Encouraging Human Operators to Appropriately Rely on Automated Decision Aids

Mary T. Dzindolet,
Cameron University
Department of Psychology
2800 Gore Blvd.
Lawton, OK 73505
(580) 581-2514
maryd@cameron.edu

Hall P. Beck, and
Appalachian State University
Department of Psychology
Boone, NC 28608
(828) 262-2725
beckhp@appstate.edu

Linda G. Pierce
Army Research Laboratory
Bldg. 3040, Room 220
Fort Sill, OK 73503-5600
(580) 442-7139
piercel@sill.army.mil

Abstract

Information technology is changing the nature of the military decision-making process. However, the underlying assumption in employing human-automated system teams, namely that the team will be more productive than the human or the automated system would be alone, is not always met. Under certain conditions, human operators err by overly relying or under-utilizing automated systems [Parasuraman and Riley, 1997]. A Framework of Automation Use [Dzindolet *et al.*, 1999] posits that cognitive, social, and motivational processes combine to predict automation use. Two studies were performed to examine social processes, controlling for cognitive and motivational processes. The framework posits that when human operators' estimates of the reliability of the automated system and manual operation are accurate, appropriate use of automation is most likely to occur. Various ways of communicating to human operators information concerning the reliability of their own and an automated aid's decisions were examined in an effort to encourage human operators to appropriately rely on automated decision aids. Both studies found alarming rates of disuse. However, provision of many sources of information concerning the reliability of the automated decision aid was successful in reducing the bias toward disuse. Results have implications for both training and system design.

1. Automated Decision Aids

Dramatic increases in the use of automation in recent years have occurred in the military [cf. Cesar, 1995]. Information technology is changing the nature of the military decision-making process by providing decision makers with more relevant, accurate, and timely information than was previously possible. The underlying assumption in employing these human-automated system "teams" is that the teams will be more productive than either the automated system or the human would be working alone. While some researchers have found support for this underlying assumption [Dalal and Kasper, 1994], others have found human operators often overly rely on (misuse) or underutilize (disuse) automated decision systems [Parasuraman and Riley, 1997]. Some speculate that this inappropriate use is the main reason why automation of decision making has not advanced as far as it has in other areas [Cohen *et al.*, 1999].

By understanding the processes that human operators use when allocating tasks to automated operations, one may be able to better design systems that encourage appropriate use of automated aids. What are the processes leading to the decision to rely on or ignore an automated system? Drawing from Lee and Moray [1992; 1994], Mosier and Skitka [1996], and Shepperd [1993],

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2000**	2. REPORT TYPE	3. DATES COVERED **00-00-2000 to 00-00-2000**

4. TITLE AND SUBTITLE **Encouraging Human Operators to Appropriately Rely on Automated Decision Aids**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Cameron University,Department of Psychology,2800 Gore Blvd,Lawton,OK,73505**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		**10**	
unclassified	**unclassified**	**unclassified**			

Dzindolet, Pierce, Beck, and Dawe [1999] created a framework that predicts automation use from cognitive, motivational, and social processes (see Figure 1).

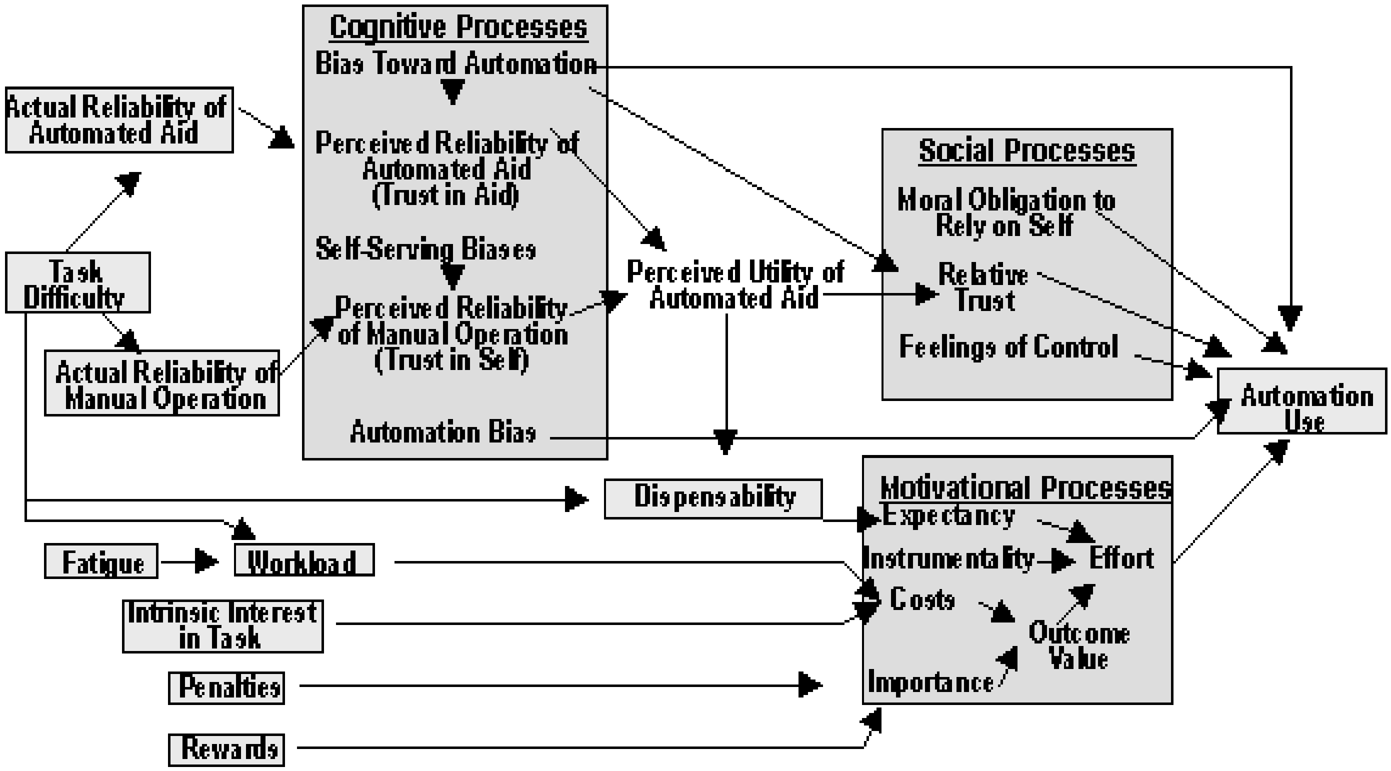

Figure 1. Framework of Automation Reliance [Dzindolet *et al.*, 1999]

1.1 *Cognitive Processes*

One reason why people may overly rely on automated systems when making decisions has to do with the manner in which they process the information provided by the automated aid. Rather than going through the cognitive effort of gathering and processing information, the decision supplied by the automated system is used. Often, this strategy is optimal; however, under certain conditions, this reliance may be inappropriate and misuse will occur. Relying on a decision aid in a heuristic manner is dubbed the automation bias [Mosier and Skitka, 1996]. Although the automation bias can explain misuse of automated decision aids, it cannot account for the disuse often found.

In order to eliminate the automation bias, participants in the two studies reported in this paper were only provided with the decision of their automated aid *after* they indicated their decision and their level of confidence in their decision. This procedure prevented participants from relying on the automated aid's decision in a heuristic manner. After all, they did not even know the automated aid's decision until after they had made their decision.

1.2 ***Motivational Processes***

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 [e.g., Bowers *et al*., 1996]. Thus, the human may feel less responsible for the outcome when working with an automated system than when working alone and may extend less effort. In the social psychological literature, this phenomenon has been dubbed social loafing or free riding [Kerr and Bruun, 1983]. One theory that has been successful in accounting for much of the findings in the social loafing literature is Shepperd's Expectancy-Value Theory [1993]. According to this theory, motivation is predicted from a function of three factors: expectancy, instrumentality, and outcome value.

1.2.1 ***Expectancy***

Expectancy is the extent to which members feel that their efforts are necessary for the group to succeed. 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 and Bruun, 1983]; misuse of automated systems is likely.

1.2.2 ***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.

1.2.3 ***Outcome Value***

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, increasing the likelihood they rely on their automated aids. Costs vary with the intrinsic interest of the task, the number of other tasks one must perform, fatigue, and cognitive overhead. Importance varies with personal importance of successfully completing the task, and the rewards and penalties of successful task completion.

In order to control for motivational processes in the two studies, we created a dependent variable for which automation reliance would require as much effort as self-reliance. Specifically, after completing 200 trials, participants were asked to indicate whether they would rather have their reward contingent on responses *they* had made in the past or on responses their *automated aid* had made. In this paradigm, self-reliance did not require more effort to be expended than automation reliance. In this way, motivational processes were virtually eliminated.

1.3 ***Social Processes***

Another process leading to automation use, misuse, and disuse has to do with the role of the computer as an expert. If human operators perceive the automated aid to be more reliable than manual operation, they are likely to place greater trust in the automated system, and rely on it. In this situation, if the system truly *is* more reliable than manual operation, appropriate automation use will take place. However, when people underestimate manual operation and/or overestimate the reliability of the automated system, then misuse will occur.

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). The outcome of the decision process has been named the perceived utility of the automated aid [Dzindolet *et al.*, 1999] and is predicted to be directly related to the relative trust of the automated aid (and automation use).

The perceived utility of the automated aid will be most accurate when the *actual* ability of the aid and the *actual* ability of the manual operator are compared. Unfortunately, the actual reliability of the aid and of the manual operator are unlikely to be accurately perceived by the operator. In reality, errors and biases are likely to occur. The larger the errors and biases, the more likely misuse and disuse are to occur.

One type of error occurs when human operators estimate their own performance. Social psychological literature has numerous examples of self-serving biases; the results from several studies examining automation use have found that human operators, without feedback, are likely to overestimate their manual ability. Dzindolet *et al.* [1999] found that without feedback, participants inappropriately assessed their performance to be superior to that of their automated aid, which led to disuse of the automated systems.

The other type of error occurs when human operators estimate the performance of their automated aid. Human operators often predict near-perfect performances from automated aids [Dzindolet *et al.*, 1999]. This may be due in part to a demand characteristic created when the human operator and the automated aid are paired. Human operators (soldiers and experimental participants) may assume they would only be provided with an automated aid of high quality.

According to the model, when human operators' estimates of the reliability of automation and manual operation are accurate, appropriate use of automation is most likely to occur. Two studies were performed to examine various ways of communicating to human operators information concerning the reliability of their own and an automated aid's decisions in an effort to encourage human operators to appropriately rely on automated decision aids.

2. **Study 1**

Disuse of an automated decision aid has been found with a paradigm in which students are instructed to view slides of Fort Sill terrain in search of a camouflaged soldier. After students indicate their present/absent decision, an automated system ("contrast detector") indicates its decision. Following

200 trials, students decide to receive rewards based on the number of correct decisions made by either the automated system or the student on 10 randomly chosen trials. Even when told that the automated system made fewer errors, the majority of students make the rewards contingent on their own decisions rather than on the decisions of the automated aid [Dzindolet *et al.*, 1999; Hawkins *et al.*, 1999; Moes *et al.*, 1999].

One reason for the disuse may be that participants expect automated systems to be near-perfect in performance [cf. Dzindolet *et al.*, 1999]. When the automated system makes an error, this expectation is violated. The violations are likely to be remembered leading human operators to underestimate the reliability of the automated aid. Providing operators with information concerning the reliability of the aid (e.g., "The aid usually makes about 10 errors") may reduce the disuse. Study 1 explored this possibility.

2.1 ***Method***

Before performing the task, half of the participants in the study were informed that the aid usually makes about ten errors in 200 trials; the other half were not given any information regarding the aid's typical performance. In addition, after performing 200 trials, half of the participants were provided feedback indicating the number of errors they and their automated aid had made. The aid's performance was manipulated such that the aid always made approximately half as many errors as the participant made. The other half of the participants were not provided with cumulative feedback. Thus, a 2 (knowledge of the number of errors usually made by the automated aid prior to performing the 200 trials) X 2 (knowledge of the number of errors made by the participant and the aid at the completion of the 200 trials) between subjects design was utilized.

2.1.1 ***Participants***

Fifty Cameron University students participated in the study. Most students received extra credit in a course offered in the Department of Psychology and Human Ecology for their participation. Guidelines set forth in the American Psychological Association Guidelines for Ethical Conduct were strictly followed.

2.1.2 ***Materials***

The workstation contained a Hewlett-Packard Vectra PC, 133 mhz CPU with 32 mb of 60 ns RAM, including an S3, Inc. Trio 64 Plug-n-Play PCI video card. The 17-inch Hewlett-Packard Ultra VGA monitor was set at High Color (16 bit) resolution, 800 X 600 pixels. Slides of Fort Sill terrain (see Figure 2) were presented for about ¾ of a second.

Figure 2. Sample Slide

2.1.3 ***Procedure***

After signing informed consent forms, participants read an instruction page along with the experimenter. They were told they would view 200 slides displaying pictures of Fort Sill terrain on a computer screen. The instructions indicated that about half of the slides contained one soldier ("target") in various levels of camouflage; the remaining slides were of terrain only. Participants were told that sometimes the soldier would be rather easy to spot; other times he would be more difficult to find. Each slide would be presented on the computer screen for about ¾ of a second. After viewing the slide, it would be replaced by a screen asking them to indicate whether or not they believed the soldier was in the slide. They were told they had as much time as they needed to make their decision. Next, they would be asked to indicate the extent to which they were certain their decision was correct. A five-point scale ranging from “highly confident” to “not at all

Participants were told that a computer routine had been written to assist them in performing their task. They were told that the routine performed a rapid scan of the photograph looking for contrasts that suggested the presence of a human being. If the contrast detector routine determined the soldier was probably present, the word "PRESENT" and a red circle would appear. If the contrast detector routine determined the soldier was probably absent, the word "ABSENT" and a green circle would appear. The aid’s decision was provided to participants only *after* they had indicated their decision and their level of decision confidence. Half the participants were told that the aid usually made about ten errors during the 200 trials.

The instructions explained that there were two possible errors that could be made. One error was made when they indicated that the soldier was present when, in fact, he was not. The other error was made when they indicated that the soldier was not present when, in fact, he was. Participants were told that both errors were equally serious and that they should attempt to avoid them.

Participants performed four practice trials, were provided an opportunity to ask questions, and then proceeded to view the 200 slides. When they completed the task, half the participants were told the number of errors they and their aid made. Finally, students were told they could earn $.50 in coupons to be used at the cafeteria in the Student Union for every correct decision made on ten randomly chosen trials. Participants had to **choose** whether the performance would be based on **their decisions** or on **the decisions of their aid**. After making their choice, students were asked to justify their choice in writing. Finally, participants completed a brief survey concerning their experience.

2.2 ***Results***

Though not reaching standard levels of significance, results indicated disuse of the automated aid was reduced only for participants provided with *both* prior information about the aid's reliability and actual information about the aid's superior performance, $\underline{X}^2(3) = 7.30$, $\underline{p} < .07$. (See Table 1).

	Knowledge of Expected Number of Errors Made by Aid *(Students were told "The contrast detector usually makes about 10 errors.")*	No Prior Information of Aid's Reliability	Total Number of Participants
Knowledge of Number of Errors Made by Aid and Participant *(This was manipulated such that the aid always made about half as many errors as the participant.)*	58.33%	100.00%	24
No Information of Actual Errors Made by Aid or Participant	84.62%	84.62%	26

Table 1. Percentage of Students Who Relied on Their Decisions Rather than the Decisions of Their Superior Automated Aid

2.3 ***Discussion***

The combination of instilling more realistic expectations of an automated decision aid's performance and cumulative feedback indicating the automated aid's superior performance reduced the bias toward disuse. However, even when provided with both types of information, the majority of students still chose to ignore the superior automated aid. Would other forms of feedback also help to reduce the bias toward disuse? Study 2 explored this possibility.

3. **Study 2**

In Study 1, participants were not provided immediate feedback concerning the accuracy of their or their aid's decisions. Providing feedback after each trial would allow participants to more realistically evaluate their performance and the performance of their automated aid. Therefore, in Study 2, half the participants were provided with feedback after each trial; half were not. As in Study 1, half the participants were provided with cumulative feedback indicating the aid's performance was superior to that of their own; half were not. Finally, prior to deciding to rely on the decisions of the automated aid or their own decisions, half the participants were informed that most other students in past studies would have earned greater rewards had they relied on the decisions of the automated aid.

3.1 ***Method***

The method was similar to that of Study 1 except that extra credit points in a psychology course were offered as the reward rather than coupons redeemable at the Student Union. Seventy female and twenty-three male undergraduates from a southeastern comprehensive university were randomly assigned to the 2 (Individual Trial Feedback) x 2 (Cumulative Results Feedback) x 2 (Prior Results Feedback) between-subjects design.

Individual Trial Feedback was given immediately after each of the 200 trials and consisted of telling the participant if the soldier was or was not present on that trial. Cumulative Feedback and Prior Results Feedback were delivered following the 200th trial. Cumulative Feedback reported the total errors made by the participant and the aid; this information indicated the aid made about half as many errors as the participant. With Prior Results Feedback, participants were told that persons who relied on the detector averaged more reward (i.e., extra credit points) than persons who relied on themselves. Each type of feedback was delivered via message boxes displayed on a computer monitor.

3.2 ***Results***

Analyses revealed that the three forms of feedback had equivalent effects on reducing the bias toward disuse. In addition, the more types of feedback that students received, the more likely they were to rely on the superior automated aid $\underline{X}^2(3) = 6.52$, $\underline{p} < .05$. See Table 2.

Number of Feedback Sources	Self	Automated Decision Aid	Total
0	16 (84%)	3 (16%)	19
1	20 (69%)	9 (31%)	29
2	19 (66%)	10 (34%)	29
3	7 (44%)	9 (56%)	16
Total	62 (67%)	31 (33%)	93

Table 2. Relationship of the Number of Feedback Sources and the Decision to Rely on Own Or on Superior Automated Decision Aid

4. Conclusions

The results of the two studies are consistent with the outcomes of earlier investigations, which find inappropriate use of automated aids. An alarming 82% of participants in Study 1 and 67% of participants in Study 2 disused their automated decision-aid.

However, provision of many sources of information concerning the reliability of the automated decision aid and the performance of the human operator was successful in reducing the bias toward disuse. It is hypothesized that this information increased the accuracy of the perceived utility of the automated aid, leading to more appropriate automation use (see Figure 1). Perhaps it is necessary for human operators to use the automated system in realistic situations and to be provided with focused feedback to develop a clear understanding of when the system will and will not be accurate. Armed with this information, human operators will be most likely to appropriately rely on the automated system. Clearly, future research in methods that encourage human operators to appropriately rely on automated systems is needed. The results from these studies will have implications for both training and system design.

5. References

[Bowers *et al.*, 1996] C. A. Bowers, R. L. Oser, E. Salas, and J. A. Cannon-Bowers. Team performance in automated systems. In Raja Parasuraman and M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 92-99). Lawrence Erlbaum Associates, Mahwah, NJ, 1996.

[Cesar, 1995] E. M. Cesar. *Strategies for defining the army's objective vision of command and control for the 21st century*, RAND, MR-487-A, Santa Monica, CA, 1995.

[Cohen *et al.*, 1999] M. S. Cohen, R. Parasuraman, and J. T. Freeman. *Trust in decision aids: A model and its training implications*. (Technical Report USAATCOM TR 97-D-4). Cognitive Technologies, Inc, Arlington, VA, 1999.

[Dalal and Kasper, 1994] N. P. Dalal, and G. M. Kasper. The design of joint cognitive systems: The effect of cognitive coupling on performance. *International Journal of Human-Computer Studies,* 40, 677-702, 1994.

[Dzindolet *et al.*, 1999] M. T. Dzindolet, L. G. Pierce, H. P. Beck, and L. A. Dawe. Misuse and disuse of automated aids. In *Proceedings of the Human Factors Society 43rd Annual Meeting* (pp. 339 - 343). Human Factors and Ergonomics Society, Santa Monica, CA, 1999.

[Hawkins *et al.*, 1999] J. L. Hawkins, M. T. Dzindolet, and H. P. Beck. *Bias against machines with decision making capabilities*. Poster session presented at the annual meeting of the Southeastern Psychological Association, Savannah, Georgia, 1999.

[Kerr and Bruun, 1983] N. L. Kerr and S. E. Bruun. Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology,* 44, 78-94, 1983.

[Lee and Moray, 1992] J. D. Lee and N. Moray. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 10, 1243-1270, 1992.

[Lee and Moray, 1994] J. D. Lee and N. Moray. Trust, self confidence, and operators' adaptation to automation. *International Journal of Man-Human Studies,* 40, 153-184, 1994.

[Moes *et al*., 1999] M. Moes, K. Knox, L. G. Pierce, and H. P. Beck. *Should I decide or let the machine decide for me?* Paper presented at the meeting of the Southeastern Psychological Association, Savannah, Georgia, 1999.

[Mosier and Skitka, 1996] K. L. Mosier and L. J. Skitka. Human decision makers and automated decision aids: Made for each other? In Parasuraman, R., and Mouloua, M. (Eds.), *Automation and human performance: Theory and applications*. Lawrence Erlbaum Associates, Mahwah, NJ, 1996.

[Shepperd, 1993] J. A. Shepperd. Productivity loss in performance groups: A motivational analysis. *Psychological Bulletin,* 113, 67-81, 1993.

[Parasuraman and Riley, 1997] R. Parasuraman and V. Riley. Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39, 230-253, 1997.