Violations of Expectations: A Study of Automation Use

Regina A. Pomranky
Cameron University
Department of Psychology
and Human Ecology
2800 W Gore Blvd.
Lawton, OK 73505
pomrankyr@sill.army.mil

Mary T. Dzindolet
Cameron University
Department of Psychology
and Human Ecology
2800 W Gore Blvd.
Lawton, OK 73505
maryd@cameron.edu

Scott A. Peterson
Cameron University
Department of Psychology
and Human Ecology
2800 W Gore Blvd.
Lawton, OK 73505
scottp@cameron.edu

Graduate Student Paper

Abstract

Although the use of automated decision systems is increasing, the performance of these human computer teams is often less than optimal (Parasuraman & Riley, 1997). This study examined the effect of violations of expectations on appropriate automation use in a paradigm that controlled motivational and cognitive processes implicated in affecting automation usage decisions (Dzindolet, Pierce, Beck & Dawe, 1999). One hundred and ninety-five Cameron University students were randomly assigned to one of the conditions described by a 2 (Aid's Performance Level: Superior or Inferior) X 2 (Provision of Aid's Decision: Decisions Present or Decisions Absent) X 3 (Type of Feedback: No Feedback, Cumulative Feedback, or Continuous Feedback) design or to a condition that examined initial expectations of an automated decision aid. Although participants initially had a bias toward automation, a bias against automation was found after 200 trials in most conditions. Only when participants were unable to detect obvious errors made by the automated aid but were provided with continuous feedback was appropriate use found, $\chi^2(11) = 34.45$, $p < .01$. Implications for future research are discussed.

Introduction

With the introduction of automation into the daily lives of humans comes many concerns as to how to encourage the human-automated team to perform optimally. Using automation appropriately, a rational decision-maker would rely on automation when doing so would increase the probability of successful mission completion and would rely on manual operations when relying on automation would decrease the probability of successful mission completion. Parasuraman and Riley (1997) identified two types of non-rational behavior often associated with automation usage decisions (AUD): Misuse and Disuse. Misuse is when a human operator tends to overly rely on an automated aid and disuse is when a human operator tends to ignore an automated aid (Parasuraman & Riley, 1997). Dzindolet, Pierce, Beck and Dawe (2000) introduced a general model of AUD in which cognitive, social, and motivational processes of the human operator combine to predict misuse, disuse, and appropriate automation use.

Cognitive Processes

Mosier and Skitka (1996) hypothesized that when people make decisions, they may tend to overly rely on automated systems due to faulty cognitive processes. 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 & 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 present study 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.

Motivational Processes

When working in a group, the responsibility for the product is diffused among 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, Oser, Salas, & Cannon-Bowers, 1996). Thus, the human may feel less responsible for the outcome when working with an automated system than when working alone and therefore may extend less effort.

In order to control for motivational processes in the present study, we used 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 controlled.

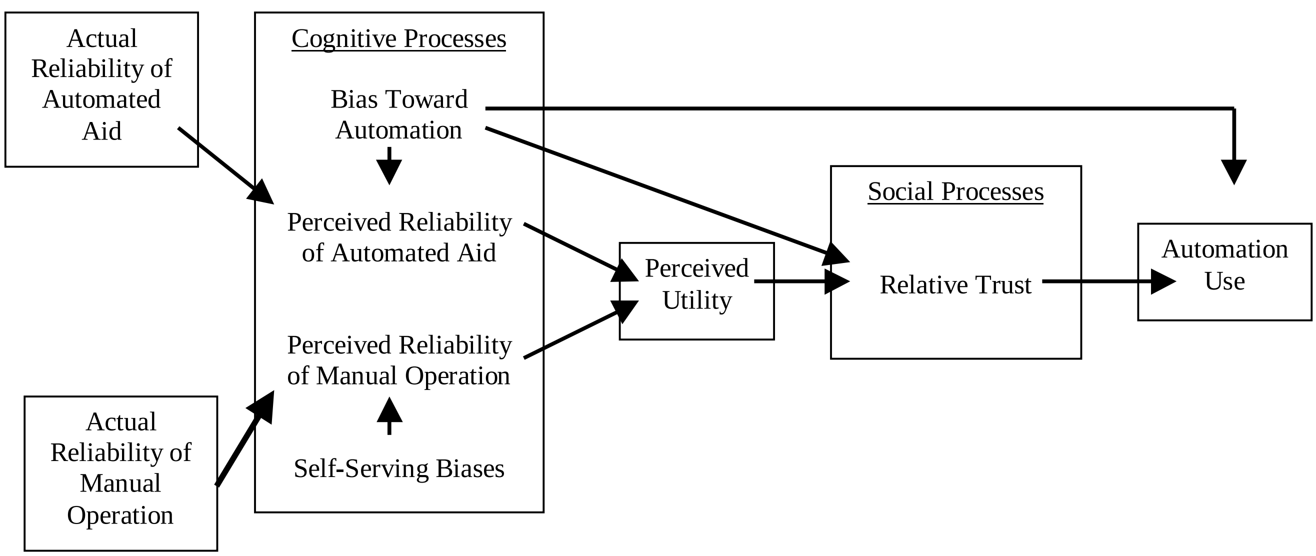

Figure 1. Portion of the Framework of Automation Use (Dzindolet et al., 2000)

Social Processes

Another process predicted to lead to automation use, misuse, and disuse is the role of the computer as an expert (Dzindolet et al., 2000). 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. If the system truly *is* more reliable than manual operation, appropriate automation use will take place. However, when people inaccurately estimate manual operation and/or the reliability of the automated system, then inappropriate automation use may occur.

Automation use is determined from the outcome of a comparison process between the perceived reliability of the automated aid and the perceived reliability of manual control. The outcome of the decision process has been named the perceived utility of the automated aid (Dzindolet et al., 2000) 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.

The goal of this study was to measure the effects of various levels of performance awareness to understand the most effective way to make human operators more aware of information concerning the reliability of their own decisions and their automated aid's decisions in an effort to reduce these biases and errors and encourage human operators to appropriately rely on automated decision aids. To focus on the perceived utility, the motivational processes were controlled and the potential for automation bias was eliminated.

One 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; Study 3). Does this bias toward automation lead to misuse and an over-reliance on automation? According to Dzindolet et al. (1999, Study 1), when controlling motivational processes and eliminating the automation bias, disuse, not misuse was likely to occur. Inconsistent with the results from Dzindolet et al.'s (1999; Study 3) study, participants were not more likely to rely on an automated aid than a human aid, showing instead a strong bias against automation. This study suggested that something happened during the experience with the aid that changed participants' bias toward automation (Dzindolet et al., 1999; Study 3) to a bias against automation (Dzindolet et al., 1999; Study 1). What occurs during the interaction with automation to change the bias from misuse to disuse? Further, what techniques can be used to induce appropriate reliance on automated systems?

Expectations held by the human operator prior to interaction with automated aids may lead to disuse of the automated aid. Prior to interaction with an automated system people seem to have a schema concerning automation, specifically that automated aids are reliable and accurate. It is this schema that leads human operators to expect almost perfect performances from their automated systems. Researchers in cognitive psychology have found that information inconsistent with a schema is likely to be well remembered and play an unduly large role in information processing (Ruble & Stangor, 1986; Smith & Graesser, 1981). When an event occurs that is in direct conflict with an operator's original expectations, the operator will be more likely to remember this event.

Any error made by the automated aid is inconsistent with the human operator's schema that the automated aid is reliable and accurate. Therefore each error made by the aid is likely to be remembered and as the task progresses it may be hard for the participant to retain an accurate picture of the aid's reliability. The pieces of contradictory information are overemphasized and exaggerated producing a distorted view of the aid's abilities. This may lead operator to underestimate the performance of the automated aid. It may be that a salient form of feedback is needed to overcome these experiences.

In addition, when an operator is presented with an easy task, quickly decides a course of action, and makes that decision with a high degree of confidence, the operator assumes the automated aid will be in concurrence. When the automated aid reaches a different decision, the operator is likely to notice the obvious error just committed by the aid and will question the reliability of his automated system. Repeated exposure to obvious errors committed by an automated aid may result in a distorted view of the automated system's perceived utility. As long as participants are able to view the decisions made by their automated systems, which will include some obvious errors, disuse will occur.

In summary, in paradigms in which the automation bias is eliminated and motivational processes are controlled, an initial bias toward automation becomes, by the end of 200 trials, a bias against automation. The focus of this study was to explore what might occur during the interaction with automation to change the bias toward automation into a bias against automation. Two possible sources of violations of expectations were examined in this study. In order to help human participants retain an accurate picture of the aid's reliability, the operator must be continuously aware of his or her aid's performance levels. Some of the participants were provided with a bar graph that displayed the number of errors made by the human operator and the automated aid during the entire task (Continuous Feedback); others were provided with a similar bar graph only at the completion of the task (Cumulative Feedback); still others were not provided with any type of feedback (No Feedback).

To eliminate the detection of obvious errors, some participants were unable to view the decisions by the automated aid (Decisions Absent); others were able to view the aid's decisions (Decisions Present).

In order to examine both misuse and disuse, some participants were paired with an automated aid that made twice as many errors as the participant (Inferior); others were paired with an automated aid that made half as many errors as the participant (Superior).

This 2 (Aid's Performance Level: Superior or Inferior) X 2 (Provision of Aid's Decision: Present or Absent) X 3 (Type of Feedback: No Feedback, Cumulative Feedback, or Continuous Feedback) design allowed for an examination of the effects of obvious errors and distorted views of the aid's ability on AUD both individually and cumulatively. In addition, one condition (Initial Expectations) of this study replicated Dzindolet et al.'s (1999, Study 3) study concerning the bias toward automation.

The purpose of this study was to examine various ways of making human operators more aware of information concerning the reliability of their own decisions and their automated aid's decisions in an effort to encourage human operators to appropriately rely on automated decision aids.

Methods

Participants

One hundred ninety-five Cameron University students volunteered to participate in this study. Some received extra credit in their course offered in the Psychology and Human Ecology Department for their participation. Those enrolled in General Psychology fulfilled a research requirement for that course. Guidelines set forth in the American Psychological Association Guidelines for Ethical Conduct were strictly followed.

Materials

The two workstations used in this study, consisted of a Hewlett Packard Vectra PC, 133 mhz CPU with 32mb of RAM, including an S3, Inc. Trio 64 Plug-n-Play PCI video card. The monitors were 17-inch Hewlett-Packard Ultra VGA and were set at High Color (16-bit) resolution, 800 x 600 pixels. The operating system used was Windows 95, version 4.00.950. Slides of Fort Sill terrain were presented as stimuli. The participants were told that they would be paired with an automated aid called a contrast detector. In addition, they were told that the contrast detector would scan the image of the photograph looking for contrast that suggested the presence of a human being. If the contrast detector indicated that a soldier was probably present, the word "present" and a green circle would appear in the contrast detector box. Conversely, if the contrast detector determined that the soldier was probably absent, the word "absent" and a red circle would appear in the contrast detector box. In reality, there was no contrast detector. The reliability was manipulated such that the aid (contrast detector) was either superior, making half as many errors, or inferior, making twice as many errors, as the participant.

Procedure

Participants were randomly assigned to one of the 2 (Aid's Performance Level: Superior or Inferior) X 2 (Provision of Aid's Decision: Decisions Present or Decisions Absent) X 3 (Type of Feedback: No Feedback, Cumulative Feedback, or Continuous Feedback) conditions or to the Initial Expectations condition. Appropriate instruction pages were provided. The experimenter read the instruction page aloud while the student followed along. After the experimenter answered the participants' questions, participants performed four practice trials. For each trial, the student viewed a slide of Fort Sill terrain for about 3/4 of a second, made their decision regarding the absence or presence of the soldier, indicated their confidence level in their absent/present decision, and viewed the decision made by the contrast detector. About half of the slides contained one soldier in various levels of camouflage (see Figure 2); the remaining slides were of terrain only.

Figure 2. Sample Slide

After viewing a slide, the participants were asked to indicate whether or not they believed the soldier was present in the slide and indicate their level of confidence in their absent/present decision on a five-point Likert-type scale.

In order to determine the effect of viewing obvious errors committed by an automated aid, some participants viewed the decision reached by the automated aid after indicating their decision (Decisions Present); others were not provided with that information (Decisions Absent). To improve the accuracy of the participants' reliability estimates of the automated aids, about one third of the participants were presented with a continuous feedback display that remained on the screen and was updated every 5 trials (Continuous Feedback). During the four practice trials the display was on the screen but did not report any information until after the completion of the fourth trial. Another third of the participants received a cumulative feedback display, presented twice; once at completion of the four practice trials and once at the end of the 200 trials (Cumulative Feedback). Both the cumulative and continuous forms of feedback were graphically represented indicating the number of errors the participant made and the number of errors their automated aid made. Other participants did not receive any feedback (No Feedback). Depending on the aid's performance level (Superior or Inferior), the feedback indicated that the aid made half as many errors or twice as many errors as the participant.

Participants in the Initial Expectations condition were provided with the decisions reached by the automated aid for each of the four practice trials. After the four practice trials, these participants completed a survey.

After completing four practice trials, all other participants performed 200 trials. Just as in the practice slides, some participants were presented with a continuous feedback display that remained on the screen throughout the 200 trials and was updated every 5 trials or a cumulative feedback display presented only after completion of the 200 trials. Both forms of feedback were graphically represented indicating the number of errors the participant made and the number of errors their automated aid made.

After completing the 200 trials, participants were told that 10 trials would be randomly selected from the prior 200. Coupons would be earned for each correct decision made for these 10 selected trials. Participants were asked to decide whether they would like to rely on the decisions made by their automated aid or whether they would like to rely on their own decisions. After indicating their preference, participants were asked to explain why they chose to rely on their own decisions or on the decisions of the automated aid.

Results

Initial Expectations

To loosely replicate the findings of Dzindolet et al. (1999; Study 3), participants in the Initial Expectations condition were asked to estimate their and their automated aid's expected performance on 200 trials after completing only four practice trials. Two items directly assessed participants' expected relative performance. T-tests comparing the mean responses to these items with the midpoint of the scale (5) did not reveal a bias toward automation. Pairs of survey items asked participants to rate how well they and their aid would perform on the upcoming trials, to indicate how much they could trust their and their aid's decisions, and to estimate the number of errors they and their aid would make. Dependent t-tests performed to compare the responses between the item pairs revealed a significant difference for only one of the pairs. No differences in *ratings* of expected performance or trust were found. However, when asked to estimate the number of errors, the bias toward automation was revealed. Participants estimated that the automated aids would make nearly 30 fewer errors than they ($\underline{M}_D = 28.73$, $\underline{t}(14) = 2.76$, $\underline{p} < .02$).

Performance Data

A chi square test of independence was used to determine whether participants' condition, described by the 2 (Aid's Performance Level: Superior or Inferior) X 2 (Provision of Aid's Decision: Decisions Present or Decisions Absent) X 3 (Type of Feedback: No Feedback, Cumulative Feedback, or Continuous Feedback) design, was independent of their decision to choose to rely on their or their aid's decisions. Results indicated that condition and automation reliance were related, $\chi^2(11) = 34.45$, $\underline{p} < .01$. Table 1 presents the frequencies and percentages of those in each condition who chose to rely on their decisions rather than those of their automated aid.

	Superior Aid	Inferior Aid
	No Feedback	
Decision	73.33% (11)	80.00% (12)
No Decision	80.00% (12)	86.67% (13)
	Cumulative Feedback	
Decision	73.33% (11)	86.67% (13)
No Decision	60.00% (9)	73.33% (11)
	Continuous Feedback	
Decision	73.33% (11)	73.33% (11)
No Decision	13.33% (2)	93.33% (14)

Table 1: Performance Data: Percentage Of Those Who Chose To Rely On Self

Consistent with prior data (Dzindolet et al., 1999, Study 1; Moes, Knox, Pierce, & Beck, 1999), a bias toward self-reliance is evident. The bias toward self-reliance was attenuated only among participants who were continually provided with feedback indicating the superior performance of their aid but were unable to view their aid's decisions.

Justification Data

After indicating their self-reliance or automation-reliance decision, participants were asked to justify their decision. Content analysis was performed with the participants' response to the prompt: "We would like to know what led you to your decision to base your performance on either your decisions or on the decisions of the aid. Please tell us everything you thought of in coming to this decision. Don't worry about spelling or grammatical errors. Use the back side of this paper if necessary." Each justification was typed into a database with the participant's number and condition. Two raters categorized each justification into separate categories. The 180 participants generated 189 justifications. Eight of the justifications were in lone categories and were not included in the final analysis.

The remaining 181 justifications fell into 14 different categories (Table 2).

Category	n	Percentage
Confidence in self	65	35.91
Detected obvious errors	33	18.23
I had fewer errors	18	9.94
Contrast detector missed less	16	8.84
Don't trust computers	8	4.42
Not confident in self	7	3.87
Aid made more errors	7	3.87
Wanted to know my performance	6	3.31
Military experience	6	3.31
Contrast detector did better	5	2.76
Computer said I made more errors	4	2.21
Compared the contrast detector to me	2	1.1
Randomly chose	2	1.1
Reward for my performance	2	1.1

Table 2: Number and Percentage of Participants using each of the Fourteen Categories of Justifications

However, 85% of the justifications fell into seven categories representing four general constructs: (1) confidence (or lack of confidence) in self (e.g., “I was not confident in what I saw”; "I chose to use 'my decisions' because I trust my observations, and I never second guess my self"), (2) detection of obvious errors (e.g. "There were a few times that I'm pretty sure I saw the soldier, but the program said he was absent"), (3) relative performance (e.g., “I had less errors than the computer”, “The contrast detector made less errors”, “The computer made more mistakes compared to mine,” and (4) trust in computers (“I don’t trust computers that much. I know a lot about their tendency for errors”). Only 15% of the justifications fell into the remaining seven categories (see Table 2).

Discussion

In order to improve human-computer performance, it is imperative that we examine the processes that lead to AUD. Dzindolet et al. (2000) generated a framework of automation use, in which they suggested that cognitive, motivational, and social processes combine to predict AUD. In order to focus on social processes, Dzindolet et al. (1999) and Moes et al. (1999) used a paradigm that eliminated the automation bias and controlled motivational processes. In this paradigm, a bias toward self-reliance has been found. The results of this study are consistent with this previous research. Nearly three quarters (73%) of the participants who were made aware of their superior aid’s decisions after every trial but were not given feedback as to the number of errors the aid or they made, chose to rely on themselves rather than on their superior aid.

This bias against automation is especially interesting considering that after limited interaction with an automated system (only four trials), participants believed that the contrast detector would commit fewer errors than they (Dzindolet et al., 1999, Study 3). In this study, participants who had completed only four trials (Initial Expectations condition) estimated that their aid would make nearly 30 fewer errors than they.

Violations of the expectation that the contrast detector would perform well could explain how the initial bias toward automation becomes a bias against automation after 200 trials. Two techniques were developed to reduce the impact of the violations of expectations: (1) The continuous feedback display served as a constant reminder of the relative performance of the aid reducing the possibility that participants could distort the perceived utility of the automated aid, and (2) Preventing participants from viewing the decisions reached by the automated aid removed the possibility of participants noticing an obvious error committed by the automated aid.

The conditions in which the best performance occurred were the ones in which participants were unable to detect errors made by their automated aid but were provided with a continuous feedback display. When paired with a superior aid, only two out of 15 participants chose to disuse the aid; when paired with an inferior aid, only one of the 15 participants chose to misuse the aid. Thus in this condition, not only was disuse virtually eliminated, misuse was not encouraged.

These results are consistent with the framework of automation use proposed by Dzindolet et al. (2000). In addition, disuse of automation was reduced without causing misuse. Therefore when controlling for the motivational processes and eliminating the automation bias, providing continuous feedback of the perceived utility of the automated aid and eliminating the detection of obvious error made by the automated aid led to appropriate AUD.

Future studies should examine the effects of these techniques on AUD in a paradigm that allows motivational and cognitive processes to flourish. In addition, future research should address how the solutions explored in this study could be best implemented. Obviously, it would be ridiculous to suggest that human operators be equipped with an automated decision making system that does not allow the user to see the aid's decision. The sole purpose of such a system is to aid the user when he/she is faced with a decision! Similarly, it is not practical or even feasible to provide a system with a continuous feedback display. There would be no way of gauging right or wrong decisions since the truth or right decision is not known by the system or the user. It is oftentimes the ambiguity of the situation that calls for the use of an automated decision making system to aid the user in making decisions. Therefore, the concepts of providing continuous feedback and not allowing the operator to see the decisions of the automated aid are not currently applicable to today’s automated decision-making systems. However, the results of this study suggest that making human operators more aware of information concerning the perceived utility of their automated aid with a more salient form of feedback brings the field one step closer to encouraging human operators to appropriately rely on them. The results of this study suggest that only when an operator’s expectations of a near perfect automated aid are not violated will appropriate use occur.

References

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.

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

Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2000). A framework of automation use, (ARL-TR-2412),Aberdeen Proving Ground, MD: Army Research Laboratory.

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.

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.

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

Ruble, D. N., & Stangor, C. (1986). Stalking the elusive schema: Insights from developmental and social-psychological analyses of gender schemas. Social Cognition, 4, 227-261.

Smith, D. A., & Graesser, A. C (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. Memory & Cognition, 9, 550-559.