THE IMPACT OF S1 VERSUS S2 THINKING ON AUTOMATION USAGE DECISIONS

A Thesis
by
COURTNEY C. CORNELIUS

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTER OF ARTS

December 2015
Department of Psychology

THE IMPACT OF S1 VERSUS S2 THINKING ON AUTOMATION USAGE DECISIONS

A Thesis
by
COURTNEY C. CORNELIUS
December 2015

APPROVED BY:

____
Hall P. Beck, Ph.D.
Chairperson, Thesis Committee

____
Chris A. Dickinson, Ph.D.
Member, Thesis Committee

____
Timothy D. Ludwig, Ph.D.
Member, Thesis Committee

____
James C. Denniston, Ph.D.
Chairperson, Department of Psychology

____
Max C. Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Abstract

THE IMPACT OF SYSTEM 1 VERSUS SYSTEM 2 ON AUTOMATION USAGE DECISIONS

Courtney Cornelius
B.S., Appalachian State University
M.A., Appalachian State University

Chairperson: Hall P. Beck, Ph.D.

Automation usage decisions (AUDs) were examined via a System 1–System 2 framework. This study examined intent errors, a type of suboptimal AUD, which occurs when the operator knowingly chooses an option with a lower probability for success. The task, composed of a series of target-detection trials, simulated firing decisions that soldiers often make during combat. Two hundred three individuals were randomly assigned to conditions in a 2 (Automation: No-aid, Aid) x 3 (Credit Points: 0, 3, 6) design. Operators in the aid condition differed from those in the no-aid condition in that they could rely on "advice" from an error-free machine, a predominately S1 task. Credit points, which had a monetary consequence, could be earned by correct (firing) decisions. The dependent variables were the total of incorrect responses and the latency of responses. Participants in the aid condition were predicted to have shorter response latencies and make fewer errors in comparison to those in the no-aid condition. Additionally, operators were expected to have longer latencies when making incorrect than correct AUDs due to greater S2 involvement. Because prior research has indicated that S2 processing tends to increase as a function of task

consequence (Kahneman, 2011), a main effect for the credit points variable was also anticipated. Results were consistent with the premise that the presence of an error-free machine produced fewer errors and shorter responses latencies from reliance on the S1 system. However, the availability of the machine only produced shorter response latencies when the operators' AUD was correct. If the AUD was incorrect latencies were longer in the aid than the no-aid condition. The credit points variable was not statistically significant in any of the analyses. Overall, this study demonstrated the value of an S1-S2 framework in helping to understand why humans make seemingly irrational choices when interacting with automated decision aids.

Acknowledgments

I would like to express my deepest gratitude to my advisor, Dr. Hall Beck, for helping me achieve a goal of becoming a Master's graduate. Throughout my entire college experience, he has provided me with treasured advice and wisdom that I will keep close to my heart. I would also like to thank my committee members, Dr. Timothy Ludwig and Dr. Christopher Dickinson, for their patience, support, and expertise. Additionally, I would like to thank my program director, Dr. Rose Mary Webb, for encouraging me throughout this process. Finally, I would like to thank my friends and family for always supporting me.

Table of Contents

List of Tables

List of Figures

The Impact of System 1 versus System 2 on Automation Usage Decisions

Courtney C. Cornelius

Appalachian State University

Abstract

Automation usage decisions (AUDs) were examined via a System 1–System 2 framework. This study examined intent errors, a type of suboptimal AUD, which occurs when the operator knowingly chooses an option with a lower probability for success. The task, composed of a series of target-detection trials, simulated firing decisions that soldiers often make during combat. Two hundred three individuals were randomly assigned to conditions in a 2 (Automation: No-aid, Aid) x 3 (Credit Points: 0, 3, 6) design. Operators in the aid condition differed from those in the no-aid condition in that they could rely on “advice” from an error-free machine, a predominately S1 task. Credit points, which had a monetary consequence, could be earned by correct (firing) decisions. The dependent variables were the total of incorrect responses and the latency of responses. Participants in the aid condition were predicted to have shorter response latencies and make fewer errors in comparison to those in the no-aid condition. Additionally, operators were expected to have longer latencies when making incorrect than correct AUDs due to greater S2 involvement. Because prior research has indicated that S2 processing tends to increase as a function of task consequence (Kahneman, 2011), a main effect for the credit points variable was also anticipated. Results were consistent with the premise that the presence of an error-free machine produced fewer errors and shorter responses latencies from reliance on the S1 system. However, the availability of the machine only produced shorter response latencies when the operators’ AUD was correct. If the AUD was incorrect latencies were longer in the aid than the no-aid condition. The credit points variable was not statistically significant in any of the analyses. Overall, this study demonstrated the value of an S1-S2 framework in helping to understand

why humans make seemingly irrational choices when interacting with automated decision aids.

The Impact of System 1 Versus System 2 on Automation Usage Decisions

A voluminous automation literature has developed over the past half century. Most investigations have focused on topics such as the development of new design models (e.g., Parasuraman, 2000; Sheridan, 2002), the effects of technology (e.g., Cho & Chang, 2008; McBride, Rogers, & Fisk, 2011), willingness to use automation (e.g., Lee & Moray, 1994; Maehigashi, Miwa, & Terai, 2014; Parasuraman & Manzey, 2010), and the impact of automation on society (e.g., Franchimon & Brink, 2009; Ullmann, 1966).

Beck, Dzindolet, and Pierce (2002) pointed out that operators frequently have the choice to perform a task manually or by using one or more levels of automation. They proposed that these automation usage decisions, which they called AUDs, could be profitably examined using a classical decision-making framework. To illustrate this model, consider a two-choice situation in which the operator elects to use the manual control or a form of automation. The optimal choice is the AUD with the greatest probability of a successful performance. Conversely, a suboptimal AUD is a decision to engage an option with a lower probability of success.

One reason that AUDs are important is that they happen with great frequency. Some common AUDs include setting an alarm clock instead of trusting your circadian rhythm, using self-checkout instead of relying on a cashier, and choosing an electric toothbrush over a manual model. Although many AUDs are not of great consequence, others change lives and alter the course of history.

A successful AUD played a critical role in one of the defining events of the last millennium. Before taking “one giant leap for mankind” (National Aeronautics and Space Administration [NASA], 1995), Neil Armstrong found himself in a precarious situation. The computing capacity on Apollo 11 was less than many modern laptops. As data flowed in, the computer overloaded and was unable to properly position the spacecraft for a flat, smooth landing. Armstrong quickly engaged manual control, located a suitable landing site, and safely brought the spacecraft to the surface of the Moon (NASA, 1995). Had Armstrong taken more than 25 additional seconds to land, a lack of fuel would have stranded him on the Moon.

Not all AUDs have desirable outcomes. In 1985, a technician accidentally programmed a lethal dose of radiation for a cancer patient (Casey, 1998). Recognizing her mistake, she switched the machine back to a lower dosage. Unfortunately, the software engineer had not programmed the machine to react to sudden corrections. Rather than decrease the dosage level, the machine sent an error message to the technician. At this point, the technician should have taken the manual alternative and shut off the machine. Instead of investigating possible malfunctions, she relied on automation and continued administering multiple doses. This AUD ultimately resulted in the patient's death from radiation poisoning.

Suboptimal AUDs may be subdivided into misuse and disuse. As the name implies, misuse is overreliance, using an automated device when the task could be better accomplished by a nonautomated or less automated alternative. Disuse is underutilization of automation; failing to use automation that would improve the likelihood of a successful performance (Parasuraman & Riley, 1997).

The Washington Metro has had a series of costly and fatal crashes originating from suboptimal AUDs. In 1996, an operator overran the Shady Grove platform. Instead of relying on the manual braking system, she depended on an automated braking system that failed to account for inclement weather. The operator of the train was killed due to automation misuse (National Transportation Safety Board [NTSB], 1996). In 2004, another Washington Metro rolled backwards, hitting a service train. It appears that disuse was the culprit because the automatic braking system was turned off (NTSB, 2006).

Beck et al. (2002) specified three reasons that are likely causes of misuse and disuse. Both automation misuse and disuse can result from recognition, appraisal, or intent errors. Recognition errors happen when the operator is unaware of an alternative means of control and chooses an option with a lower probability of success. For example, assume that a global positioning system is faster and more accurate than the manual alternative, a paper map. An individual new to technology may continue to rely on a paper map simply because they do not recognize that there is an option with a higher probability of success.

Appraisal errors occur when the operator incorrectly assesses which option will have the highest probability of a successful outcome. Consider a driver who uses cruise control to maintain a safe speed instead of manually pressing the gas pedal. Although cruise control may be efficient for a stretch of open roads, this AUD would be an appraisal error if a driver failed to recognize that cruise control is suboptimal in a high-traffic scenario.

Although appraisal errors are well-established, intent errors have been seldom studied in the automation literature. Intent errors are made when the operator knows which AUD will have the highest probability for success yet chooses the suboptimal option. This type of suboptimal AUD may have contributed to the loss of Amelia Earhart and her navigator, Fred Noonan. In addition to manual methods of using landmarks for navigation, Earhart's plane was equipped with an automated radio system. For illustrative purposes, let us assume that the radio navigation was the more accurate option when flying across the vastness of the

Pacific. If Noonan knew that the radio navigation was superior, yet relied on traditional navigation techniques, then he would have committed a lethal intent error.

A case can be made that intent errors are of trivial concern because rarely would an individual knowingly elect a form of control that reduces the chance of success. Surely, a pilot would not rely on manual control if she or he knew that the automated pilot system would be more likely to land the aircraft safely. Beck, Dzindolet, and Pierce (2005) did not test pilots, but they did simulate this situation by providing operators with “advice” from a machine that never made errors. Participants could maximize their success by agreeing with the machine’s never failing answer. Although operators knew that the machine never made errors, results show that 82% of participants did not utilize advice from the machine on all 100 trials. This finding suggested that a high proportion of individuals are susceptible to making intent errors.

After establishing that intent errors, as well as appraisal errors, are commonplace, Beck, Dzindolet, and Pierce (2007) addressed the question of how to control these suboptimal AUDs. The design was a 2 (Feedback: Present, Absent) x 2 (Scenario Training: Present, Absent) x 2 (Machine Performance: Inferior, Superior), where the operator’s AUD was the main dependent variable. The task was composed of 200 trials in which participants could rely on their own performance or an automated device to determine if a soldier was present or absent in a series of briefly displayed photographs.

Feedback was used to control for appraisal errors. In the feedback-present condition, operators received a cumulative total of the errors that they and the machine made. This ensured that the participants knew precisely how many total errors had been committed by the machine and themselves. Therefore, suboptimal AUDs should only ensue if the individual disregarded feedback. In the feedback-absent manipulation, operators were not told whether they were more or less accurate than the machine. In order to make an optimal AUD, these individuals had to avoid both appraisal and intent errors.

A cognitive-behavioral technique, developed by Beck et al. (2007) called scenario training, was employed to decrease intent errors. Scenario training took the operator through the thought processes of an optimal decision maker. For instance, in one of four scenarios, participants were asked whether it would be preferable to rely on investment advice from a stockbroker, the nonautomated option, or a computer program, the automated option. For two of the four scenarios, the optimal AUD was automated control and for the other two scenarios the optimal decision was manual control.

Beck and his colleagues (2007) randomly assigned participants to either an inferior or superior machine condition. In the inferior machine condition, the machine made approximately twice as many errors at the operator. Misuse occurred if the individuals in this condition relied on the machine's performance. In the superior machine condition, the detector made only half as many errors as the operator. The participants committed disuse if they chose to rely on themselves instead of the superior machine.

Because there were very few suboptimal AUDs in the inferior machine conditions, no statistical analyses were performed with these groups. On the other hand, there was a main effect for feedback in the superior machine condition. Fewer suboptimal AUDs were found in the feedback-present groups than the feedback-absent groups. Although there was no main effect for scenario training, there was a statistically significant Feedback x Scenario Training interaction.

No statistical difference was found between operators who received neither feedback nor scenario training and those who had scenario training but no feedback. The most interesting finding was that scenario training reduced the frequency of suboptimal AUDs in groups receiving feedback. Individuals who were given feedback coupled with scenario training made the fewest suboptimal AUDs. The results of this study demonstrated that there are manipulations that mitigate the occurrence of intent errors in human-automation interactions.

An issue of importance is the causes of intent errors. Beck, McKinney, Dzindolet, and Pierce (2009) suggested that an operator's proclivity to depend on oneself or a machine may be influenced by their views of automation. American folklore includes a tale about this critical type of human-machine interaction. John Henry was a legendary railroad man who died in a race with a steam drill. Consequently, responding to automation as a challenger is called a John Henry Effect (Beck et al., 2009)

One manifestation of a John Henry Effect is the reluctance of an operator to rely on automation to perform a task. Not all human-machine interactions are likely to result in a John Henry Effect. For example, most people do not experience anguish or the need to compete with their washing machines. On the other hand, the notion of robotic lovers is an idea that many people find disconcerting. Beck et al. (2009) proposed that an operator's personal investment increases the likelihood of a competitive John Henry Effect. In turn, John Henry Effects should augment reliance on the manual alternative, reducing misuse and increasing disuse.

Beck et al. (2009) tested this hypothesis using a 2 (Operator: Self-reliant, Other-reliant) x 2 (Machine Reliability: Inferior, Superior) x 14 (Block Trials) design. All participants were told that their objective was to obtain as many credit points as possible. Credit points could be exchanged for a monetary reward. For operators in the self-reliant conditions, each trial began with a credit screen in which they elected to base an upcoming credit point on their performance or the performance of a combat identification device (CID). A counter on this screen kept a running total of errors. After a number of trials, feedback eliminated the possibility of suboptimal AUDs resulting from appraisal errors. Operators knew whether the automated option or the manual option was more accurate. The CID made approximately half the mistakes as the operator in the superior reliability condition and twice as many mistakes in the inferior reliability condition.

Other-reliant operators differed from the self-reliant condition by having the choice to rely on the CID or another human participant. The participants were yoked; individuals were matched with respect to sex and the machine reliability variable. For instance, assume a male in the self-reliant condition worked with the inferior machine, clicked fire on trial 25, and saw that the CID held fire. Then, a male in the other-reliant condition who worked with the inferior machine saw that the previous participant fired on trial 25 and that the CID held fire.

Error variances were greater in the superior machine condition than in the inferior machine condition. As a result, the machine reliabilities were separately examined by using two 2 (Operator: Self-reliant, Other-reliant) x 14 (Trial Blocks) ANOVAs. The dependent variable in both analyses was the frequency of suboptimal AUDs. In each analysis, the most important finding was a statistically significant main effect for the operator variable. In the inferior machine condition, self-reliant operators were less likely than other-reliant operators to rely on the machine; they committed fewer suboptimal AUDs.

In the superior machine condition, suboptimal AUDs occurred when the operator chose to rely on a human response rather than automation. Once again, there was a statistically significant main effect for the operator variable. Self-reliant operators made more suboptimal AUDs than other-reliant operators. In comparison to other-reliant participants, self-reliant individuals were more resistant to cumulative feedback.

As posited by Beck et al. (2009), self-reliant operators were less likely to rely on the automated option than other-reliant operators. It is important to note that greater John Henry Effects did not always result in a decline in performance. When automation was the inferior option, relying on human control was the optimal choice. It was when the human control was the suboptimal option that John Henry Effects reduced the probability of task success.

In sum, substantial progress has been made in understanding AUDs. Viewing suboptimal AUDs as recognition, appraisal, and intent errors has proven to be a useful framework (Beck et al., 2002). Although recognition errors and appraisal errors are well-accepted, intent errors remain controversial. Some might be skeptical of the significance of intent errors, contending that they are rare occurrences confined to a few irrational operators. Research by Beck et al. (2007) has indicated that intent errors are not atypical, but are commonplace and need to be addressed. Furthermore, it has been shown that intent errors can be reduced through interventions, such as scenario training (Beck et al., 2007). Personal investment has been established to be one cause of intent errors; however, the research on the causes of intent errors is still in its infancy.

System 1-System 2 Processing

Another strategy for investigating causes of intent errors is to explore the decision-making underpinnings of suboptimal AUDs. Numerous studies (e.g., Evans & Stanovich, 2013; Sloman, 1996) have found that the System 1 (S1)-System 2 (S2) framework (Stanovich & West, 2000) provides a useful conceptual scheme for examining decision-making (see Table 1 for a list of S1-S2 characteristics).

In general, the default for problem solving is S1. This system houses innate skills and learned biases and heuristics. The automaticity associated with S1 generally aids people by allowing them to interact quickly and almost effortlessly with various stimuli (Kahneman, 2011). Consider how quickly you are able to recognize that a yellow light at an intersection indicates the need to slow down your car. Automotive havoc would result if there was a constant need to take time to consider the meaning of a yellow light before responding.

How might you react to an intersection where the traffic lights are not working? This is where S2 takes the lead and allows you to consider options for unexpected situations. S2 is the more deliberate, effortful, and conscious of the two systems (Kahneman, 2011). Although it is not free from the associations and biases of an individual, it allocates more resources to examine a situation in greater depth. Because S2 employs more time and energy than S1, it is most efficient to engage S2 only when the situation justifies the need for additional cognitive resources (Kahneman & Frederick, 2002).

It is important to recognize that the two systems are not independent of one another. Rather, the activity of the two systems ebb and flow depending on the perceived task at hand. Optimal performance is generally attained when both systems interact to provide appropriate behavior. Most individuals experience regular conflict of the two systems, however, and the communication of the systems does not always lead to the most successful response (De Neys & Glumicic, 2008). As AUD research has demonstrated, decision-making errors are common in the human-automation interaction. The swiftness of System 1 may lead to a rash decision with negative consequences, but the contemplation of System 2 may unnecessarily waste cognitive resources.

For instance, the classic Stroop task could be interpreted through the lens of the dual-process framework. In the typical Stroop task, individuals must override semantic processing of a word and only name the color in which the word is printed (Stroop, 1935). Experimenters generally compare a congruent condition, in which the color word matches the ink color (e.g., the color word RED inked in red), and a non-congruent condition, where the word is inked in a different color (e.g., the color word RED inked in blue). Task interference occurs when the individual mistakenly says the printed word rather than the font color. Looking from the angle of the S1-S2 framework, this interference can be viewed as a result of the automatic processing of reading from S1. Even with deliberate effort from S2 to focus on the color, this interference is difficult for most individuals to overcome.

As the interest in the dual-process framework has grown, so have the number of manipulations. Atler, Oppenheimer, Epley, and Eyre (2007) explored how changing the font of a two-statement syllogism may alter the levels of each system used. They proposed that people in the difficult-to-read font (disfluent) condition would engage S2 at higher levels and subsequently answer more questions correctly than those in the easy-to-read font (fluent) condition. Consistent with their hypothesis, participants in the disfluent condition had a higher frequency of correct responses (64%) than those in the fluent condition (43%).

In the aptly named article “How to open the door to System 2,” Bourgeois-Gironde and Vanderhenst (2009) analyzed experimental conditions that may be likely to override System 1 and enlist System 2. In one of their experiments they altered the commonly used stamp and envelope problem by providing possible answer choices for the volunteer to select.

The researchers hypothesized that by listing answers to be evaluated, participants would initiate System 2 processing and therefore reduce error-prone intuition. As predicted, individuals in the evaluation condition had a higher rate for correct answers (49%) compared to those who had no answers to guide their decision-making (31%).

Design

The primary purpose of this thesis was to examine the relationship of S1-S2 thinking and AUDs. To my knowledge, a S1-S2 analysis has not been applied to the study of operators' decisions to rely on automated or manual control. The task was composed of a series of target-detection trials. Participants were informed that the task objective is to distinguish "friendly" from "enemy" helicopters. Credit points, which had a monetary consequence, could be earned by correct decisions.

The design was a 2 (Automation: No-aid, Aid) x 3 (Credit Points: 0, 3, 6). Operators in the aid condition differed from those in the no-aid condition in that they had an error-free machine to rely on for correct responses. That is, they only had to follow the CID's "advice" to avoid committing errors. The credit points variable indicated the amount of credit points operators could gain on a given trial. The dependent variables were the total of incorrect responses and the latency of responses. Latencies were recorded in seconds.

Hypotheses

Errors. A main effect was predicted for the automation variable, with fewer errors occurring in the aid condition than the no-aid condition. Participants in the aid condition should have had a much easier task than those in the no-aid condition. Operators paired with the aid were able to rely almost exclusively on S1 processing. For participants in the aid condition, most disagreements with the error-free CID were intent errors, because they were aware that the machine does not make mistakes.

The proceeding hypothesis assumes that operators paired with aid condition utilized information from the CID. If, however, participants in the aid condition did not use this information, then the only effect from the CID would be to serve as a distraction. Should this be the case, performance will be better in the no-aid condition than the aid condition.

Prior research has indicated that S2 processing tends to increase as a function of task consequence (Kahneman, 2011). This suggests that participants will make fewer errors if the trials are worth 6 credit points than if they are worth 0 or 3 credit points. Thus, a main effect for the credit points variable was anticipated. At present, there is insufficient information to make a prediction regarding a possible interaction.

Latency. Many studies (e.g., Atler et al., 2007; Sloman, 1996) have reported that S2 processing is associated with longer response latencies than S1 processing. In the current study, S1 processing should be more pronounced in the aid condition because these participants had a less cognitively demanding task; they were also able to rely on information from the error-free CID. Therefore, a main effect was predicted for the automation variable with shorter latencies in the aid condition than the no-aid condition. Because S2 processing engages as a function of the consequence magnitude (Kahneman, 2011), latencies were predicted to be longer in the 6 credit points condition than the 0 and 3 credit points conditions.

Application of an S1-S2 model also suggests that the latencies of participants will depend on whether or not they respond correctly on a given trial. It was hypothesized that uncertainty will be positively correlated with the likelihood of errors and that uncertainty will initiate S2 processing (Kahneman, 2011). This analysis suggests that operators should take longer to respond when they answer incorrectly than when they answer correctly. A 2 (Automation: Aid, No-Aid) x 2 (Correctness: Correct, Incorrect) mixed design ANOVA was performed to test these predictions.

Method

Participants

Two hundred three undergraduates (106 females, 97 males) enrolled in psychology classes at Appalachian State University volunteered for course credit. The sample was composed primarily of underclassmen (56%); ages ranged from 18 to 32 years (M = 20.28 years, SD = 2.033). Random assignment to conditions was employed with the stipulation that the aid and no-aid levels contained an equal number of participants, as well as an equal number of females and males. Procedures were approved by an institutional review board (see Appendix A for approval form, 2015) and were treated in accordance with the American Psychological Association's (2010) guidelines of ethical conduct.

Materials

The workstation was an Intel Core 2 Duo E6550, 2.33-GHz central processing unit equipped with 4.00 GB of random-access-memory, a mouse, and a keyboard. Resolution was true-color (32-bit), 1280 x 1024 pixels. An internet housed Visual Basic program presented the slides and recorded responses. The targets were black-and-white photographs of 60 Black Hawk ("friendly") and 60 Hind ("enemy") helicopters previously used by Beck et al. (2009).

Procedure

Participants were randomly assigned to either the no-aid condition or the aid condition. Operators worked individually at cubicle workstations in groups ranging in size from 1 to 10. The arrangement did not allow operators to view one another's screen or responses. They were told that they would respond to a series of target identification trials. The task simulated decisions made by soldiers in deciding to "fire" at an enemy helicopter or "hold fire" in the presence of a friendly helicopter.

Two practice trials were conducted to ensure that the operators understood the instructions. Before the actual trials begin, the experimenter asked the participants if they have any questions (see Appendix B for verbatim instructions). After completing a consent form (see Appendix C), operators began the task.

No-aid condition. Photographs showing complete friendly and enemy helicopters were placed next to the participant's computer for reference (see Figure 1). Operators were informed that they would be working alongside a "perfect machine" using instructions similar to Beck et al. (2005).

> I need to inform you that this machine is perfect. The combat identification device or CID is perfect at detecting the presence and absence of an enemy helicopter. Whenever an enemy helicopter is present, the CID will always correctly state that the enemy is present. Whenever an enemy helicopter is absent, the CID will always state that the enemy is absent. Stated another way, if the CID indicates that the enemy helicopter is absent, it is correct. The CID will never erroneously indicate that the enemy helicopter is absent. Similarly, if the CID indicates that the enemy helicopter

> is present, it is correct. The CID will never erroneously indicate that the enemy is present (p. 213).

At the start of each trial, a target photograph appeared on the screen with a ‘Continue’ button beneath it. Some of the images showed the entire helicopter, whereas most displayed portions of the helicopter. Once the participant viewed the image on the monitors, they would click the ‘Continue’ button to proceed to the next screen (see Figure 2.A.1).

An “Operator” screen then appeared in which the participant was instructed to click a button labeled “Fire” if an enemy helicopter was displayed or “Hold Fire” button if a friendly helicopter was shown (see Figure 2.B). Latencies, measured from the onset of the target photograph to the offset of the operator screen, were recorded for all trials.

Operators were also told that their goal was to earn as many credit points as possible. If the participant was correct in her or his decision-making, they earned the number of credit points shown on the operator screen (see Figure 2.C). To illustrate, if six credit points were available, a text box on the operator screen would state “6 Credit Points will be available on this trial.” The number of credit points (0, 3, and 6) varied as a function of 40 trial blocks; the sequence was counterbalanced. In addition, participants were told that a monetary compensation of $5.00 would be awarded at the end of the task if their total number of credit points was above the median performance of previous operators.

Each trial concluded with a “Results” screen (see Figure 2.D). This screen provided a summary of (1) the helicopter’s identity, (2) the CID’s decision, (3) the operator’s attempted identification, and (4) whether or not they earned credit points. After the operator pressed a button to advance past the results screen, they were presented with another target photograph signaling the start of a new trial.

After the 120th trial, all data were automatically submitted to an SQL database. A screen then directed the participant to raise his or her hand to signal task completion. The experimenter checked to see if the student had earned enough credit points to receive compensation. All participants, regardless of credit points, received class participation credit. Upon completion of the session, students were thanked for volunteering.

Aid condition. The procedure was identical to the no-aid condition except for the image at the beginning of each trial. In this condition, there was a text box displayed above the image of the helicopter that indicated whether the CID has decided to "fire" or "hold fire" (see Figure 2.A.2).

Results

Of the 104 individuals in the aid condition, only 23 (22.1%) agreed with the CID on every trial. It is possible that some incorrect responses were not intent errors. For instance, an operator might inadvertently click the wrong button or they might test to see if in fact the machine was indeed error-free. Presumably, these non-intent errors would not occur repeatedly. Although we cannot be certain that any particular response was an intent error, it is reasonable to assume that those operators in the the aid who made three or more suboptimal AUDs (44.3%) committed errors of intent.

Two 2 (Automation: No-aid, Aid) x 3 (Credit Points 0, 3, 6) mixed designs ANOVAs were performed; alpha set at .05. Automation was a between-subjects variable and credit points was a within-subjects variable. Total incorrect responses and latencies served as dependent measures. The data were screened for outliers using a procedure recommended by Tabachnick and Fidell (2012). This resulted in the elimination of one participant.

It was hypothesized that participants in the aid condition would make fewer errors than those in the no-aid condition. As predicted, a main effect was found for automation (Aid: $M = 4.11$, $SD = 5.86$; No-Aid: $M = 23.99$, $SD = 7.86$), $F(1, 200) = 418.265$, $p < .001$, $\eta_p^2 = .677$. It was also anticipated that participants across conditions of the automation variable would make fewer errors when trials were worth 6 points rather than 0 or 3 credit points. This proposition was not supported, $F(2, 400) = 2.216$, $p = .110$, $\eta_p^2 = .011$. In addition, the Automation x Credit Points interaction was not statistically significant, $F(2, 400) = .900$, $p = .407$, $\eta_p^2 = .004$.

It was predicted that response latencies would differ in the aid and no-aid conditions. In support of that hypothesis, a statistically significant main effect was obtained for automation, $F(1, 200) = 13.674$, $p < .001$, $\eta_p^2 = .064$. Mean latencies were shorter in the aid than the no-aid condition (Aid: $M = 2.92$, $SD = 1.52$; No-Aid: $M = 3.68$, $SD = 1.39$). It was also anticipated that latencies would increase as a function of task consequence. The main effect for credit points was not statistically significant, providing no support for that hypothesis, $F(2, 400) = .881$, $p = .415$, $\eta_p^2 = .004$. Also, the Automation x Credit Points interaction was not statistically significant, $F(2, 400) = .672$, $p = .511$, $\eta_p^2 = .003$.

To further examine intent errors, a 2 (Automation: Aid, No-Aid) x 2 (Correctness: Correct, Incorrect) mixed design ANOVA was conducted. Automation was a between-subjects variable and correctness was a within-subjects variable. The latencies for correct responses and incorrect responses served as the dependent measures. To reduce the effects of selection bias, a filter was applied to include only individuals in the aid condition who made more than 2 errors (44.3%). Because this action removed the better performers in the aid condition, only the bottom 44.3% of operators in the no-aid condition were entered in this analysis.

A statistically significant main effect was found for the correctness variable in that the latencies were longer for incorrect responses ($M = 6.57$, $SD = 5.33$) than correct responses ($M = 3.18$, $SD = 1.28$), $F(1, 87) = 32.813$, $p < .001$, $\eta_p^2 = .270$. A significant main effect was also obtained for automation, $F(1, 87) = 8.303$, $p = .005$, $\eta_p^2 = .087$ (Aid: $M = 5.94$, $SD = 4.88$; No-Aid: $M = 3.82$, $SD = 1.72$). The most interesting finding was a statistically significant Automation x Correctness interaction, $F(1, 87) = 12.777$, $p < .001$, $\eta_p^2 = .128$ (Incorrect Responses: Aid: $M = 8.70$, $SD = 8.28$; No-Aid: $M = 4.44$, $SD = 2.37$; Correct Responses: Aid: $M = 3.17$, $SD = 1.48$; No-Aid: $M = 3.19$, $SD = 1.07$).

Exploratory Analysis

One question that later presented itself was whether or not the operators in the aid condition were following an optimal decision-making procedure. The ideal strategy in the aid condition would be one that produces no errors and minimizes effort. This would be achieved if operators paired with the aid condition ignored the photograph and focused exclusively on the CID's decision. This strategy may be difficult to accomplish because the presence of the photograph may elicit S1 processing to shift attention away from the CID response. That is, individuals may direct their attention to the helicopter image even when the photograph conveys no useful information.

To determine if participants in the aid condition followed the optimal strategy, four additional volunteers in the aid condition were instructed to complete the task by only reading the CID's decision and ignoring the helicopter image. Mean latencies were found to be longer for the 23 participants in the aid condition who made no errors (*M* = 2.70, *SD* = 1.22) compared to the four volunteers (*M* = 1.89, *SD* = 0.16).

Discussion

The purpose of this thesis was to explore the relationship between S1-S2 thinking and intent errors. Participants in the aid condition were informed that they were paired with a machine that made no errors on a target identification task. In order to receive a monetary compensation, individuals paired with the aid only had to agree with the error-free machine on every trial, an S1 processing task. While performing the task, they received trial-by-trial feedback making clear that any disagreement with the CID was an error on their part. On the other hand, operators in the no-aid condition had to rely on their own target detection skills.

It is reasonable to propose that relative to the no aid condition, operators given an aid had an easier task, one that is less likely to require S2 processing. Based on that premise, I made a series of predictions regarding participants' performance on the target detection trials.

Prior research (Beck et al., 2005) led to the expectation that access to an error-free machine would not eliminate misidentifications. This prediction was confirmed. Only 22.1% of participants in the aid condition obtained perfect scores. Even if we allow for the possibility of an accidental response or an initial test of the CID's accuracy, 44.3% of the operators in the aid condition made three or more errors. This finding strongly suggests that among these participants at least some of the incorrect AUDs were errors of intent. It is noteworthy that these operators were aware that misidentifications incurred a monetary cost.

Comparison to other studies (Beck et al., 2007) suggests the frequency of intent errors would be greater if the machine was superior to the operator but not error-free. In future investigations it would be useful to determine how intent errors varied as a function of operator versus machine performance.

As predicted, a main effect of automation was obtained; operators made fewer errors in the aid condition than the no-aid condition. It was also anticipated that participants across conditions would make fewer errors when trials were worth 6 points rather than 0 or 3 credit points. Results did not support this prediction. Neither the main effect for the credit points variable nor the Credit Points x Automation interaction was statistically significant.

There are a number of reasons why the credit points manipulation may have been ineffective in this study. For one, the $5.00 compensation may not have been large enough to elicit a differential response. Another probable reason may be that people are not used to thinking in terms of credit points. Dispensing money for each correct response may have a more powerful effect on AUDs. Of course, carry-over effects may have also obscured any differential impact of credit points.

A statistically significant main effect for the automation variable was also found when latencies served as dependent measures. The aid condition had shorter latencies overall when compared to the no-aid condition. This finding was consistent with the premise that the aid facilitated S1 processing. As was previously noted, S1 processing tends to be faster than S2 processing. With S1 processing engaged, participants did not have to expend cognitive effort beyond agreeing with the CID. Latencies did not differ significantly in regards to credit points.

If uncertainty tends to be higher when operators commit misidentifications, latencies should be longer for incorrect AUDs than correct AUDs. A statistically significant main effect for the correctness variable provided support for this proposition. The most interesting finding was that the difference of latencies of correct and incorrect responses depended on whether the operator was assigned to the aid or the no-aid condition. A statistically significant interaction (see Table 2) illustrates a major finding in that the operator in the aid condition took longer to respond when they made incorrect AUDs than the participants who worked without an aid. This finding demonstrates that in certain circumstances the presence of an automated aid can slow operators' responses.

Although this study focused on incorrect AUDs, intent errors can be manifested in a number of ways. In the aid condition, the optimal strategy, one that would eliminate errors with the least possible effort, would be to always follow the CID's "advice" and ignore the photographs of the helicopters. An exploratory analysis found that the four additional volunteers, who were instructed to only read the CID's response, exhibited faster latencies than participants in the aid condition who made no misidentifications. This result suggests that even operators who made no misidentifications committed intent errors by spending time attending to the photographs of the helicopters.

In this study, operators who responded slowly experienced relatively few negative consequences. However, in other situations longer response latencies could prove catastrophic. For example, taking twenty seconds to decide whether or not to fire could be fatal in a combat scenario. Therefore, an important area for future inquiry would be to examine AUDs in time intensive situations.

Limitations

Consistent with data from S1-S2 research (Kahneman, 2011), latencies are used to gauge the decision-processing time of an individual. Because latency served as a dependent variable in this study, there was no separate manipulation check to indicate a distinction between S1 and S2 processing.

Like every study, questions of external validity must be considered. Because the sample consisted of university students, it is reasonable to consider whether similar results would be obtained with other populations. In addition, the lack of realism in the target-detection task may have lower the consequence magnitude, affecting the latencies and reflection of S1-S2 processing. For example, decision-making in a laboratory cannot simulate the life and death decisions of the battlefield.

Future Research

The results of the present study are consistent with the findings of early investigations (e.g., Beck et al., 2005; 2009) in that intent errors had a detrimental effect on performance. Beck et al. (2007) developed a cognitive-behavioral technique, scenario-training, that when combined with feedback, reduced intent errors. This intervention guided the participant through the thought processes of an optimal performer. Clearly additional procedures need to be developed to mitigate the occurrence of intent as well as appraisal errors.

Additional studies are needed to further explore the relationship between intent errors and response latencies. As previously seen, longer latencies, presumably from greater S2 processing, were related to incorrect responses across conditions. It is necessary to examine more closely the changes in latency, especially when feedback from the machine counters what the participant perceives to be true.

Beck et al. (2009) examined the effects of personal investment on intent errors. Work on this issue needs to be expanded to determine how operator characteristics affect intent errors. Finding the right balance of control between the individual and the machine may be a necessary adjustment for successful human-machine interactions, one that may depend on an operator's personality, attributes of the machine, and work environments.

With that said, it would be worthwhile to investigate how varying the machine's efficiency may impact the operator's response. For example, would an individual be more likely to rely on a machine that had an 80% probability of success versus 100% as seen in this study? Perhaps, we would find that the perfect machine poses too much of a threat to a human's ego for control and a slightly lower probability for success would be less intimidating.

Furthermore, an interesting study would be to explore operators' personal attributions as to why they are committing intent errors. From an outsider's perspective, intent errors may appear irrational. As show in this study, a substantial percentage of operators repeatedly committed intent errors, reducing their chances of obtaining a monetary reward. How do we account for this type of behavior? One possibility is that there are contingences other than performance consequences operating. Despite their illogical appearance, it should not be assumed that intent errors are irrational from the perspective of the operator.

Conclusions

In this study, an S1-S2 framework was reasonably successful in predicting the effects of an automated decision aid on intent errors and response latencies. Usually suboptimal AUDs are attributed to appraisal errors. That is, it is assumed that the operator misjudged the relative accuracy of the automated option versus the manual option. The results of this investigation are consistent with other previous experiments (e.g., Beck et al., 2002; 2005) in demonstrating that this is an overly simplistic model of human-automation interaction. Often operators know how to best respond, yet for reasons we do not fully comprehend, they knowingly impair their performance rather than relying on machines. The findings of this and other studies (e.g., Beck et al., 2007; 2009) will help to elucidate the reasons that operators appear to respond irrationally when interacting with machines that simulate human cognitions.

References

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*, 569-576. doi:10.1037/0096-3445.136.4.569

American Psychological Association. (2010, June 1). Ethical principles of psychologists and code of conduct. Retrieved from http://www.apa.org/ethics/code/index.aspx

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 (Ed.), *Advances in human factors and cognitive engineering: Automation*, (Vol 2, pp. 37-78). Amsterdam, Netherlands: JAI.

Beck, H. P., Dzindolet, M. T., & Pierce, L. G. (2005). Taking the advice of a decision aid: I'd rather be wrong. In *Proceedings of the 49th annual meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA: Human Factors and Ergonomics Society.

Beck, H. P., Dzindolet, M. T., & Pierce, L. G. (2007). Automation usage decisions: Controlling intent and appraisal errors in a target detection task. *Human Factors*, *49*, 429-437. doi:10.1518/001872007X200076

Beck, H. P., McKinney, J., Dzindolet, M. T., & Pierce, L. G. (2009). Effects of human—machine competition on intent errors in a target detection task. *Human Factors*, *51*, 477-486. doi:10.1177/0018720809341746

Bourgeois-Gironde, S., & Vanderhenst, J. B. (2009). How to open the door to System 2: Debiasing the bat and ball problem. In S. Watanabe, A.P Bloisdell, L. Huber, A. Young (Eds.), *Rational animals, irrational humans* (pp. 235-252). Tokyo, Japan: Keio University Press.

Casey, S. (1998). *Set phasers on stun: And other true tales of design, technology, and human error* (2nd ed.). Santa Barbara, CA: Aegean.

Cho, S. D., & Chang, D. R. (2008). Salesperson's innovation resistance and job satisfaction in intra-organizational diffusion of sales force automation technologies: The case of South Korea. *Industrial Marketing Management, 37,* 841-847. doi:10.1016/j.indmarman.2008.04.004

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106*, 1248-1299. doi:10.1016/j.cognition.2007.06.002

Evans, J. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives On Psychological Science*, *8*, 223-241.

Franchimon, F., & Brink, M. (2009). Matching technologies of home automation, robotics, assistance, geriatric telecare and telemedicine. *Gerontechnology, 8,* 88-93. doi:10.4017/gt.2009.08.02.007.00

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York, NY: Cambridge University Press.

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies, 40,* 153-184. doi:10.1006/ijhc.1994.1007

Maehigashi, A., Miwa, K., & Terai, H. (2014). Use of automation and role of trust. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 21,* 100-112.

McBride, S. E., Rogers, W. A., & Fisk, A. D. (2011). Understanding the effect of workload on automation use for younger and older adults. *Human Factors, 53,* 672-686. doi:10.1177/0018720811421909

National Aeronautics and Space Administration. (1995). The First Lunar Landing. Retrieved from https://www.hq.nasa.gov/alsj/a11/a11.landing.html

National Transportation Safety Board (1996). *Annual Report to Congress*. Washington, DC.

National Transportation Safety Board (2006). Collision Between Two Washington Metropolitan Area Transit Authority Trains at the Woodley Park-Zoo/Adams Morgan Station in Washington, D.C., November 3, 2004. *Railroad Accident Report NTSB/RAR-06/01*. Washington, DC.

Parasuraman, R. (2000). Designing automation for human use: Empirical studies and quantitative models. *Ergonomics, 43,* 931-951. doi:10.1080/001401300409125

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52,* 381-410. doi:10.1177/0018720810376055

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*, 230-253. doi:10.1518/001872097778543886

Sheridan, T. B. (2002). *Humans and automation: System design and research issues.* Hoboken, NJ; Santa Monica, California: John Wiley & Sons Inc.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645-665. doi:10.1017/S0140525X00003435

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3-22. doi:10.1037/0033-2909.119.1.3

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662. doi:10.1037/h0054651

Ullmann, J. E. (1966). Automation and the price of progress. *American Journal of Orthopsychiatry, 36*, 70-76. doi:10.1111/j.1939-0025.1966.tb02291.x

Table 1

System 1 and System 2 Characteristics

System 1	System 2
Fast	Slow
High Capacity	Limited Capacity
Biased	Normative
Automatic	Controlled
Associative	Rule-Based

Note. Based on research by Evan and Stanovich (2013) and Kahneman (2011).

Table 2

Latencies as a Function of Automation and Response Correctness

	Aid		No-Aid				
	M	*SD*	*M*	*SD*	*F*	*df*	*p*
Correctness					12.777	(1, 87)	.001
Correct Responses	3.17	1.48	3.19	1.07			
Incorrect Responses	8.70	8.28	4.44	2.37			

Note. Latencies recorded in seconds.

FRIEND **ENEMY**

Figure 1. Reference helicopter photographs located on each participant's workstation.

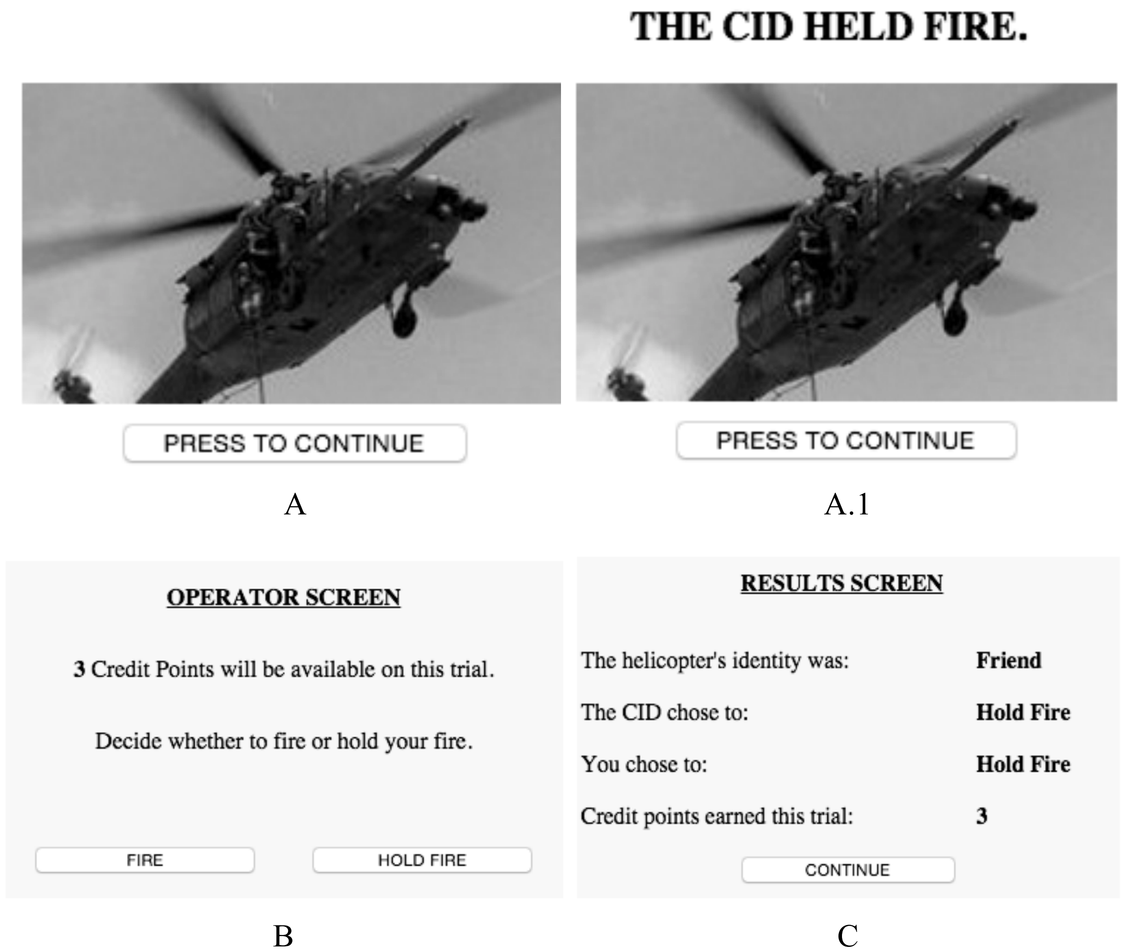

Figure 2. Sequence of screens composing a detection trial: photograph displayed for a no-aid participant (A.1) versus an aid operator (A.2), Operator (B), and Results Screens (C) shown for both conditions.

APPENDIX A

Institutional Review Board Approval Form

Appalachian
STATE UNIVERSITY

INSTITUTIONAL REVIEW BOARD
Office of Research Protections
ASU Box 32068
Boone, NC 28608
828.262.2130
Web site: http://www.orsp.appstate.edu/protections/irb
Email: irb@appstate.edu
Federalwide Assurance (FWA) #00001076

To: Hall Beck
Psychology
EMAIL

From: Dr. Lisa Curtin, Institutional Review Board Chairperson
Date: 2/06/2015
RE: Notice of IRB Approval by Expedited Review (under 45 CFR 46.110)
Study #: 15-0200

Study Title: Human Computer Interaction
Submission Type: Initial
Expedited Category: (7) Research on Group Characteristics or Behavior, or Surveys, Interviews, etc.
Approval Date: 2/06/2015
Expiration Date of Approval: 2/05/2016

The Institutional Review Board (IRB) approved this study for the period indicated above. The IRB found that the research procedures meet the expedited category cited above. IRB approval is limited to the activities described in the IRB approved materials, and extends to the performance of the described activities in the sites identified in the IRB application. In accordance with this approval, IRB findings and approval conditions for the conduct of this research are listed below.

Regulatory and other findings:

The IRB determined that this study involves minimal risk to participants.

Approval Conditions:

Appalachian State University Policies: All individuals engaged in research with human participants are responsible for compliance with the University policies and procedures, and IRB determinations.

Principal Investigator Responsibilities: The PI should review the IRB's list of PI responsibilities. The Principal Investigator (PI), or Faculty Advisor if the PI is a student, is ultimately responsible for ensuring the protection of research participants; conducting sound ethical research that complies with federal regulations, University policy and procedures; and maintaining study records.

Modifications and Addendums: IRB approval must be sought and obtained for any proposed modification or addendum (e.g., a change in procedure, personnel, study location, study instruments) to the IRB approved protocol, and informed consent form before changes may be implemented, unless changes are necessary to eliminate apparent immediate hazards to participants. Changes to eliminate apparent immediate hazards must be reported promptly to the IRB.

Approval Expiration and Continuing Review: The PI is responsible for requesting continuing review in a timely manner and receiving continuing approval for the duration of the research with human participants. Lapses in approval should be avoided to protect the welfare of enrolled participants. If approval expires, all research activities with human participants must cease.

Prompt Reporting of Events: Unanticipated Problems involving risks to participants or others; serious or continuing noncompliance with IRB requirements and determinations; and suspension or termination of IRB approval by an external entity, must be promptly reported to the IRB.

Closing a study: When research procedures with human subjects are completed, please complete the Request for Closure of IRB review form and send it to irb@appstate.edu.

Websites:

1. PI responsibilities:
http://researchprotections.appstate.edu/sites/researchprotections.appstate.edu/files/PI%20Responsibilities.pdf

2. IRB forms: http://researchprotections.appstate.edu/human-subjects/irb-forms

APPENDIX B

Instructions for Target Detection Task

Instructions for the No-Aid Condition

This experiment looks at performance of participants on a target identification task. You will notice that there are pictures of helicopters next to your monitor. Throughout the trials you will be required to identify "friendly" from "enemy" helicopters. On every trial, a combat identification device, also known as a CID, will also make a decision on every helicopter that you are shown. I need to inform you that this machine is perfect. The CID is perfect at detecting the presence and absence of an enemy helicopter. I am now going to go over the instructions on how to complete the task.

The goal of this experiment is to earn as many "credit points" as possible. Some trials will be worth three credit points and some will be worth six credit points. In addition, there will also be experimental trials that will be worth zero credit points. If your performance is above the median of previous players, you will earn $5.00 after the completion of the task. On the other hand, if your performance falls below the median, you will not receive any monetary compensation. Regardless of the total of credit points that you earn, you will receive three ELCs for participation in this study.

The first screen on each trial will contain a photograph of a helicopter. Half of the slides will be of friendly helicopters and half will be of enemy helicopters. Pictures of both a friendly and an enemy helicopter are placed next to your monitor for reference. In this study, there will be one type of friendly helicopter and one type of enemy helicopter. However, the markings and the way that the friendly and enemy helicopters are painted will vary from slide to slide. Some of the photographs you will observe will be of complete helicopters. In other pictures, only part of the helicopter will be shown (i.e., you may see a whole helicopter or just a wheel or propeller).

There are two possible errors that can be made. One error is to fire upon a friendly helicopter. The other type of error is to hold fire when the helicopter was an enemy. Both errors should be avoided. They are equally serious.

After you have been given the opportunity to view a slide of the helicopter, you will click continue to see the operator screen. On this screen you will have the opportunity to make your decision regarding the identity of the previously seen helicopter. If you detect that the helicopter is an enemy, you will click the “Fire” button. However, if you think that the helicopter is friendly, you should click the “Hold Fire” button. In addition, below these two buttons you will see a total number of credit points available on this trial. This will indicate how much a correct decision will be worth.

Once you have made your target detection decision, the “Results” screen will then appear. On this screen you will see 1) the identity of the helicopter, 2) what the CID decided to do, 3) whether you decided to fire or hold fire, and 4) how many credit points you earned, if any, on that trial.

We will begin with two practice trails to be certain you understand these directions. Please ask questions during the practice session. You will not be able to ask questions during the actual session. Once the program begins, no interaction can take place.

Now please answer the questions in front of you. I will talk through the two practice trials. There trials are simply for your understanding. I will tell you what to click and ask you questions throughout these samples to make sure you fully comprehend the task and how to earn credit points. The outcomes of these trials do not count toward the five dollars or affect the number of credit points. Please ask questions as we go.

To start the practice trails, type in the code ‘426’. You will first see the image of the helicopter on the monitor. Notice on your reference sheet that this is a friendly helicopter. Please press the continue button. Now you are seeing the Operator screen. Notice the amount of credit points offered at the top of this screen. The decision on this trial is worth 3 points. Below the credit points, there are the two buttons labeled “fire” and “hold fire”. You would click fire if you thought the helicopter was a what? Enemy correct. You would click hold fire if you thought the helicopter was a what? Friend, correct. Click hold fire. Now you see the Results screen. It says that the helicopter was a friend, the CID decided to hold fire, you decided to hold fire, and you earned 3 credit points. Do you have any questions? Ok, please click continue.

Notice on your reference sheet that this is an enemy helicopter. Now on the Operator screen you can see that this trial is worth 0 credit points. Please click fire. The Results screen indicates that the helicopter was an enemy, the CID decided to fire, you decided to fire, and you earned 0 credit points because there were no credit points available.

Now that we have gone over the practice trials and you understand your task, it is time for me to collect your consent form. Like all participants in psychological investigations, you are guaranteed certain rights. Among these is the right to confidentiality; neither I nor anyone else will know your answers. You also have the right to terminate the experiment at any time.

During the task, please do not click the back button. This will terminate the task and you will lose your chance to earn $5.00. In addition, some pictures will require that you scroll down to click the continue button. After the last trial, please remain seated until further instructions. Are there any questions? Now you may type in the code ‘1990’ and begin.

Instructions for the Aid Condition

This experiment looks at performance of participants on a target identification task. You will notice that there are pictures of helicopters next to your monitor. Throughout the trials you will be required to identify "friendly" from "enemy" helicopters. On every trial, a combat identification device, also known as a CID, will also make a decision on every helicopter that you are shown.

I need to inform you that this machine is perfect. The CID is perfect at detecting the presence and absence of an enemy helicopter. Whenever an enemy helicopter is present, the CID will always correctly state that the enemy is present. Whenever an enemy helicopter is absent, the CID will always state that the enemy is absent. Stated another way, if the CID indicates that the enemy helicopter is absent, it is correct. The CID will never erroneously indicate that the enemy helicopter is absent. Similarly, if the CID indicates that the enemy helicopter is present, it is correct. The CID will never erroneously indicate that the enemy is present. I am now going to go over the instructions on how to complete the task.

The goal of this experiment is to earn as many "credit points" as possible. Some trials will be worth three credit points and some will be worth six credit points. In addition, there will also be experimental trials that will be worth zero credit points. If your performance is above the median of previous players, you will earn $5.00 after the completion of the task. On the other hand, if your performance falls below the median, you will not receive any monetary compensation. Regardless of the total of credit points that you earn, you will receive three ELCs for participation in this study.

The first screen on each trial will contain a photograph of a helicopter. Half of the slides will be of friendly helicopters and half will be of enemy helicopters. Pictures of both a friendly and an enemy helicopter are placed next to your monitor for reference. In this study, there will be one type of friendly helicopter and one type of enemy helicopter. However, the markings and the way that the friendly and enemy helicopters are painted will vary from slide to slide. Some of the photographs you will observe will be of complete helicopters. In other pictures, only part of the helicopter will be shown (i.e., you may see a whole helicopter or just a wheel or propeller). In addition, you will see text on this image that states what the CID has detected. For example, it may say that CID has decided to fire, which would mean that the helicopter has been identified as an enemy helicopter.

There are two possible errors that can be made. One error is to fire upon a friendly helicopter. The other type of error is to hold fire when the helicopter was an enemy. Both errors should be avoided. They are equally serious.

After you have been given the opportunity to view a slide of the helicopter, you will click continue to see the operator screen. On this screen you will have the opportunity to make your decision regarding the identity of the previously seen helicopter. If you detect that the helicopter is an enemy, you will click the “Fire” button. However, if you think that the helicopter is friendly, you should click the “Hold Fire” button. In addition, below these two buttons you will see a total number of credit points available on this trial. This will indicate how much a correct decision will be worth.

Once you have made your target detection decision, the “Results” screen will then appear. On this screen you will see 1) the identity of the helicopter, 2) what the CID decided to do, 3) whether you decided to fire or hold fire, and 4) how many credit points you earned, if any, on that trial.

We will begin with two practice trails to be certain you understand these directions. Please ask questions during the practice session. You will not be able to ask questions during the actual session. Once the program begins, no interaction can take place.

Now please answer the questions in front of you. I will talk through the two practice trials. There trials are simply for your understanding. I will tell you what to click and ask you questions throughout these samples to make sure you fully comprehend the task and how to earn credit points. The outcomes of these trials do not count toward the five dollars or affect the number of credit points. Please ask questions as we go.

To start the practice trails, type in the code ‘426’. You will first see the image of the helicopter on the monitor. Please note that the text says that the CID will hold fire and your reference indicates that this is a friendly helicopter. Please press the continue button. Now you are seeing the Operator screen. Notice the amount of credit points offered at the top of this screen. The decision on this trial is worth 3 points. Below the credit points, there are the two buttons labeled “fire” and “hold fire”. You would click fire if you thought the helicopter was a what? Enemy correct. You would click hold fire if you thought the helicopter was a what? Friend, correct. Click hold fire. Now you see the Results screen. It says that the helicopter was a friend, the CID decided to hold fire, you decided to hold fire, and you earned 3 credit points. Do you have any questions? Ok, please click continue.

Notice the picture is the enemy helicopter on your reference and that the CID decided to fire. Now on the Operator screen you can see that this trial is worth 0 credit points. Please click fire. The Results screen indicates that the helicopter was an enemy, the CID decided to fire, you decided to fire, and you earned 0 credit points because there were no credit points available.

Now that we have gone over the practice trials and you understand your task, it is time for me to collect your consent form. Like all participants in psychological investigations, you are guaranteed certain rights. Among these is the right to confidentiality; neither I nor anyone else will know your answers. You also have the right to terminate the experiment at any time.

During the task, please do not click the back button. This will terminate the task and you will lose your chance to earn $5.00. In addition, some pictures will require that you scroll down to click the continue button. After the last trial, please remain seated until further instructions. Are there any questions? Now you may type in the code '1990' and begin.

APPENDIX C

Consent Form

APPALACHIAN STATE UNIVERSITY

Informed Consent for Participants in Research Projects Involving Human Subjects

Title of Project: Humans Computer Interaction

Investigator(s): Hall P. Beck, Courtney C. Cornelius, Eric Isley & Amanda Osgood

I. Purpose of this Research/Project

The purpose of this investigation is to assess the effectiveness of a simulated combat identification device on a photo recognition task.

II. Procedures

This investigative session involves completing a task on the computer and will require approximately one hour of your time. You will be shown photographs of helicopters on the computer monitor. Your task is to distinguish "friendly" from "enemy" helicopters.

III. Risks

It is extremely unlikely that you will incur psychological, legal, or social harm from your participation in this study. Slight psychological discomfort may be present but the discomforts are minimal. If you feel uncomfortable then you may withdraw at any time without penalty to yourself or a record of your participation. In addition, you may consult the researchers conducting this experiment.

IV. Benefits

Your participation in this study will allow you to gain a greater appreciation for how psychological investigations are conducted. Not only will you study scientific research in the classroom, you will be able to actively learn about it in the laboratory. Society will benefit in that the results of this study may help psychologists better understand how people interact with computers and other forms of technology. Since many jobs are becoming technological or completely automated, this research could have important ramifications.

V. Extent of Anonymity and Confidentiality

Your answers will be saved on a data file along with the information about your age, sex, and any potential military experience. Though you will have the option of signing this consent form, your name will not be associated with or otherwise linked to your responses to task.

VI. Compensation

You can earn up to 3 ELCs for your participation. There are other research options and non-research options for obtaining extra credit or ELCs. One non-research option to receive 1 ELC is to read an article and write a 1-2 page paper summarizing the article and your reaction to the article. More information about this option can be found at: psych.appstate.edu/research. You may also wish to consult your professor to see if other non-research options are available.

In addition, **if** your performance is above the median for your condition, you will receive $5.00.

VII. Freedom to Withdraw

You are free to leave/withdraw from the investigation at any time without penalty.

VIII. Approval of Research

This research project has been approved on February 6, 2015 by the Institutional Review Board (IRB) at Appalachian State University. This approval will expire on February 5, 2016 unless the IRB renews the approval of this research.

IX. Participant's Responsibilities

I voluntarily agree to participate in this study and complete the target identification task. By signing this form, I confirm that I am at least 18 years of age.

X. Participant's Permission

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent.

____Date____________________

Participant's Signature

Should you have any questions about this research or its conduct, you may contact:

Hall P. Beck, Ph. D. **beckhp@appstate. edu**

Investigator E-mail

Questions regarding the protection of human subjects may be addressed to the IRB Administrator, Research Protections, Appalachian State University, Boone, NC 28608 (828) 262-2692, irb@appstate.edu

Vita

Courtney C. Cornelius was born in Jacksonville, Florida, the daughter of Heidi L. Loos and Larry B. Cornelius. After completing her schoolwork at Ardrey Kell High School in June 2008, Courtney enrolled in Appalachian State University. She received her Bachelor of Science in Psychology in May 2012. In the fall of 2013, she entered the Master of Arts degree in General Experimental Psychology program at Appalachian State University.