Understanding the Human-Computer Team[1]

Mary T. Dzindolet
Cameron University
2800 Gore Blvd.
Lawton, OK 73505
UNITED STATES

Linda G. Pierce
Army Research Laboratory
Fort Sill
Oklahoma
UNITED STATES

Hall P. Beck
Appalachian State University
Dept of Psychology
Boon, NC 28608
UNITED STATES

Information technology is changing the nature of work. Access to more relevant, accurate, and timely information than was previously possible has drastically increased the use of automation across society, including the military (cf. Cesar, 1995). The requirement for automation is highlighted by the increasing complexity of the military mission, which is expanding to include the full spectrum of conflict, from humanitarian assistance and peace-keeping to small-scale contingencies and major wars. To enhance situational awareness and a myriad of other tasks, the military decision-maker is often provided with automated aids. The underlying assumption in providing these automated aids to military personnel is that the human-computer "team" will be more productive than either the human or the automated aid would be working alone. Some researchers have found support for this underlying assumption (Corcoran, Dennett, & Carpenter, 1972; Dalal & Kasper, 1994; Parasuraman, 1987; Thackray & Touchstone, 1989a), others have found human operators often overly rely on (misuse) or underutilize (disuse) automated decision systems (Parasuraman & Riley, 1997).

One strategy used to optimize human-computer performance has been to call on system designers to create automated aids that are increasingly more reliable. Increasing the automated aid's reliability is assumed to lead to increased human-computer "team" performance. As with human teams, however, increasing the reliability of one team member's performance will not necessarily affect the team's performance. Sorkin and Woods' (1985) signal detection analysis revealed that optimizing an automated aid's performance would not always optimize the human-computer team's performance on a monitoring task. Specifically, they found that using a response criterion that yielded the best performance for the automated aid (i.e., highest detection rates and fewest false alarms) did not yield the best performance for the human-automated team. Although "synergy" can be found in human-computer teams (e.g., Dalal & Kasper, 1994), it is not likely to be gained by optimizing human-alone or computer-alone performances. The interaction between the automated aid and human operator must be considered.

By understanding the processes that human operators use when deciding to rely on their decisions or on an automated aid's decisions, one may be able to better design systems and train military decision-makers on the appropriate use of automated aids. This is especially important in the current military environment in which

[1] This paper is based on a technical report published by the Army Research Laboratory in 2001 by Dzindolet, Beck, Pierce, & Dawe.

Paper presented at the RTO SCI Symposium on "Critical Design Issues for the Human-Machine Interface", held in Prague, Czech Republic, 19-21 May 2003, and published in RTO-MP-112.

automation is used to solve problems. For example, the high fratricide rates experienced during Operation Desert Storm led the U.S. Army to two technological solutions: the Battlefield Combat Identification System (BCIS) and the Combat Identification for the Dismounted Soldier (CIDDS; Doton, 1996). Both of these systems were designed as decision aids for the identification of friendly troops, the BCIS for armor gunners and the CIDDS for the individual soldier. These systems provide the soldier the ability to "interrogate" a potential target by sending a microwave or laser signal that, if returned, identifies the target as a "friend." Unanswered signals produce an "unknown" response. When the hardware on all friendly troops is functioning properly, the combat identification systems can identify a "friend" at a confidence level above 99%. Presumably this rate of reliability is greater than that of most soldiers, thus, when the hardware is functioning properly, soldiers should rely on the decisions made by the combat identification systems. However, the hardware may not always function properly; it may be lost or damaged in battle. In addition, the likelihood of receiving an "unknown" response to a friendly target may change during battle in ways that are difficult to predict. For example, the probability that an "unknown" response identifies an enemy may be very high until a large number of allied forces without the transponders (the hardware that generates the "friendly" signals) enter the area. Under these conditions, soldiers should ignore the combat identification systems. Unfortunately, several laboratory studies indicated that sub-optimal use of the combat identification systems had the potential of increasing, rather than decreasing, fratricide (Dzindolet, Pierce, Beck, Dawe, & Anderson, 2000). An understanding of human-automation decision-making is necessary. The purpose of this paper is to present a general framework we believe will be useful in guiding research to understand human operators' task allocation decisions (see Figure 8). We will discuss the results from many studies that support the framework and provide suggestions for the design of future automated decision aids and training procedures for operators relying on such aids.

DISUSE

A rational decision-maker will rely on an automated aid when doing so will maximize gains and/or minimize losses. Failure to rely on an aid in this situation constitutes disuse (Parasuraman & Riley, 1997). Disuse is defined as "underutilization of automation" (p. 233). Anecdotal evidence supports disuse; Parasuraman and Riley (1997) described many real-world incidences in which disastrous results occurred due to people ignoring automated warning signals. Further, laboratory experiments have found disuse in paradigms in which the aid's decisions are provided only after the human operators have indicated their decision (Dzindolet, Pierce, Beck, & Dawe, 1999; Moes, Knox, Pierce, & Beck, 1999). For example, Moes et al. (1999) found the majority of college students (67%) chose to ignore the decisions of an automated aid even after being provided with feedback that their automated aid made half as many errors as they did during 200 prior trials. In addition, Riley (1996, Study 1) found students favored manual operation over automated control, even when doing so was clearly not an optimal strategy.

MISUSE

When ignoring an automated aid will maximize gains and/or minimize losses, the rational decision-maker will ignore the aid and rely on his or her decisions. Relying on the automated aid in this circumstance would constitute misuse (Parasuraman & Riley, 1997). Misuse is defined, "as overreliance on automation" (p. 233). Parasuraman, Molloy and Singh (1993) and Singh, Molloy, and Parasuraman (1997) found misuse among operators performing monitoring functions. They labeled this behavior complacency, "a psychological state characterized by a low index of suspicion" (Wiener, 1981, p. 117). Misuse has also been found with automated decision aids (Dzindolet, Pierce et al., 2000; Layton, Smith, & McCoy, 1994; Mosier & Skitka, 1996).

FRAMEWORK OF AUTOMATION USE

What are the processes leading to appropriate use, misuse, and disuse of an automated aid? Mosier and Skitka (1996) outlined three possible reasons why people inappropriately use automation: (1) cognitive miser hypothesis, (2) authority hypothesis, and (3) diffusion of responsibility. These three hypotheses parallel three processes: cognitive, social, and motivational, which have been implicated as the causes of productivity loss found in groups (e.g., Mullen, Johnson, & Salas, 1991). Since many researchers have considered a human-computer "team" to be a group in which one member happens not to be human (e.g., Bowers, Oser, Salas, & Cannon-Bowers, 1996), we hypothesize that these same three processes can lead to sub-optimal performance in human-computer teams. The Framework of Automation Use (Dzindolet, Beck, Pierce, & Dawe, 2001) hypothesizes that cognitive, social, and motivational processes of the human operator combine to predict automation use. A discussion of each process as it relates to inappropriate use is presented in the next three sections.

Cognitive Processes: Cognitive Miser Hypothesis / Automation Bias

Mosier and Skitka (1996) hypothesized that people may overly rely on automated systems when making decisions due to faulty cognitive processing. In addition to the large body of literature examining errors due to flawed cognitive processing in *individual* decision-making (Tversky & Kahneman, 1973), the social cognition literature is replete with examples of less-than-ideal cognitive processing while working in teams or groups. Errors and biases have been identified in various domains of social psychology (e.g., illusory correlation and in-group differentiation/out-group homogeneity in stereotype formation (Hamilton & Sherman, 1989; Mullen & Hu, 1989), the halo effect and the negativity bias in impression formation (Skowronski & Carlston, 1989), the confirmation bias and self-handicapping in attribution (Arkin & Baumgardner, 1985; Leyens & Yzerbyt, 1992). Rather than logically processing relevant pieces of information, people often adopt effort-saving strategies called heuristics. Mosier and Skitka (1996) coined the term "automation bias" to refer to, "the tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing" (p. 205).

The fact that the automated system provides a decision may lead the decision-maker to rely on this information in a heuristic manner (see Figure 1). Rather than going through the cognitive effort of gathering and processing information, the information supplied by the automated systems is used (Mosier & Skitka, 1996). Conceivably, this may occur in various degrees. In its most extreme form, the decision reached by the automated aid is immediately adopted. In a less extreme form, the decision reached by the aid may be given an inappropriately large role in the human's decision-making process. For example, Layton, Smith, and McCoy (1994) found many pilots provided with an automated aid's poor en-route flight plan did not explore other solutions (e.g., they did not generate actual flight plans on screen) as much as pilots not provided with the automated aid's decision.

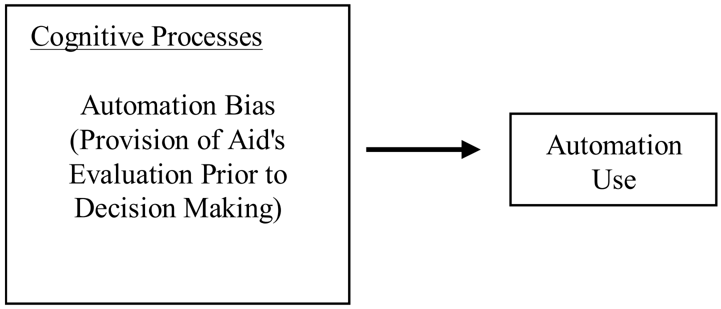

Figure 1: Automation Bias.

The automation bias will lead participants to rely on the automated aid. Oftentimes, this strategy will be appropriate. However, under certain conditions, this reliance may be inappropriate leading to misuse.

Although the automation bias plays a role in predicting automation use, it cannot account for the full range of potential findings. Specifically, automation bias would not be able to account for the anecdotal evidence for disuse when an automated aid's evaluation is provided to humans prior to their reporting a final decision. The information received from the automated aid is predicted to influence the human's response. Ignoring an automated aid's decision would not be expected.

Indirect evidence to support the existence of automation bias can be found by examining task allocation decisions across two paradigms. In one paradigm, the automation bias is allowed to flourish; in the other, it is prevented from occurring. In both paradigms, participants view about 200 slides displaying pictures of Fort Sill terrain on a computer screen (see Figure 2 for a sample slide). About half of the slides contain one soldier ("target") in various levels of camouflage; the remaining slides are of terrain only.

Figure 2: Sample Slide Containing Soldier.

Sometimes the target is easy to detect (as in the above slide); other times it is more difficult to find. Each slide is presented on the computer screen for about ¾ of a second. Participants do not need to make their absent-present decision alone; they are provided with the decision of an automated decision aid. Specifically, participants are told that a computer routine had been written to assist them in performing their task. They are told that the routine performs a rapid scan of the photograph looking for contrasts that suggest the presence of a human being. If the contrast detector routine determines the soldier is probably present, the word "PRESENT" and a red circle appears. If the contrast detector routine determines the soldier is probably absent, the word "ABSENT" and a green circle appears.

The next screen asks the students to indicate whether or not they believe the soldier was in the slide. They are given as much time as they need to make their decision. Participants are informed that there are two possible errors that they can make. One error is made when they indicate that the soldier is present when, in fact, he is not. The other error is made when they indicate that the soldier is not present when, in fact, he is. Participants are told that both errors are equally serious and that they should attempt to avoid them.

Finally, the participants indicate the extent to which they are certain their decision is correct. A five-point Likert scale ranging from "highly confident" to "not at all confident" is provided. Because participants view the automated decision aid's decision before making their own decision, the automation bias is allowed to flourish.

To analyze reliance in this paradigm, participants' overall error rate is determined [p(error)]. To examine error type, the participants' error rate is determined for the trials in which the aid gave the correct decision [p(error | aid correct)] and for the trials in which the aid gave an incorrect decision [p(error | aid error)]. Misuse, or overreliance on the contrast detector, exists when p(error | aid error) > p(error | aid correct). Disuse, or underutilization of the contrast detector, exists when p(error | aid error) = p(error | aid correct).

The reliability of the contrast detector has been varied in different studies. For example, Dzindolet, Pierce, Beck, Dawe, and Anderson (2001) varied the reliability of the automated decision aid to be 60%, 75%, or 90% accurate. In a study reported in Dzindolet, Pierce, Peterson, Purcell, and Beck (2002), the contrast detector was 100% accurate when it reported the target was absent, but only 67% accurate when it reported the target was present.

Repeatedly, misuse is found. Dzindolet et al. (2001) found that regardless of the reliability of the automated aid, participants were more likely to err by relying on their decision aids than by ignoring them, p(error | aid error) = .27; p(error | aid correct) = .13. Similarly, participants in the Dzindolet, Pierce, Peterson et al. (2002) study were unable to adjust their reliance strategy to the aid's reliability. Regardless of the aid's decision, misuse was found, p(error | aid error) = .37; p(error | aid correct) = .15. In fact, participants not provided with the automated decision aid outperformed those provided with the aid.

Therefore, when provided with the aid's decisions first, thereby allowing the automation bias to occur, participants were more likely to misuse than disuse their automated aids. Given the propensity for participants to rely on the decision aid – even an unreliable aid, we were surprised by the results of one study in which the automated decision aid was always correct (Batka, Beadles, & Dzindolet, 2003). Misuse was not possible in this study; reliance on the aid was the best strategy. Yet, 87% of the time, participants chose to disagree with the automated aid.

Changing this paradigm slightly, we were able to create a paradigm in which the automation bias is eliminated. Participants are *not* provided with the decisions reached by the contrast detector until *after* they have indicated their decision and their level of confidence in their decision. Without the automation bias, will participants still be more likely to misuse than disuse their automated aids?

Participants in these studies view the slide of Fort Sill terrain for about 3/4 of a second, indicate their decision, rate their confidence, and, only then, are provided with the contrast detector's decision. After completing 200 slides, students are told they can earn $.50 in coupons to be used at the university cafeteria (or in some studies, extra credit in a course) for every correct decision made on ten randomly chosen trials. Participants have to *choose* whether the performance will be based on ***their decisions*** or on ***the decisions of their aid***. After making their choice, students are asked to justify their choice in writing.

Rather than misusing the contrast detector, participants in these studies disuse the automated aid (Dzindolet, Peterson, Pomranky, Pierce, & Beck, in press; Dzindolet, Pierce, Beck, & Dawe, 2002). Even among participants provided with feedback that their aid's performance was far superior to their own, the majority choose to rely on their own decisions rather than on the decisions of the automated aid! For example, Beck, Dzindolet, Pierce, Poole, and McDowell (2000) found 83% of participants in their study chose self-reliance rather than relying a the decisions of a superior automated decision aid.

Therefore, when the automation bias could play a role in the decision to rely on automation, misuse occurred more than disuse. However, when the automation bias was eliminated by providing the automated aid's decisions only *after* participants recorded their decision, disuse, not misuse, was found on a subsequent task allocation decision.

Although we hypothesize that the automation bias plays a role in automation use and misuse, we do not believe it is the only important variable. At the very least, processes predicting the bias toward disuse found in these studies need to be identified. Analyses of the justifications of the task allocation decisions provided by participants indicated other variables might play a role (Dzindolet, Pierce, Beck, & Dawe, 2002; Dzindolet et al., in press). For example, nearly one-quarter (23%) of the participants in Dzindolet, Pierce, Beck and Dawe's (2002) study justified their disuse by stating they did not trust the automated aid as much as they trusted themselves. Similarly, many of the students who chose self-reliance in the Dzindolet et al. (in press) Study 2 justified their decision by stating that they did not trust the automated decision aid.

Social Processes: Authority Hypothesis/Trust in Automation

A second explanation of the inappropriate use of automated aids has to do with the role of the computer as the expert. According to Mosier and Skitka's (1996) authority hypothesis, people rely on the automated system's decision because they believe it to be more reliable, and thus place greater trust in it. Many other researchers have indicated trust is one important variable in predicting automation use (Cohen, Parasuraman, & Freeman, 1998; Lee & Moray, 1992; 1994; Moray, Inagaki, & Itoh, 2000; Muir, 1987; 1994; Singh, Molloy, & Parasuraman, 1993; Tan & Lewandowsky, 1996). Overly trusting an automated aid will lead human operators to misuse; lack of trust in a superior aid will lead to disuse.

Parasuraman and Riley (1997) described many real-world incidences in which disastrous results occurred due to people ignoring automated warning signals they saw as untrustworthy. Ignoring automated alarm systems that have previously signaled in error has been dubbed the cry wolf effect (Bliss, 1997; Bliss, Dunn, & Fuller, 1995). Publicizing the alarm's trustworthiness was one strategy that proved effective in reducing the cry wolf phenomenon (Bliss et al., 1995).

Muir (1987; 1994), one of the first researchers to focus on automation trust, relied on the literature of human trust to understand human operator's trust of automation. She hypothesized that automation that is predictable, dependable, and inspires faith that it will behave as expected in unknown situations will be seen as more trustworthy. Trust is gained in the areas of persistence, technical competence, and fiduciary responsibility.

To test some of Muir's hypotheses, Lee and Moray (1992; 1994) examined the effect of trust in automation on task allocation decisions. Participants controlled a simulated orange juice pasteurization plant for two hours each day for three days. The simulation included three subsystems, each of which could be operated manually or automatically. Participants could allocate tasks any way they wished and could change their allocations easily. As part of the experiment, whether controlled automatically or manually, one of the subsystems failed periodically. Lee and Moray (1992; 1994) were especially interested in task allocation changes after these failure events, since Muir predicted that after failure events, trust would decline rapidly and slowly increase as the system performed without making errors. At the end of each session, participants completed a questionnaire concerning their trust in the automation and self-confidence in performing the tasks manually.

Results indicated strong individual differences in automation use. Some participants were prone to use manual control; others were prone to use automation. Singh et al. (1993) created a scale to determine individual differences in the propensity to misuse automated aids that may have been able to predict which individuals would be prone to rely on automation and which would be prone to rely on manual operation.

Inconsistent with Muir's hypotheses, the Lee and Moray (1992; 1994) participants rarely changed their allocation task decisions. Once the human operator assigned a subsystem to automated or manual control, he or she was unlikely to change the decision during that session – even after failure events.

To predict trust in automation dynamically, Cohen, Parasuraman and Freeman (1998) extended Muir's work and created an especially promising model, Argument-based Probabilistic Trust (APT). This model defines trust as the perceived probability of the system's reliability given certain situations. Rather than trust being correlated with the predicted reliability of the system (e.g., Muir's (1994) theory), trust *is* the predicted reliability of the system.

Whether predicted reliability is related to trust (e.g., Muir, 1994) or is trust (e.g., APT model), we do not believe predicted reliability will directly predict automation use. The predicted reliability of an automated aid is meaningful only in the context of some standard of performance. What standard did the Lee and Moray (1992; 1994) subjects use? Automation use was found to be related to subjects' estimates of the ability of the automated aid *relative to estimates of their own ability.* Thus, we hypothesize that automation use is determined from the outcome of a comparison process between the perceived reliability of the automated aid (trust in aid) and the perceived reliability of manual control (trust in self). We call the outcome of the decision process the perceived utility of the automated aid (see Figure 3). If one perceives the ability of the aid to be greater than one's own, perceived utility of the aid will be high. If one perceives the ability of the aid to be inferior to one's own, perceived utility of the aid will be low. People may misuse an automated aid because the perceived utility of the aid is overestimated; they may disuse an aid when the perceived utility of the aid is underestimated.

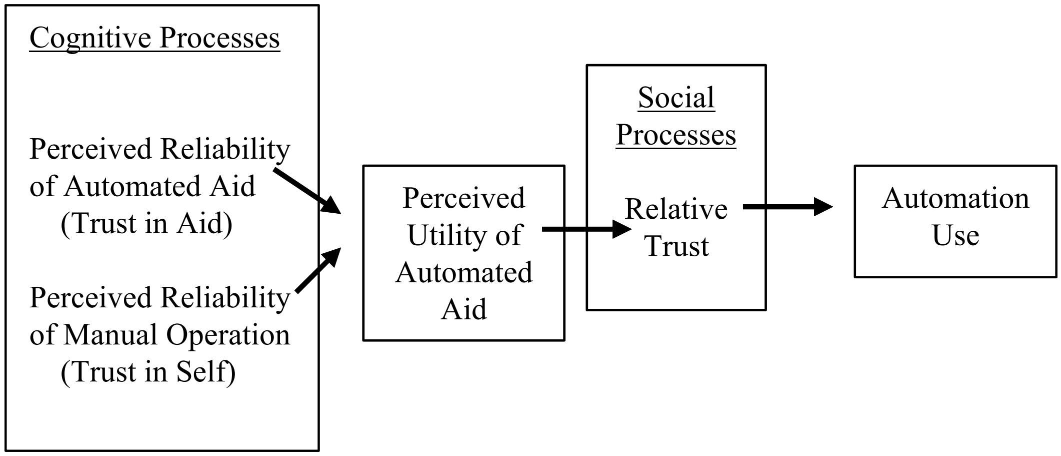

Figure 3: Relative Trust.

Since accurately perceiving the utility of the aid will lead to appropriate automation use, it is very important that we gain an understanding of how this perception is formed. The perceived utility of the aid will be most accurate when the *actual* ability of the aid and *actual* ability of the manual operator are compared. However, the "actual" ability is never known. Perceived ability is determined through a function of actual ability and error. The larger the error, the more likely misuse and disuse is to occur. We suspect that at least two types of errors occur.

One type of error occurs when human operators estimate their own performance. Human operators tend to overestimate their own ability. Social psychological literature is fraught with examples of self-serving biases. Humans exaggerate their contribution to a group product (appropriation of ideas, Wicklund, 1989), overestimate the number of tasks they can complete in a given period of time (planning fallacy, Buehler, Griffin, & Ross, 1994), are overconfident in negotiations (Neale & Bazerman, 1985), and inflate their role in positive outcomes (Whitley & Frieze, 1985). Thus, we hypothesize that human operators will be likely to overestimate their manual ability.

The other type of error occurs when human operators estimate the performance of their automated aid. Prior to working with the aid, the human must rely on stereotypes formed concerning the performance of automated aids. Although individual differences exist, a bias toward automation leads many people to predict near-perfect performances from automated aids. Dzindolet, Pierce, Beck, and Dawe (2002) told half the participants they would be provided with the decisions reached by the contrast detector before they made their soldier absent/present decision for each of 200 slides. Other participants were told they would be provided with the decisions reached by the prior participant before making their soldier absent/present decision for each of the 200 trials. The instructions informed the participants that their aid (human or automated) was *not* perfect. When asked to estimate the number of errors the automated aid would make in the upcoming 200 trials, participants predicted, on average, the aid would make only 14.19 errors (i.e., correct nearly 93% of the time). When asked to estimate the number of errors the human aid would make, participants predicted, on average, 46.17 errors (i.e., correct only 77% of the time – only 27% better than chance). Other researchers have also found a bias toward automation. For example, Dijkstra, Liebrand, and Timminga (1998) found students judged advice from an expert system to be more rational and objective than the same advice from a human advisor. However, Lerch, Prietula, and Kulik (1997) found greater trust in automated systems than human advisors only when the automation was an expert system and the human advisor was a novice.

They did not find participants were more likely to trust an automated expert system than a human expert advisor.

In summary, the perceived reliability of the automated aid is determined by the actual reliability of the automated aid and by the bias toward automation. The perceived reliability of manual operation is determined by the actual reliability of the human operator and by self-serving biases.

We hypothesize that what is important in determining automation use, though, is not the perceived reliability of the aid or the perceived reliability of manual control, but the result of a comparison process between the two, perceived utility. Increasing the reliability of the aid will not increase automation use unless the aid's perceived reliability surpasses that of manual operations.

This can explain the inconsistent findings concerning automation reliability and human-computer performance. For example, Riley (1994) and Moray, Inagaki, and Itoh (2000) found people were more likely to overly rely on a more reliable than a less reliable automated system. Yet, Dzindolet et al. (2001) did not find participants to rely on more reliable decision aids than less reliable ones.

Similarly, Parasuraman, Molloy, and Singh (1993) and Singh, Molloy, and Parasuraman (1997) did not find reliability to affect automation use with a monitoring task. Participants were required to simultaneously perform three flight simulation tasks. Two of the tasks required manual operation, but a third, monitoring task was automated. The automation varied in its reliability and in its variability. Some participants worked with an automated system that was consistently correct 87.5% of the time; others worked with a system that was consistently correct 56.25% of the time. Others worked with machines that varied in reliability. Some students started with automated systems that were accurate 87.5% of the time for the first ten minutes of the session of the experiment, but then alternated in reliability to 56.25% and 87.5% during ten-minute intervals for the rest of the experimental session. Other participants started working with a less reliable automated system that alternated to a more reliable every ten minutes during the experimental session.

Overreliance on automation was found by participants in both of the constant conditions (Parasuraman, Molloy, & Singh, 1993) and the saliency of the automation failure signals did not affect the results (Singh, Molloy & Parasuraman, 1997). Thus, when the reliability of the automation varied, participants appropriately relied on the automated monitoring system. However, when the system was consistent in its reliability (either high or low), participants tended to overly rely on the system's monitoring capabilities. Singh, Molloy and Parasuraman (1997) hypothesize that complacency was due to the participants being too trusting of the automation. "Participants may have begun the automated sessions with an equal amount of trust in the automation. However, as the reliability of the automation fluctuated for the variable group their trust may have declined. Therefore, the variability group may have been more skeptical of the automation, and, thus, been more vigilant for automation failures" (p. 28).

Why did reliability not affect automation use in the Parasuraman, Molloy and Singh (1993) and Singh, Molloy and Parasuraman (1997) studies? This finding is unexpected according to the APT model and Muir's model of trust. Our model predicts the automation's reliability will affect automation use only if it leads the human operator to perceive the aid's reliability as greater than his or her own.

In summary, we hypothesize that people may misuse an automated aid when the perceived utility of the aid is overestimated and disuse an aid when the perceived utility of the aid is underestimated. The perceived utility of the system results from a comparison between the automated system's perceived ability and one's own perceived ability. Perceived ability is hypothesized to be affected by actual ability and various biases (self-serving and bias toward automation).

Reducing the biases should decrease inappropriate automation use. Providing accurate feedback will assuage the self-serving biases. Dzindolet, Pierce, Beck, and Dawe (2002) found that providing feedback of the contrast detector's superior performance after each trial and/or at the end of the 200 trials reduced disuse. Beck et al. (2001) found that disuse could be reduced by providing participants multiple forms of feedback of their own and of the superior aid's performance. However, in both of these studies, some disuse was found – even when participants were continually reminded of the number of errors both they and their aid had made (and the aid had made half as many errors).

Understanding the reliability of the automated aid may be difficult for human operators, especially if they begin with a bias toward automation. Dzindolet et al. (in press) hypothesize that when a participant views an easy slide, quickly spots a target, and makes his or her decision with a high degree of confidence, he or she assumes that the automated aid will be in concurrence. When the automated aid indicates the target is absent, the participant is likely to notice the obvious error just committed by the aid. Without an understanding of why this error was made, this obvious error violates the trust the operator has in the aid's decisions. Trust may diminish slowly or may immediately drop to a low level. As long as participants are able to view the decisions made by their automated aids, obvious errors can be detected setting in motion the violation of trust. In some of the conditions in Study 2 of Dzindolet et al. (in press), participants could not view the aid's decisions. Results indicated appropriate automation reliance only among participants who received continuous feedback of their aid's superiority and were unable to view errors made by the decision aid.

In Study 3 of Dzindolet et al. (in press), participants were provided with an explanation as to *why* their automated aid might err. This manipulation was successful in increasing the participants' level of trust in the automated decision aid and reliance on the aid. Unfortunately, participants provided with inferior aids were just as likely to rely on them as those provided with superior aids. Thus, although providing participants with a rationale as to their aid's errors was successful in reducing disuse, in certain conditions, it led to misuse.

In addition to trust (through perceived utility) affecting automation use, we conjecture that two other social processes may affect automation use: feelings of control and moral obligation to rely on self. Figure 4 illustrates the social processes.

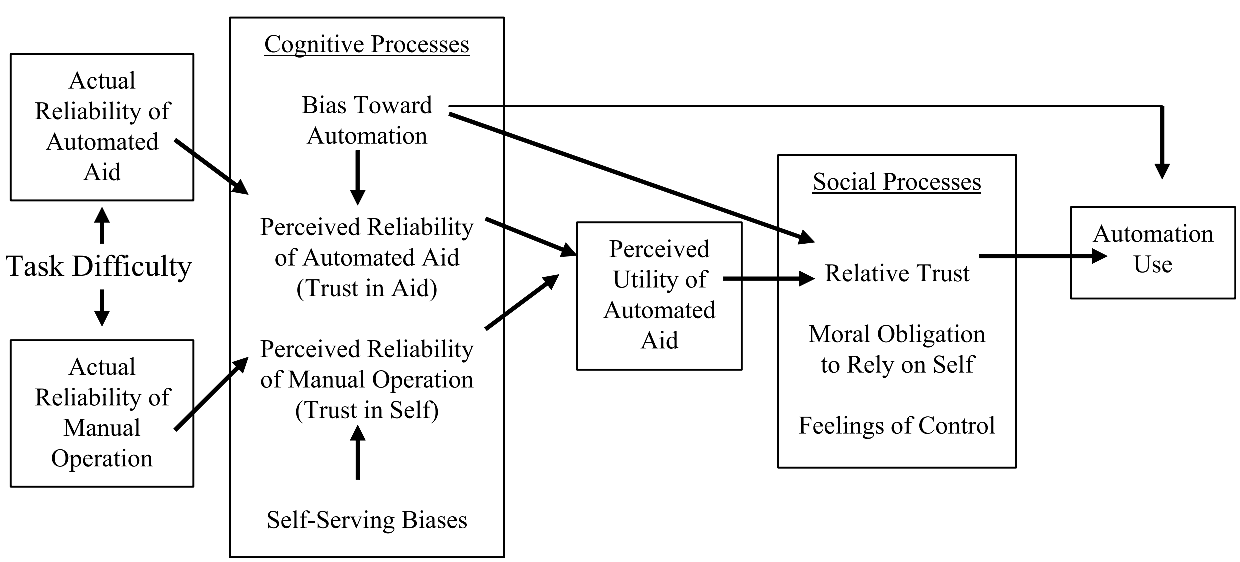

Figure 4: Social Processes.

Analyses of the justifications of the task allocation decisions provided by participants in one of the Dzindolet, Pierce, Beck, and Dawe (2002) experiments revealed that 71% of the students, who were provided cumulative feedback that indicated that the aid made about an equal number of errors as the participant, justified self-reliance with statements indicating they would not earn more rewards if they relied on the aid. Since the task allocation decision would not affect the size of their reward, why did participants opt for self-reliance? We hypothesize that self-reliance provides participants with an illusion of control. Langer (1983) has found that people often behave illogically in order to have an illusion of control.

In addition, many participants (though more working with human aids (n=24, 43.64%) than automated aids (n=9, 16.67%), $\underline{X}^2 = 8.58$, $\underline{p} < .01$) justified self- reliance with statements concerning a moral obligation to rely on oneself. One student wrote, "I would rather the amount of coupons I receive be based on my performance- it seems more 'fair' to myself." Another wrote, "I feel anything earned should be based on how well I did or didn't do."

Much research remains to be performed to explore the social processes and their effect on automation use. Only with a more clear understanding of these processes will we be able to suggest ways that misuse and disuse can be reduced.

The variables which affect perceived utility are of special interest to us because perceived utility is not only predicted to affect trust, but also to affect the last of the processes, motivational processes.

Motivational Processes: Diffusion of Responsibility

A third explanation of the over-reliance on automation discussed by Mosier and Skitka (1996) involves the idea that when working in a group, the responsibility for the group's product is diffused among the group members. Several researchers have thought of the human-computer system as a dyad or team in which one member is not human (Bowers, Oser, Salas, & Cannon-Bowers, 1996; Scerbo, 1996; Woods, 1996). Thus, the human may feel less responsible for the outcome when working with an automated system than when working without one. The person may not be as motivated to extend as much effort when paired with an automated system as when working alone. In the social psychological literature, this phenomenon has been dubbed social loafing (cf. Latane, Williams & Harkins, 1979) or free riding (Kerr & Bruun, 1983).

One theory which has been successful in accounting for much of the findings in the social loafing literature is Shepperd's Expectancy-Value Theory (1993; 1998). According to this theory, motivation is predicted from a function of three factors: expectancy, instrumentality, and outcome value.

Expectancy

The first factor, expectancy, is the extent to which members feel that their efforts are necessary for the group to succeed (see Figure 5). When members feel their contributions are dispensable, or when one's individual contribution is unidentifiable or not evaluated, one is likely to free ride, or work less hard (Kerr & Bruun, 1983; Williams & Karau, 1991).

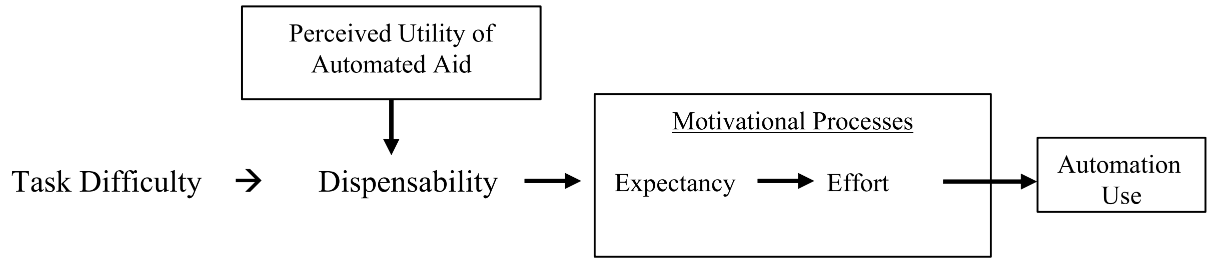

Figure 5: Expectancy.

With a human-computer system, individual contributions tend to be identifiable and evaluated, thus these variables are not thought to affect motivational processes. However, when the perceived utility of a system is high, one is likely to feel his or her efforts are more dispensable when working with a system low in perceived utility. Thus, we would expect human operators to be likely to misuse an automated system deemed more reliable than themselves in the same way people free ride on group members deemed more reliable than themselves.

Task difficulty, which is predicted to affect perceived utility (see Figure 4), has also been found to directly affect dispensability. In fact, one of the methods researchers have used to make group members feel their efforts are indispensable has been to imply that the difficulty of the task makes the demands on each group member particularly high (Shepperd, 1993). For example, Harkins and Petty (1982) asked participant to generate as many uses as they could for either a knife (easy task) or for a detached door knob (difficulty task) either alone or in nine-member groups. Although they found social loafing with the easy task (group members did not generate as many ideas as those working alone), they did not find social loafing with the difficult task.

In summary, the more dispensable the human operator is made to feel, the lower expectancy will be; effort will likely be low and the likelihood that the automated aid will be relied upon will be high. In some instances, this will lead to automation misuse.

Instrumentality

Instrumentality, the extent to which members feel that the group's successful performance will lead to a positive overall outcome, is also predicted to affect effort. Members who feel the outcome is not contingent on the group's performance are less likely to work hard. Thus, inappropriate use should be high among members who feel their group's performance is irrelevant. In one study, Shepperd (1998) varied instrumentality. Half the participants were told that the seven groups with the highest number of ideas generated (out of ten groups) would earn a reward. Other participants were told that the members of the four groups with the highest number of ideas generated (out of 40 groups) would have their names entered into a lottery. One name would be drawn and that person would earn a reward. Thus, in the former condition, members had a 7 in 10 chance of attaining the reward; in the latter condition, there was only a 1 in 200 chance of attaining the reward. In addition to the optimistic bias (participants estimated their chance of winning to be about 1 in 4), he found that performance suffered when instrumentality was lowered.

On the battlefield, there will be many soldier-computer teams. If the human determines that the overall outcome is not contingent on his or her human-computer team (either because he estimates other teams are more able to do the task or that his human-computer team is dispensable), then the human will put little effort into the task.

Outcome Value

Finally, the value of the outcome is predicted to affect motivation. Outcome value is the difference between the importance of the outcome and the costs associated with working hard. Increasing the costs or minimizing the importance of the reward will lead members to put forth less effort. More effort will be extended toward tasks that lead to valuable outcomes without requiring much cost. Costs vary with the number of other tasks one must perform, fatigue, intrinsic interest of the task, and cognitive overhead (see Figure 6).

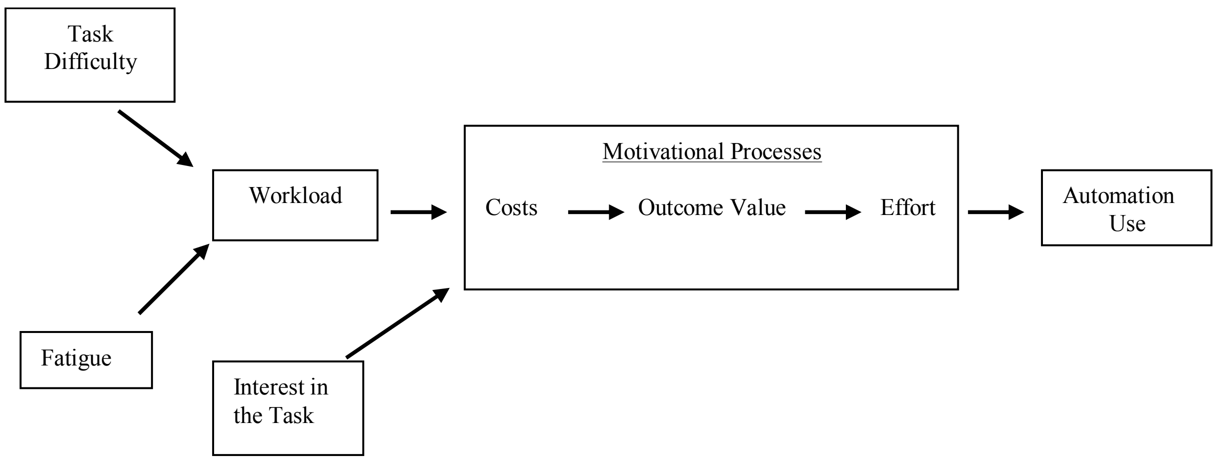

Figure 6: Costs.

Workload. As workload increases, the cost of performing a specific task increases, thereby increasing automation use (and the potential for misuse). In Study 2 of Parasuraman, Molloy, and Singh (1993), misuse did not occur among participants who were required only to perform the automated task. Overreliance on automation was found only among participants who had to manually perform two additional tasks. Similarly, Thackray and Touchstone (1989b) did not find differences in detection of failures between aided and unaided participants performing a single task. Thus, consistent with the framework, increasing the workload increased the likelihood for misuse.

Fatigue. The more fatigued the human operator is, the higher the cost of performing an additional task. Therefore, we would expect automation use to be greater with more fatigued human operators. In some instances, this will lead to appropriate automation use, but in other instances, this will lead to misuse of the automated aid.

Interest in the Task. The costs of performing a task that is intrinsically interesting are less than the costs of performing a less interesting task. For example, the cost of performing boring, redundant tasks is higher than the cost of performing interesting tasks. If the task is boring enough, the costs might outweigh the importance; outcome value will be decreased, effort will be lower, and automation use and the potential for misuse will rise.

In summary, if costs of manual operation are larger than importance due to large workload, fatigue of the human operator, and/or boredom of the task, then outcome value will be low. The amount of effort expended on the task is predicted to be low. Thus, the likelihood for automation use and the potential for misuse will be great.

Costs will only affect automation use, however, if they are deemed greater than the importance of the outcome. If the importance of the outcome is deemed greater than the cost, outcome value will be high, thereby increasing effort. The likelihood for automation use is low in this situation making the possibility for disuse potentially large.

Importance is predicted to be affected by the rewards of successful task completion and the penalties of task failure (see Figure 7).

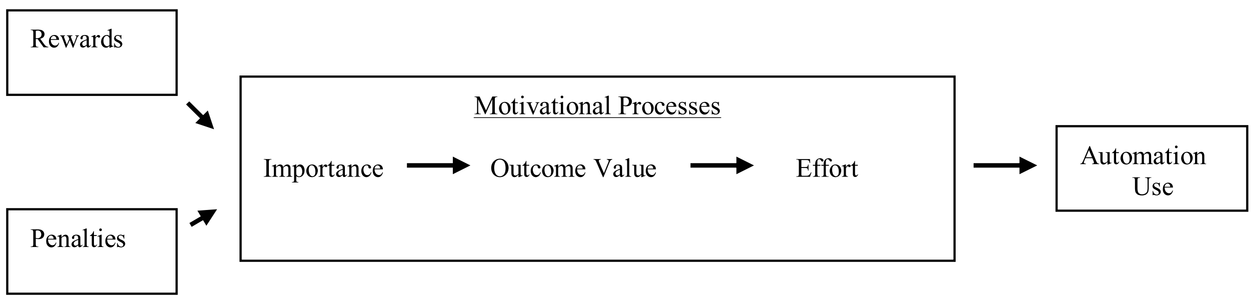

Figure 7: Importance of Outcome.

Rewards. When successful completion of the task leads to highly valued resources (e.g., money, prestige), and if the rewards are not outweighed by the costs of succeeding, then outcome value will be high and participants are predicted to work hard. Human operators will not rely on the automated aid; rather they will put forth the effort to complete the task. Oftentimes, this strategy will be appropriate; sometimes disuse will occur.

Penalties. Similarly, when grave penalties are a consequence of successful completion of the task not occurring, specifically, when the penalties of failure outweigh the costs of succeeding, then outcome value will be high and participants are predicted to work hard. In some situations, this will lead to an optimal task allocation decision; sometimes it will lead to disuse of automated aids.

On the battlefield, sub-optimal human-computer performance can be lethal; outcome value is extremely high. In combat, disuse may be more of a problem than misuse. Among such highly motivated people, misuse may not exist at all. This is consistent with findings from some interviews with Gulf War soldiers, who indicated they turned off their automated systems.

For this reason, it is imperative that some of the research testing the model be performed in more combat-like environments. At the very least, researchers should examine automation reliance while varying the consequences for successful task performance. Although it is intuitively appealing, albeit, somewhat obvious, to predict that instituting positive consequences for successful performance would lead to improved automation reliance, this may not necessarily occur. In the Dzindolet, Pierce, Beck, and Dawe (2002) and Moes et al. (1999) experiments, participants earned coupons or extra-credit for correct decisions. Yet, in all three studies, disuse prevailed. Of course, in these studies, the consequences were not varied. We predict that increasing the rewards and/or the penalties for correct decisions would have decreased disuse. Future research needs to determine the validity of such predictions.

In summary, penalties and rewards will affect the importance of successful completion of the task. If importance is greater than costs associated with performing the task, outcome value will be high, leading effort to be expended, and automation use to be less likely. In some situations, this will lead to an optimal task allocation decision; sometimes it will lead to disuse of automated aids.

CONCLUSIONS

The Framework of Automation Use predicts that cognitive, motivational, and social processes work together to cause misuse, disuse, and appropriate automation use. Many factors affect each of the processes (see Figure 8), and may therefore, affect automation use. The reliability of the automated aid, the reliability of manual operation and several cognitive biases (including self-serving and the bias toward automation) combine to affect the perceived utility of the aid. When the perceived utility of the aid is high, the operator is likely to trust the aid and feel dispensable; his or her efforts are not necessary for the task to be completed. Automation use is predicted to be high through both social and motivational processes. Fatigue, workload, intrinsic interest in the task, penalties for task failure, and rewards for task completion combine to affect the outcome value, which also will affect the effort the human will expend on the task than the likelihood he or she will rely on the automated aid (see Figure 8).

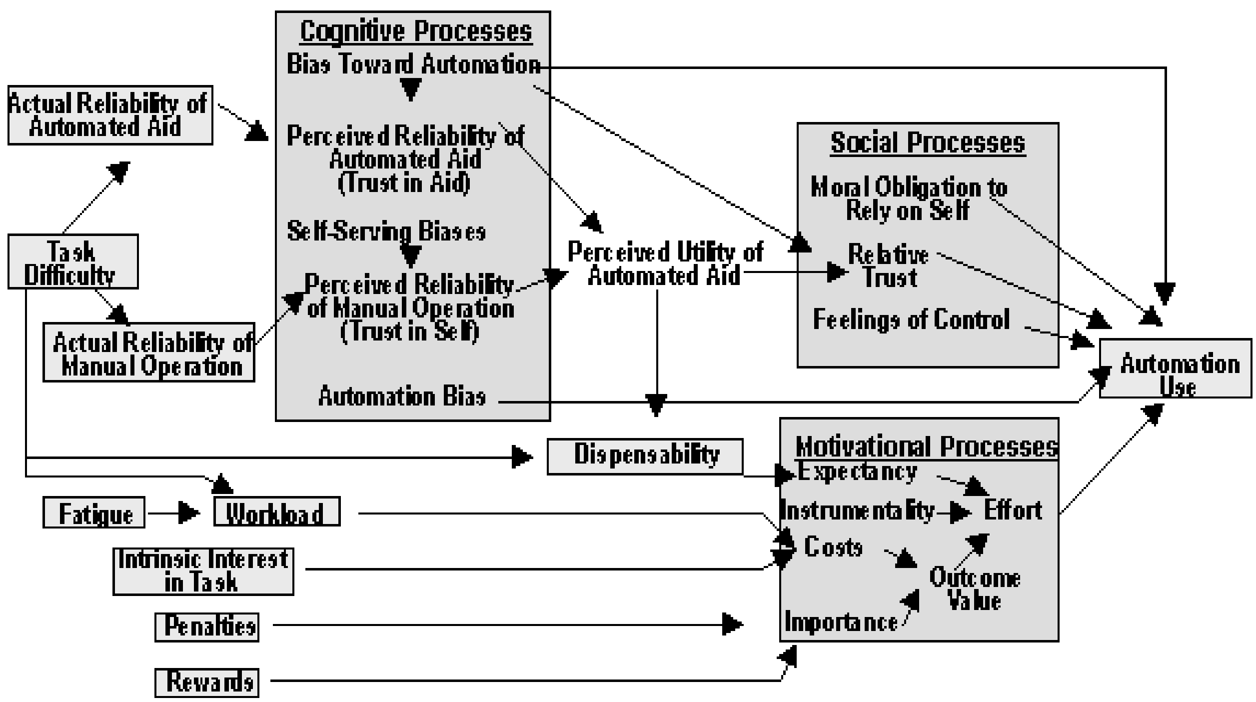

Figure 8: Framework of Automation Use (Dzindolet, Beck, Pierce & Dawe, 2001).

In addition to this theoretical perspective, a strong look at the empirical evidence gained from the studies performed with the paradigms discussed at the beginning of the paper may help to guide system designers and trainers. Figure 9 presents such a summary.

SUMMARY OF EMPIRICAL FINDINGS	
1.	Under some conditions, people provided with an automated decision aid perform worse than people not provided an automated decision aid.
2.	Improving the reliability of the automated decision aid does not necessarily improve the performance of the human-automated team.
3.	People are more likely to appropriately rely on a human aid than appropriately rely on an automated decision aid. They overestimate the automated aid's performance prior to interaction with the aid and underestimate the automated aid's performance after interaction with the aid.
4.	The following techniques may lead to more appropriate automation reliance: • Reduce people's initial trust in an automated decision aid • Provide people with continuous feedback and eliminate the possible detection of obvious errors • Provide people with various types of information concerning their performance and performance of the automated decision aid.
5.	Appropriate automation use is not enhanced with training in conditions in which the aid is reliable or in conditions in which the aid is not reliable.
6.	People trust both an inferior and superior automated aid and rely on it more when trained as to why the aid might err.

Figure 9: Summary of Empirical Findings.

System designers and trainers can use the empirical findings and the Framework of Automation Use to help create situations which encourage human operators to appropriately rely on automated decision aids. Future researchers need to further examine the framework, determine the effect of each of the cognitive, social, and motivational processes on automation use, and examine the interaction of the three processes. We believe the framework will prove useful to researchers interested in reducing automation misuse and disuse.

REFERENCES

Arkin, R., & Baumgardner, A.H. (1985). Self-Handicapping. In J.H. Harvey & G. Weary (Eds.), *Basic Issues in Attribution Theory and Research*. New York: Academic Press.

Beck, H.P., Dzindolet, M.T., & Pierce, L.G. (2002). Applying a Decision-Making Model to Understand Misuse, Disuse, and Appropriate Automation Use. In E. Salas, C.A. Bower, N. Cooke, J. Driskell, & D. Stone (Eds.) *Advances in Human Factors and Cognitive Engineering*, Vol 2, PLACE: JAI Press.

Bowers, C.A., Oser, R.L., Salas, E., & Cannon-Bowers, J.A. (1996). Team Performance in Automated Systems. In R. Parasuraman, & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications. Human Factors in Transportation*. Mahwah, NJ: Lawrence Erlbaum Associates.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "Planning Fallacy": Why People Underestimate Their Task Completion Times. *Journal of Personality and Social Psychology, 67,* 366-381.

Cesar, E.M. (1995). *Strategies for Defining the Army's Objective Vision of Command and Control for the 21st Century*, Santa Monica, CA: RAND, MR-487-A.

Cohen, M.S., Parasuraman, R., & Freeman, J.T. (1998, July). Trust in Decision Aids: A Model and its Training Implications. *Proceedings of the 1998 Command and Control Research and Technology Symposium*. Washington, DC: CCRP.

Corcoran, D.W., Dennett, J.L., & Carpenter, A. (1972). Cooperation of Listener and Computer in a Recognition Task: II. Effects of Computer Reliability and "Dependent" versus "Independent" Conditions. *Journal of the Acoustical Society of America, 52,* 1613-1619.

Dalal, N.P. & Kasper, G.M. (1994). The Design of Joint Cognitive Systems: The Effect of Cognitive Coupling on Performance. *International Journal of Human-Computer Studies, 40*, 677-702.

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., & Beck, H.P. (in press). The Role of Trust in Automation Reliance. Paper submitted to *International Journal of Human Computer Studies: Special Issue on Trust and Technology*.

Dzindolet, M.T., Pierce, L.G., Beck, H.P., & Dawe, L.A. (2002). The Perceived Utility of Human and Automated Aids in a Visual Detection Task. *Human Factors, 44,* 79-94.

Dzindolet, M.T., Pierce, L.G., Pomranky, R.A., Peterson, S.A., & Beck, H.P. (October, 2001). Automation Reliance on a Combat Identification System. *Proceedings of the Human Factors and Ergonomics Society Meeting*, Santa Monica, CA: HFES.

Dzindolet, M.T., Pierce, L.G., Beck, H.P., Dawe, L.A., & Anderson, B.W. (2001). Predicting Misuse and Disuse of Combat Identification Systems. *Military Psychology, 13(3),* 147-164.

Dzindolet, M.T., Beck, H.P., Pierce, L.G., & Dawe, L.A. (2001). *A Framework of Automation Use (Rep. No. ARL-TR-2412).* Aberdeen Proving Ground, MD: Army Research Laboratory.

Hamilton, D.L., & Sherman, S.J. (1989). Illusory Correlations: Implications for Stereotype Theory and Research. In D. Bar-Tal, C.F. Graumann, A.W. Kruglanski & W. Stroebe (Eds.), *Stereotyping and prejudice: Changing conceptions.* New York: Springer-Verlag.

Haplin, S.M., Johnson, E.M., & Thornberry, J.A. (1973). Cognitive Reliability in Manned Systems. *IEEE Transactions on Reliability, R-22,* 165-170.

Harkins, C.G., & Petty, R.E. (1982). Effects of Task Difficulty and Task Uniqueness on Social Loafing. *Journal of Personality and Social Psychology, 43,* 1214-1229.

Kerr, N.L., & Bruun, S.E. (1983). Dispensability of Member Effort and Group Motivation Losses: Free-Rider Effects. *Journal of Personality and Social Psychology, 44,* 78-94.

Latane, B., Williams, K.D., & Harkins, S. (1979). Many Hands Make Light the Work: The Causes and Consequences of Social Loafing. *Journal of Personality and Social Psychology, 37,* 822-832.

Layton, C., Smith, P.J., & McCoy, C.E. (1994). Design of a Cooperative Problem-Solving System for En-Route Flight Planning: An Empirical Evaluation. *Human Factors, 36,* 94-119.

Lee, J.D., & Moray, N. (1992). Trust, Control Strategies and Allocation of Function in Human-Machine Systems. *Ergonomics, 35,* 1243-1270.

Lee, J.D., & Moray, N. (1994). Trust, Self-Confidence, and Operators' Adaptation to Automation. *International Journal of Human-Computer Studies, 40,* 153-184.

Leyens, J., & Yzerbyt, V.Y. (1992). The In-Group Overexclusion Effect: Impact of Valence and Confirmation on Stereotypical Information Search. *European Journal of Social Psychology, 22,* 549-569.

Moes, M., Knox, K., Pierce, L.G., Beck, H.P. (1999). *Should I Decide or Let the Machine Decide for Me?* Poster presented at the meeting of the Southeastern Psychological Association, Savannah, GA.

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive Automation, Trust, and Self-Confidence in Fault Management of Time-Critical Tasks. *Journal of Experimental Psychology: Applied, 6,* 44-58.

Mosier, K.L., & Skitka, L.J. (1996). Human Decision-Makers and Automated Decision Aids: Made for Each Other? In R. Parasuraman, & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications. Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates.

Muir, B.M. (1987). Trust Between Humans and Machines, and the Design of Decision Aids. *International Journal of Man-Machine Studies, 27,* 527-539.

Muir, B.M. (1994). Trust in Automation: Part I. Theoretical Issues in the Study of Trust and Human Intervention in Automated Systems. *Ergonomics, 37,* 1905-1922.

Mullen, B., & Hu, L. (1989). Perceptions of Ingroup and Outgroup Variability: A Meta-Analysis Integration. *Basic and Applied Social Psychology, 10,* 233-252.

Mullen, B., Johnson, C., & Salas, E. (1991). Productivity Loss in Brainstorming Groups: A Meta-Analytic Integration. *Basic & Applied Social Psychology, 12,* 3-23.

Neale, M.A., & Bazerman, M.H. (1985). The Effects of Framing and Negotiator Overconfidence on Bargaining Behaviors and Outcomes. *Academy of Management Journal, 28,* 34-49.

Parasuraman, R. (1987). Human-Computer Monitoring. *Human Factors Special Issue: Vigilance: Basic and Applied Research, 29,* 695-706.

Parasuraman, R., Molloy, R., & Singh, I.L. (1993). Performance Consequences of Automation-Induced "Complacency." *International Journal of Aviation Psychology, 3,* 1-23.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors, 39,* 230-253.

Riley, V. (1994). Human Use of Automation. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 55,* 2425.

Scerbo, M.W. (1996). Theoretical Perspectives on Adaptive Automation. In R. Parasuraman, & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications. Human Factors in Transportation.* Mahwah, NJ: Lawrence Erlbaum Associates.

Shepperd, J.A. (1993). Productivity Loss in Performance Groups: A Motivation Analysis. *Psychological Bulletin, 113,* 67-81.

Shepperd, J.A. (1998). *Expectancy Value Theory*. Presentation presented at the Midwestern Psychological Association, Chicago, IL.

Singh, I.L., Molloy, R., & Parasuraman, R. (1993). Automation-Induced "Complacency": Development of the Complacency-Potential Rating Scale. *International Journal of Aviation Psychology, 3,* 111-122.

Singh, I.L., Molloy, R., & Parasuraman, R. (1997). Automation-Induced Monitoring Inefficiency: Role of Display Location. *International Journal of Human-Computer Studies, 46,* 17-30.

Skowronski, J.J., & Carlston, D.E. (1989). Negativity and Extremity Biases in Impression Formation: A Review of Explanations. *Psychological Bulletin, 105,* 131-142.

Sorkin, R.D., & Woods, D.D. (1985). Systems with Human Monitors: A Signal Detection Analysis. *Human-Computer Interaction, 1,* 49-75.

Tan, G., & Lewandowsky, S. (1996). *A Comparison of Operator Trust in Humans Versus Machines.* Presentation of the First International Cyberspace Conference on Ergonomics: http://www.curtin.edu.au/conference/cyberg/centre/paper/tan/paper.html.

Thackray, R.I. & Touchstone, R.M. (1989). Detection Efficiency on an Air Traffic Control Monitoring Task With and Without Computer Aiding. *Aviation, Space & Environmental Medicine, 60,* 744.748.

Thackray, R.I. & Touchstone, R.M. (1989). Effects of High Visual Taskload on the Behaviours Involved in Complex Monitoring. *Ergonomics, 32,* 27-38.

Tversky, A., & Kahneman, D. (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology, 5,* 207-232.

Whitley, B.E., Jr., & Frieze, I.H. (1985). Children's Causal Attributions for Success and Failure in Achievement Settings: A Meta-Analysis. *Journal of Educational Psychology, 77,* 608-616.

Wicklund, R.A. (1989). The Appropriation of Ideas. In P.B. Paulus, (Ed.), et al., *Psychology of Group Influence*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wiener, J.L. (1981). A Theory of Human Information Processing for Economists. *Dissertation Abstracts International, 42,* 809.

Williams K.D., & Karau, S.J. (1991). Social Loafing and Social Compensation: The Effects of Expectations of Coworker Performance. *Journal of Personality and Social Psychology, 61,* 570-581.

Woods, D.D. (1996). Decomposing Automation: Apparent Simplicity, Real Complexity. In Parasuraman, R., & Mouloua, M. (Eds.), *Automation and Human Performance: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.