience ranging from 2,300 to 26,000 hr ($M = 13{,}229$, $SD = 6{,}577$). The average glass cockpit experience of participants was 2,407 hr ($SD = 1{,}668$). Twenty captains and 20 first officers, from the same airline, were paired together to fly as a crew. Due to a limited participant pool, it was not possible to have an equal number of single-person crews. However, 8 participants flew the experimental scenarios in a single-person configuration to establish a baseline of performance.

Equipment and Materials

Crews flew in the mini-Advanced Concepts Flight Simulator (ACFS). This part-task simulation facility employs four Silicon Graphics color monitors to present primary and secondary flight displays, navigation and communication information, and electronic checklists, as well as EICAS and Flight Management System (FMS) instruments (see Figure 1). The displays are similar to what would be found in a B747–400. Pilots interacted with the controls and displays of the aircraft through a touchscreen overlaying the instruments. The mini-ACFS is set up for two-person crews but can also be configured to be flown by one person. Pilots received clearance and information from an air traffic controller, who was stationed in the area outside the simulator booth.

Communications with Air Traffic Control (ATC) were conducted via datalink, or digitized communication. Clearances that involved changes in flight parameters or communications frequencies could be automatically loaded into the appropriate control unit by pressing the LOAD button (see Mosier et al., 1998). Autoloaded clearances appeared on the relevant display for confirmation. Communications with ATC could also be conducted orally if desired. Electronic checklists (see Mosier, Palmer, & Degani, 1992) for normal and abnormal procedures were accessed through a menu display and were displayed on a monitor in front of the pilot. In the event of a system abnormality or failure, the relevant checklist could be accessed directly by touching the warning message on the EICAS screen.

FIGURE 1 Mini-Advanced Concepts Flight Simulator setup and display.

Design and Procedures

Participants flew under conditions that varied as a function of (a) crew size, that is, whether they worked alone or with another person (two levels); (b) one of three levels of training (systems-only training that included instructions on the aircraft systems and operations and showed pilots where to verify automated actions and directives, training that emphasized that they must verify automated actions and directives, or training that incorporated information about automation bias, errors people tend to make in automated contexts, and how these errors can be avoided); and (c) whether or not participants received a prompt to verify automated functioning each time a clearance was autoloaded and each time the EICAS displayed a warning message. The prompts appeared on the FMS or EICAS display screen, and read “VERIFY.” Dependent variables were the number of omission and commission errors participants made across conditions.

The study was a mixed design. Type of training (automation bias–verification) was a between-crews factor, and display (presence–absence) was a within-crews factor. Each crew received basic systems training on the mini-ACFS and served as its own control. Crews flew several practice approaches. Following this, each crew flew three approaches into San Francisco (SFO), into which several automation events had been programmed. The first experimental leg, flown after systems training only, served as the control leg. Prior to Legs 2 and 3, crews were given additional training as determined by assignment to experimental condition. Within each training group, half of the crews received a textual display prompt during Leg 2 and the other half during Leg 3.

Participants were instructed to communicate with ATC through datalink, although voice communications would always be available to them if desired or needed. Most clearances from ATC (e.g., route changes or changes in altitude, speed, heading, or frequency) could be autoloaded into the appropriate flight system, and correct loading could be verified by checking the mode control panel (MCP), PFD, or navigation system as appropriate. During training and practice, pilots practiced autoloading each type of clearance and verifying proper loading and execution on the appropriate display.

Automation events. Six automation failures during the three data legs offered the possibility for pilots to make omission errors if they did not verify proper automation functioning. These were similar to the events in Mosier et al. (1998) and involved misloads of frequency, altitude, arrival, and runway and navigation frequency clearances, as well as the incorrect execution of a heading clearance. Verification information for each flight-related event was available on the appropriate display as it would be in the aircraft.

One event offered the opportunity for a commission error and concerned the false automated warning of an engine fire during a go-around or missed approach at SFO (Mosier et al., 1992; Mosier et al., 1998). As they descended to approximately 250 ft, crews received a verbal ATC command to execute a missed approach and go-around due to traffic on the runway. On climb out, an automated ENGINE FIRE message appeared on the EICAS, but this was contradicted by normal engine parameters and the absence of any other indicators. Touching the EICAS message brought up the electronic checklist for Engine 1 fire and shutdown. Crews had to determine whether there really was a fire and decide whether or not to shut down the supposedly affected engine while executing the missed approach.

This event was a direct replication of the false engine fire event in Mosier et al. (1998). As before, pilots were reminded during training that an engine fire would be indicated by a number of cues, including a loud aural warning, a red master warning light, an illuminated red fire handle on the display monitor, possible deterioration of engine indicators, and an electronically generated EICAS message. All of these cues were demonstrated and their locations pointed out as part of the training process. During the experimental event, only the EICAS message appeared, which was linked to the electronic checklist, but the message was not verified by any of the other expected cues.

Debrief. Participants were asked to complete a paper-and-pencil questionnaire to obtain demographic information, and they also answered several questions regarding the engine fire event. Pilots were asked to check off the cues that were present during the event and, in a separate item, rate how confident they would be that an engine fire was present, given only one of the five possible cues listed previously, on a 7-point scale ranging from 1 (*possibly a fire*) through 7 (*definitely a fire*). They were also asked to respond on a 7-point scale, ranging from 1 (*disagree*) through 7 (*agree*) to the following statement:

> While performing a go-around, you receive an EICAS message indicating an engine fire. All other indicators are normal. In general, it would be *safer* to *shut down* the engine in case it really is on fire (leaving you with only one engine to perform the go-around), rather than to retard the throttle of the indicated engine and complete the go-around procedure with the engine running.

RESULTS

Overall omission error rates were slightly, but not significantly, better for crews than for solo pilots (43% vs. 52%), $F(1, 26) = .89$, *ns*. By event, error rates for two-person crews were: frequency, 60%; altitude, 10%; heading, 65%; arrival, 55%, runway, 16%; and navigation frequency, 47%. For single-person crews, the error rates were: frequency, 75%; altitude, 25%; heading, 43%; arrival, 88%; runway, 25%; and navigation frequency, 50%. Data for the solo pilots were comparable to the baseline error rate found in the previous single-pilot study (55%; Mosier et al., 1998). No further statistical analyses were performed on solo-pilot data.

No significant effects on errors were found for training type or display presence, $F(1, 16) = .14$ and .09, *ns*, respectively, for crews. Omission error performance on the experimental legs was best predicted by performance on the control leg, $r(18) = .47$, $p < .05$. A significant effect was found for event, $F(5, 95) = 4.04$, $p < .01$, with altitude and runway being corrected more often than frequency, arrival, heading, or nav frequency. Unexpectedly, in 21% of the total of frequency, altitude, arrival waypoint, and heading events, automation discrepancies were caught and verbally acknowledged by the crews, but no corrective action was taken.

With respect to the single opportunity for a commission error, all but two of the two-person crews (and all of the solo crewmembers)[2] responded to the false engine fire EICAS event by shutting down the supposedly affected engine on go-around. On the debriefing questionnaire, pilots responded that the presence of an EICAS message by itself would not be sufficient to ensure that an engine fire was definitely present ($M = 4.2$, $SD = 2.2$). Pilots did not agree with the statement that it would be safer, in the event of only an EICAS message while performing a go-around, to shut down the supposedly affected engine rather than to retard the throttle and leave it running ($M = 3.3$, $SD = 2.15$). In 43% of the solo pilots and 74% of the two-person crews, one or both members erroneously remembered at least one other diagnostic cue as being present during the event. It is interesting to

[2]Due to technical difficulties, one solo pilot did not experience this event.

note that none of the four crewmembers (two crews) who left the engine running at idle recalled any extra indicators.

DISCUSSION

According to the results described previously, the presence of a second crewmember is not enough to eradicate, or even significantly reduce, automation bias and associated errors. It appears that the positive impact of having a second crewmember is not, by itself, sufficient to compensate for the automation-bias-inducing characteristics of glass cockpits. Consistent with past findings (e.g., Mosier et al., 1998), aircrew behavior was not impacted by external manipulations. Training crews on automation bias or to verify correct automated functioning had no effect on automation-related omission errors, and neither did display prompts that reminded crews to verify.

It is interesting to note that the best predictors of crew omission errors were performance before any experimental manipulation (i.e., on the control leg) and criticality of the error event. Crews that caught errors before any extra training or display manipulations were more likely to catch them on subsequent legs, and errors in the most critical aspects of flight (e.g., altitude capture) were more likely to be caught than those in less critical functions (e.g., communications). Given the extensive training that crews receive on prioritization of flight tasks (e.g., aviate, navigate, and communicate) and the penalties that then can incur for being off altitude and flight path, it is not surprising that criticality emerged as a variable predictive of automation-related errors. This performance characteristic had also been found in the previous solo-pilot study (Mosier et al., 1998).

The fact that performance during the control leg was more predictive of later performance than any external manipulation suggests that the nature of pilot interaction with automation is, in part, a product of individual difference characteristics. Individual differences among pilots in attitudes toward automation and in automation use have been found in previous research, and they have been shown to be related to performance with automated systems (e.g., Helmreich, 1984; McClumpha & James, 1994; Riley, 1996). These differences are associated with the interaction between personal factors, such as trust in automated systems and self-confidence, and more objective characteristics, such as automation reliability and consistency, workload, and cognitive overhead associated with automation use (Parasuraman & Riley, 1997). In addition, our earlier investigation of the impact of decisional accountability on automation bias revealed that an internalized sense of accountability for performance was significantly related to automation-related performance and was associated with fewer omission errors (Mosier et al., 1998).

It is possible that some people may be less susceptible to automation bias because they are more conscientious with respect to double checking automated functions. In support of this notion is the finding that errors were often caught by one crewmember more frequently than the other, regardless of whether he was flying the aircraft or acting as pilot not flying. For 14 out of the 20 crews, one of the pilots was responsible for all, or all but one, of the automation catches for his crew. The relation between individual characteristics and interaction with automation is, as Parasuraman and Riley (1997) suggested, a complex one; further research is needed to clarify this relation.

One of the most puzzling factors in this study has to do with the 21% of the total of frequency, altitude, arrival waypoint, and heading events in which crews clearly saw and acknowledged a discrepancy between what they expected and what the automation was doing but did not take corrective action. This could be in part due to a *simulator effect,* that is, crews may have been less aggressive than usual in correcting errors because they were in a simulator and no real risk was present. It could also indicate that crews are wary of correcting automation errors, perhaps because errors occur so infrequently that they may question their own interpretation of what is happening (e.g., Did the frequency load incorrectly? Or, probably more likely, did we misread the frequency?). Because they are used to multiple redundant systems and systems checks, crews may also prefer to wait until the discrepancy becomes a critical factor with the hope that errors will correct themselves.

The false engine fire event offered the only opportunity in this study for a commission error, and almost all of the crews responded by following the erroneous warning, despite their reported beliefs that an EICAS message by itself would not be sufficient to ensure that an engine fire was actually present, and their failure to agree with the notion that the safest course of action in this event would be to shut down the supposedly affected engine. It should be noted that crew actions during the false engine fire event may have been directly related to particular airline policies. At least one of the airlines whose pilots were involved in the study trains crews to assume a fire is present whether or not all the indicators are consistent. Crews, then, may be following company procedure that is in conflict with what they think is the safer course of action, and organizations may be fostering tendencies toward automation bias through their policies and training.

One of the most intriguing facets of the false engine fire event, consistent with previous findings (Mosier et al., 1998), was the persistence of the phantom-memory phenomenon in crews that did shut down the engine. Crews may not be detecting automation anomalies because they believe that they are seeing information that verifies automated cues. In support of this, it should be noted that both crews who left the engine running and available for use knew that no other indicators were present. It is interesting to note that most crews did verify in this event—they very carefully verified that they were shutting down the correct engine rather than the wrong one.

The results of this study suggest that crews are failing to take into account all relevant information—at the core of the problem is decreased situational awareness and vigilance. This implies that if designers of systems expect operators to attend to all information available to them, they need to design sources of information to be equally salient and available. Design that does not account for the human tendency to take shortcuts can certainly not be labeled *human centered*.

The knowledge that the roots of many automation-related errors can be traced to cognitive shortcuts has major implications for training interventions. Successful interventions should be focused on helping to ensure that pilots notice and take into account a broader array of information (i.e., increased situational awareness) before coming to premature closure on a decision. Given that automation bias—both errors of omission and commission—have, at their root, vigilance and SA issues, there is considerable hope that through design changes and training interventions their effects can be minimized. Assuming we can develop strategies to encourage pilots as information processors to be more vigilant or to present nonautomated information in as salient a manner as automated information, we can be reasonably confident that operators can and will make rational and safe decisions.

ACKNOWLEDGMENT

This work was supported by Grants NCC2–798, NCC2–837, and NAS2–832 from the NASA Ames Research Center, Moffett Field, CA.

REFERENCES

Cannon-Bowers, J. A., Salas, E., & Grossman, J. D. (1991, June). *Improving tactical decision making under stress: Research directions and applied implications.* Paper presented at the International Applied Military Psychology Symposium, Stockholm.

Fischoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge, England: Cambridge University Press.

Fiske, S. T., & Taylor, S. E. (1994). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Foushee, H. C. (1984). Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. *American Psychologist, 39,* 885–893.

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 45–99). New York: Academic.

Helmreich, R. L. (1984). Cockpit management attitudes. *Human Factors, 26,* 583–589.

Hollenbeck, J. R., Ilgen, D. R., Sego, D. J., Hedlund, J., Major, D. A., & Phillips, J. (1995). Multilevel theory of team decision making: Decision performance in teams incorporating distributed expertise. *Journal of Applied Psychology, 80,* 292–316.

Jentsch, F., Barnett, J., Bowers, C. A., & Salas, E. (1999). Who is flying this plane anyway? What mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors, 41,* 1–14.

McClumpha, A., & James, M. (1994). Understanding automated aircraft. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Recent research and trends* (pp. 314–319). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

McIntyre, R. M., & Salas, E. (1995). Measuring and managing for team performance: Lessons from complex environments. In R. A. Guzzo, E. Salas, & Associates (Eds.), *Team effectiveness and decision making in organizations* (pp. 9–46). San Francisco: Jossey-Bass.

Mosier, K. L., Palmer, E. A., & Degani, A. (1992). Electronic checklists: Implications for decision making. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 7–11). Santa Monica, CA: Human Factors Society.

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 Associates, Inc.

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. (1997). Patterns in the use of cockpit automation. In M. Mouloua & J. Koonce (Eds.), *Human–automation interaction: Research and practice* (pp. 167–173). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. D. (1998). Automation bias: Decision making and performance in high-tech cockpits. *International Journal of Aviation Psychology, 8,* 47–63.

National Transportation Safety Board. (1982). *Air Florida, Inc., Boeing 737–222, N62AF, collision with 14th Street Bridge, near Washington National Airport, Washington, DC, January 13, 1892* (Rep. No. NTSB–AAR–82–08). Washington, DC: Author.

National Transportation Safety Board. (1993). *Aborted takeoff shortly after liftoff, Trans World Airlines Flight 843 Lockheed L–1011, N11002, John F. Kennedy International Airport, Jamaica, New York, July 30, 1992* (Rep. No. NTSB–AAR–93–04). Washington, DC: Author.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39,* 230–253.

Riley, V. (1996). Operator reliance on automation: Theory and data. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 19–36). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32,* 880–892.

Sarter, N. B., & Woods, D. D. (1997). Team play with a powerful and independent agent: Operational experiences and automation surprises on the Airbus A–320. *Human Factors, 39,* 553–569.

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision making? *International Journal of Human–Computer Studies, 51,* 991–1006.

Skitka, L. J., Mosier, K. L., & Burdick, M. D. (2000). Accountability and automation bias. *International Journal of Human–Computer Studies, 52,* 701–717.

Skitka, L. J., Mosier, K. L., Burdick, M. D., & Rosenblatt, B. (2000). Automation bias and errors: Are crews better than individuals? *International Journal of Aviation Psychology, 10,* 85–97.

Torrance, E. P. (1954). Some consequences of power differences on decision making in permanent and temporary three-man groups. In A. P. Hare (Ed.), *Handbook of small group research.* New York: Free Press.

Manuscript first received September 1999