2009

Automation Dependency Under Time Pressure

Stephen Rice

David Keller

Gayle Hunt

David Trafimow

Repository Citation

Rice, S., Keller, D., Hunt, G., & Trafimow, D. (2009). Automation Dependency Under Time Pressure. *2009 International Symposium on Aviation Psychology*, 611-616.
https://corescholar.libraries.wright.edu/isap_2009/14

AUTOMATION DEPENDENCY UNDER TIME PRESSURE

Stephen Rice, David Keller, Gayle Hunt, David Trafimow
New Mexico State University
Las Cruces, New Mexico

Previous research has identified many factors that affect human dependence on automated systems. Some of these factors include automation reliability, types of errors, and training. This study introduces a new factor, time pressure, which is directly related to operator dependence on automated aids. Participants were asked to perform a simulated UAV target-detection task with the aid of diagnostic automation. Two factors were manipulated in this study: reliability of the automation and time pressure. The results indicate that participants faced with time pressure were more likely to depend on the automation than participants who had more time to evaluate the recommendations. The subsequent increase in dependence due to time pressure was beneficial to overall performance when the automation was highly reliable. In conditions with low reliability, overall human-automation performance suffered due to time pressure. The results imply a potential technique for eradicating the problem of under-dependence on highly reliable automated systems.

The term automation can be described as a mechanism which serves to substitute or enhance human performance. Recent research indicates that while a human-automation team often outperforms a human alone, it rarely measures up to the sole performance of the automation (e.g. Dixon, Wickens, & McCarley, 2007; Dixon & Wickens, 2006; Rice, in press; Rice, Clayton, Wells & Keller, in press; Rice & Hunt, 2009; Rice & Keller, 2009; Rice, Keller, Hunt & Trafimow, 2009; Rice & McCarley, 2008; Rice, Trafimow, Clayton & Hunt, in press). In short, this finding reveals that automation, when left to its own devices, is often more accurate than a human-automation team where the individual is allowed to override the automation. There has been a growing concern in regards to human under-dependence on automation. Specifically, it may be harmful when humans second-guess automation, because they too often disagree with the automation when it is correct, and too often agree with the automation when it is incorrect.

Based on this information, the most intuitive option may be to take the human out of the equation altogether; however, there could be several serious consequences of this action. For example, there are real-world episodes of automation failure (e.g. the auto-pilot in an aircraft fails), and in this type of situation, it is essential that a human operator is present and able to regain control of the system. Although current research indicates that human-automation teams are typically less accurate overall than automation alone, this does not eliminate the possibility that with practice and training, the human-automation team may eventually become *more* accurate than automation alone. Removing the human from this equation would eradicate the possibility of obtaining performance levels exceeding the abilities of the automation. Thus, it is important to investigate any possibility of advancing human-automation performance, rather than doing away with the human factor altogether.

Automation can be divided into four different stages, modeled after human cognitive processing (Parasuraman, Sheridan, & Wickens, 2000). These stages are information acquisition, diagnosis, response selection, and response execution. For this study, our focus will fall primarily with the second stage, or diagnostic automation, which is a common function found in settings such as Unmanned Aerial Vehicles (UAVs). An example of diagnostic automation would include warning alarms and target-detection.

There are several different ways in which humans can interact with diagnostic automation (Parasuraman & Riley, 1997). One of these ways is disuse, which indicates a neglect of the automation. This often occurs when the automation consistently performs poorly (e.g. high false alarm rates), causing the operator to frequently ignore the suggestions made by the aid. This type of interaction may be beneficial in the situation where the human operator is more accurate than a consistently poor and unreliable automation. On the other hand, disuse can be highly dangerous when a human operator's performance is inferior to the automation, leading to sub-optimal results. In this situation it is imperative to correct this disuse behavior.

In reference to an automated aid, dependence is a behavior that is typically mediated by a subjective assessment of trust (a mental state). Trust in automation can be defined as an attitude indicating one's level of confidence that the aid will be successful in helping reach one's goals (Lee & See, 2004).

There are many factors that may cause a shift in trust, which in turn causes a shift in dependence. One major factor is the accuracy of the automated aid. A study by Parasuraman, Molloy & Singh (1993) revealed that when the automation is exceptionally reliable, a human operator may depend too heavily on it and may fail to detect its infrequent errors. Conversely, Dixon and Wickens (2006) found that when the automation is scarcely reliable, a human operator may refrain from any dependence on the aid, to the extent that they may ignore even its correct predictions.

A second factor affecting the level of trust in automation is the type of error the automation makes; that is, if the automation misses a target or makes a false alarm. According to Meyer (2001; 2004), trust and dependence on automation are affected by which of these two errors the automation is more prone to make. Specifically, when the automation is prone to false alarms, lower operator compliance (response when the automation reports an event) is typically the result. When the automation is prone to misses, lower operator reliance (response when the automation does not report an event) is typically the result. A study conducted by Dixon, Wickens and McCarley (2005) found that false alarms affected reliance as well as compliance, but agreed that the two errors affected dependence in different ways. Two additional studies took these findings one step further, demonstrating that the different error types affected different cognitive processes, which lead to different behaviors regarding dependence (Rice, in press).

It is important to note that trust is not the only factor that may influence operator dependence on automation. For example, if an operator is required to complete multiple tasks simultaneously, she may have no choice but to depend more heavily on the automation in order to keep focus on other tasks. In this case, without a change in trust, dependence has still increased.

Another situation where dependence may be affected without a change in trust is when it is of utmost importance to pay attention to warning signals, regardless of how accurate the automation may be. For example, the Traffic Collision Avoidance System (TCAS) is highly prone to false alarms, but in order to avoid the risk of an airplane collision, operators must respond to each and every warning alarm, even knowing that the TCAS has a high false alarm rate.

Figure 1 demonstrates how various factors may contribute to dependence on automated aids. This model is not inclusive of all factors that may affect dependence. It is simply meant to show how some factors may be mediated by trust (e.g. opacity of automation, prior information, and automation reliability) and others may not.

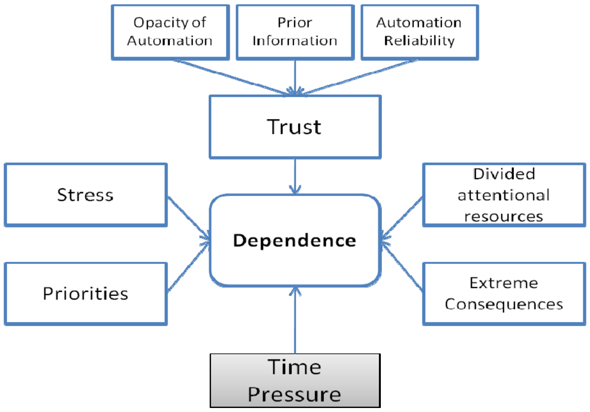

Figure 1. A Model of Automation Dependence.

The purpose of this current study is to introduce another possible factor, time pressure, which may directly influence dependence on automation, while bypassing trust. We believe that when time pressure is introduced (in the absence of outside stressors or multiple tasks), operators will exhibit different behaviors than if there was an abundance of time to make a decision. We predict that those experiencing time pressure will be more dependent on the automation than those with no time pressure. Because the reliability of the automation remains the same across time conditions, there is little reason to suspect that there will be any change in trust, so any difference in behavior should be the direct result of the time manipulation.

Such an outcome may be both beneficial and harmful, depending on the situation at hand. With our manipulations, we plan to demonstrate how automation dependence may greatly influence human-automation performance both for better and for worse.

For this study, two hypotheses are proposed. First, participants given less time to make a decision will depend more highly on the automation than participants given more time to make the same decision. Second, as a result of the increased dependence, participants under a time constraint will perform with greater overall accuracy than their counterparts when the automation is highly reliable, and will have lower overall accuracy when the automation is less reliable.

Methods

260 participants (143 female, 117 male) from New Mexico State University participated in this experiment in exchange for course credit. Participant ages ranged from 17-38, with a mean age of 20.1 (*SD* = 2.63). All participants were screened for normal or corrected-to-normal vision.

The experiment was presented to each participant using E-Prime 1.1 on a Dell computer with a 20” monitor, using 1024 x 768 resolution. E-Prime recorded accuracy and agreement rates for each trial. Images with no target consisted of 50 aerial photographs of Baghdad. Images with a target were created by placing a small tank image onto the 50 aerial photographs using Photoshop CS3. Altogether, there was a total of 100 images to be presented as stimuli—50 with the target present and 50 with the target absent.

An automated aid was used with four different reliability ratings (100%, 95%, 80%, and 65%). Reliability ratings were randomly assigned between subjects and all errors were false alarms. As a control, a condition with no automated assistance was included.

After signing an informed consent, participants were seated 21” from the experiment display with a chinrest controlling head position. Specific instructions were presented on the computer screen and any additional questions were answered by the experimenter. Instructions explained that participants were going to be presented with 100 aerial images of Baghdad and their task was to decide if an enemy tank was present in each image. If a tank was detected, they should respond by pressing the “J” key. If no tank was detected, they should respond by pressing the “F” key. All participants were asked to respond as accurately as possible within their time limits. Participants were informed about the automated aid, given the exact reliability rating, and told that the automation would only err by false alarm. Finally, participants were told that they would have either 2 or 8 seconds to view each image.

Participants were instructed to press any key when they were ready to begin. Each trial began with an automated recommendation stating either “The automation has detected a tank!” or “The automation has determined that there is no tank present!” After 1000 ms, the image was presented. Each image was presented for either 2 seconds (speeded condition) or 8 seconds (unspeeded condition), after which participants made their decision by pressing either the “J” or the “F” key and were then presented with feedback for each trial. Images were presented in random order and each participant viewed all 100 images only once.

Immediately following the completion of the computer portion of the experiment, participants were asked to complete a questionnaire (Dixon & Wickens, 2006) regarding the reliability of the automation and their own trust in the aid.

Including the baseline condition, there were five levels of automation reliability. In addition, there were two time manipulations, resulting in a total of 10 conditions. All subjects were randomly assigned to only one condition in this between-subjects design.

Results

The analyses that follow are separated into two parts: accuracy performance and dependency effects. The percentage of correct trials to total trials was used to measure accuracy. These data can be found in Figure 2. The measure *d’* was not used in this analysis because a very high number of perfect scores were found in the 100% and 95% conditions. However, it should be noted that the effects of *d’* were almost identical to those of accuracy.

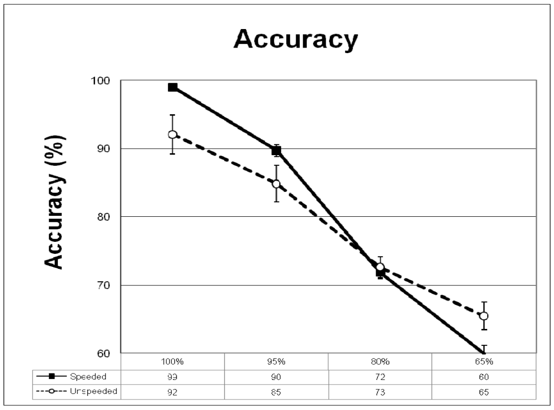

Figure 2. Accuracy Data as a Function of Automation Reliability and Time.SE bars are included.

The first analysis was performed on the Baseline conditions in order to confirm that the given task was challenging enough to warrant the aid of an automated system. It was also important to determine that the time manipulation would cause a tradeoff between speed and accuracy. Indeed, when participants experienced time pressure, their performance suffered tremendously, $t(50) = 4.43$, $p < .001$, $d = 1.25$. In fact, baseline performance in the speeded condition was barely above random performance.

A two-way between-participants ANOVA on the 8 automation conditions, using Reliability and Time as factors, revealed a main effect of Reliability, $F(4, 250) = 161.77$, $p < .001$, $\eta_p^2 = .72$, and no main effect of Time, $F(1, 250) = 1.15$, $p > .10$, $\eta_p^2 = .005$. However, there was an interaction between Reliability and Time, $F(4, 250) = 9.06$, $p < .001$, $\eta_p^2 = .13$, indicating that time pressure was beneficial to general performance at certain reliability levels, but was harmful at other levels, as seen in Figure 3.

Planned comparisons revealed that the 100s (speeded) condition produced higher accuracy than the 100u (unspeeded) condition, $t(50) = 2.43$, $p < .01$, $d = .69$, and the 95s condition produced higher accuracy than the 95u condition, $t(50) = 1.73$, $p < .05$, $d = .49$; however, the 65s condition produced lower accuracy than the 65u condition, $t(50) = 2.32$, $p = .01$, $d = .66$. There was no significant difference between the 80s and 80u conditions, $t(50) = .43$, $p > .10$, $d = .12$. The 95s conditions produced higher accuracy than the 80s condition, $t(50) = 13.77$, $p < .001$, $d = 3.89$, which in turn produced higher accuracy than the 65s condition, $t(50) = 7.65$, $p < .001$, $d = 2.16$.

The rate of participant agreement with the automation was used as a measure of operator dependence, as it is assumed that high dependence on automation indicates high rates of agreement with the automation (Dixon & Wickens, 2006; Dixon, Wickens & McCarley, 2007). It must be noted that high agreement rates to automation hits or correct rejections could possibly be due to high performance both by the automation and the operator independently. In the situation where the automation errs only by false alarm and the operator agrees with this failure, it is only reasonable to assume that this is due to high rates of dependence. For this reason, agreement rates for hits, false alarms, and correct rejections are all analyzed separately.

Planned comparisons revealed that the 100s condition generated higher agreement rates than the 100u condition, $t(50) = 2.51$, $p < .01$, $d = .71$, the 95s condition generated higher agreement rates than the 95u condition, , $t(50) = 1.67$, $p = .05$, $d = .43$. There were no differences between the 80s and 80u conditions, $t(50) = 1.08$, $p > .10$, $d = .31$, or between the 65s and 65u conditions, $t(50) < 1.0$. These data provide some support for the notion that participants under a time constraint tend to have higher agreement rates with the automation.

Planned comparisons revealed that the 100s condition generated higher agreement rates than the 100u condition, $t(50) = 2.33$, $p = .01$, $d = .66$, the 95s condition generated marginally higher agreement rates than the 95u condition, , $t(50) = 1.50$, $p = .07$, $d = .42$. There were no differences between the 80s and 80u conditions, $t(50) < 1.0$, or between the 65s and 65u conditions, $t(50) < 1.0$. These data are consistent with the automation hits data.

Planned comparisons revealed that the 95s condition generated higher agreement rates than the 95u condition, $t(50) = 3.04$, $p < .01$, $d = .86$, the 80s condition generated higher agreement rates than the 80u condition, $t(50) = 2.51$, $p < .01$, $d = .71$, and the 65s condition generated higher agreement rates than the 65u condition, $t(50) = 1.84$, $p < .05$, $d = .52$. These data confirm participants typically have higher rates of agreement with the automation when they are under a time constraint even when the automation is incorrect.

Date collected from the trust questionnaires were not surprising, since participants were told exactly how reliable the automation would be prior to beginning their task. Reliability ratings after the experiment did not differ significantly as a function of time manipulation (all ps > .10). Furthermore, participants' ratings of general trust in the automation also did not differ significantly as a function of time manipulation (all ps > .10). These data provide evidence that levels of trust in the automation were not significantly affected in this study and that the time manipulation did, in fact, affect operator dependence directly.

Discussion

The results of this study have both theoretical and practical implications. As was discussed in the introduction, many factors influence trust, which may have an effect on an operator's level of dependence on automated aids (Dixon & Wickens, 2006; Dixon, Wickens & McCarley, 2007; Parasuraman & Riley, 1997; Rice & McCarley, 2008; Rice, Clayton & McCarley, 2008; Rice, Clayton, Wells & Keller, 2008). This model (see Figure 1) implies that trust (a cognitive process) is a mediator between external factors and dependence on automation (a behavioral response).

There are also many external factors that may affect operator dependence directly, which may also be seen in Figure 1. The current study has introduced a new factor, time pressure, which affects dependence directly and is not mediated in any way by trust.

The results of this study clearly show that when participants were faced with a situation involving time pressure, they were more compliant with the automation than those who had more time to make a decision. While this general finding regarding hits and correct rejections may also be explained by superior performance by both the automation and the operator independently, compliance during trials where the automation produced a false alarm were also higher in the time pressure condition. This finding can only be explained by a higher level of dependence in speeded conditions, regardless of automation reliability.

As suspected, this increased dependence had both positive and negative effects. Positively, when the automation was highly reliable, the overall performance of the human-automation team was improved as a function of the time pressure. The reason for this improvement was clearly due to the added dependence on the automation, which is typically more accurate than the human alone (as discussed in the introduction). The simple addition of a time factor was able to produce significant improvements in human-automation performance.

Negatively, the increased dependence was not specific to the highly reliable conditions. Dependence also increased in conditions with very unreliable automations, leading to poor overall human-automation performance. In these conditions, participants with more time were able to override the automation's suggestions with some degree of confidence and ultimately performed better.

In an effort to ensure that time pressure was not mediated by trust, three deliberate steps were taken. First, participants were clearly told the precise reliability rating of the automation and exactly what type of errors to expect. Second, participants received feedback after each trial in order to allow them to gauge for themselves how accurate the automation really was. Finally, participants were asked to fill out a survey upon completion of the experiment, which ultimately indicated no difference in trust between the speeded and unspeeded trials. Based on this information, it is logical to assume that time pressure, in fact, is not mediated by trust.

There are at least two major practical implications warranted by the results of this study. First, it is essential that designers of automated aids carefully study the environment that their devices will be used in. In an environment with a great deal of time pressure, they must consider that the operator will likely depend highly on the automation, regardless of how reliable it is. If the aid is not highly reliable, a dangerous situation could occur. The opposite is also true. If there is no time pressure, an operator may have ample time to second-guess the automation even if it is highly accurate, in which case a dangerous situation could also occur. Designers must take this into consideration and adjust the environments accordingly as to avoid catastrophic situations.

The second practical application is in the training of operators. When an automation is known to be highly reliable, it may be beneficial to train operators using a time pressure situation in order to increase their dependence on the automation. Should operators learn to depend highly on the automation without constantly questioning its recommendations, the human-automation performance will likely improve. Despite the findings in this study, however, future research should be done in regards to this training technique to discover its long-term effects, especially after the time pressure has been removed. Further, this training method is unlikely to be beneficial in situations where the automated aid is no more reliable than the unaided human. When using low-reliability automated systems, a different training technique should be used to teach operators how to comply most effectively with the automation.

Conclusion

This experiment has shown that the external factor of time pressure causes a higher level dependence on automated aids without affecting operator trust in the automation. It must be highly stressed that this increased dependence is only beneficial when the automation is highly reliable. In this study, performance suffered greatly when the automation had a lower level of reliability. Though previous studies have identified other factors that may increase dependence on automation (e.g. increasing trust), these other factors are very difficult to manipulate in an applied setting. This study has introduced a simple factor that can easily be applied to a situation where operator dependence on automation is necessary and beneficial. To those in aviation and other related fields, it is of utmost importance to be able to detect dangers from an aerial viewpoint very quickly and accurately. Thus, it is essential to further our knowledge and understanding of how operators may triumph over their instinct to second-guess these highly reliable automated systems.

Acknowledgments

The authors wish to thank Amy Wells and Jackie Chavez for their help in collecting data. This research was funded by an Air Force grant (Index #111915). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author.

References

Dixon, S. & Wickens, C. (2006). Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human Factors, 48*(3), 474-486.

Dixon, S. R., Wickens, C. D., & Chang, D. (2005). Mission control of multiple unmanned aerial vehicles: A workload analysis. *Human Factors, 47*(3), 479-487.

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse the misses? *Human Factors. 49*(4), 564-572.

Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, disuse, and abuse. *Human Factors, 39*(2), 230-253.

Parasuraman, R., Sheridan, T.B., & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, 30*(3), 286-297.

Rice, S. (in press). Examining single and multiple-process theories of trust in automation. *Journal of General Psychology*.

Rice, S., Clayton, K. & McCarley, J. (in press). The Effects of Automation Bias on Operator Compliance and Reliance. *Human Factors Issues in Combat Identification*.

Rice, S., Clayton, K., Wells, A. & Keller, D. (2008). Manipulating Trust Behaviors in a Combat Identification Task. *Proceedings of the 2008 Human Factors in Combat Identification Workshop.*

Rice, S. & Keller, D. (2009). Determining reliability of multiple automation aids using system-wide trust: A comparison of four hypotheses. *Proceedings of the 5th International Conference on Technology, Knowledge and Society*.

Rice, S., Keller, D., Hunt. G. & Trafimow, D. (2009). The effects of time pressure on automation dependency. *Proceedings of the 15th Annual International Symposium of Aviation Psychology*.

Rice, S. & McCarley, J. (2008). The Effects of Automation Bias and Saliency on Operator Trust. *International Congress of Psychology*.

Rice, S. & Hunt, G. (2009). Proving a multiple-process model of trust in automation. *Proceedings of the 15th Annual International Symposium of Aviation Psychology*.

Wickens, C. D. & Hollands, J. G. (2000). Engineering Psychology and Human Performance, 3rd Edition. Upper Saddle River, NJ: Prentice Hall.