Is Ignorance Bliss? Role of Credibility Information and System Reliability on User Trust in Emergent Technologies

Poornima Madhavan

Department of Psychology
Old Dominion University
Norfolk, VA 23529-0001, USA

ABSTRACT

Understanding the relationship between the credibility of an emergent system and users' perceptions would improve our knowledge of human interaction with emergent technologies in the present day. The purpose of this research was to examine how much background information about an emergent system must be given to naïve users to engender appropriate trust and utilization. Participants performed a simulated airline luggage-screening task with the assistance of an emergent system that was 70% reliable on one half of the task and 90% reliable on the other half. Participants were assigned to one of two groups: (1) no information (participants were merely told they would receive the assistance of an automated system with no additional details), and (2) information (participants were provided with background information about the system's functions, including the fact that it was a recently developed emergent system whose credibility had yet to be established). Results revealed that background information led to higher trust and better utilization than no information, even when the system was portrayed as emergent without established credibility. Regardless of the actual level of system expertise, information about system functions engendered more appropriate utilization.

Keywords: Emergent Technology, Automation, Trust, Credibility, Reliability, Luggage Screening

INTRODUCTION

Emergent (or "emerging") technologies refer to technical innovations which represent progressive developments, contemporary advances and innovation in a variety of fields for competitive advantage. Specifically, emergent technologies denote technology developments that broach new territory in some significant way in their field. Some examples of currently emerging technologies include advances in cognitive science, biotechnology, robotics, information technology and artificial intelligence. Therefore, it is reasonable to conclude that the primary purpose of such emergent technologies is to assist and improve human performance in a variety of domains over and beyond the current state.

Despite continuous attempts at progressively improving the design of technology and automated aids as stated above, research has extensively documented that automated aids are seldom 100% accurate; they are prone to errors, as are their human counterparts. Specifically, human operators typically display different reactions to automated aids such as: disuse, misuse, a lack of critical consideration (deVries, Midden and Bouwhis, 2003), as well as correct use. Disuse, usually attributed to lack of trust, describes a situation in which a reliable aid is under-utilized. The opposite occurs in situations of too much trust (or over-trust), wherein individuals may rely uncritically on the system, becoming oblivious to errors. Thereby, misuse occurs when individuals over-depend on an aid. Correct use describes the situation in which an accurate allocation of reliance is delineated to an aid such that reliable aids are utilized and unreliable aids are not (Younho, Bisantz and Gattie, 2006).

As alluded to in the above paragraph, research has revealed that appropriate use of automation is influenced strongly by operators' trust in an automated aid (e.g. Dijkstra, 1999). Researchers have found that initial trust in automation tends to be high due to the existence of a bias toward automation or a "perfect automation schema" (Dzindolet, Pierce, Beck, Dawe and Anderson, 2001). However, this positive bias toward automation leads operators to be very sensitive to the errors made by automation, leading to a sharp decline in trust and dependence when machines generate errors (Dzindolet et al., 2001). Furthermore, Madhavan and Wiegmann (2007) found that operators using newly developed automated aids (framed as "novices") were likely more sensitive to 'easy' errors (i.e., instances in which the automated aid missed a target but the operator easily detected it) due to operators' prematurely low expectations of newly developed technologies. Such expectancy-driven judgments possibly led to a greater degree of miscalibration of trust relative to those receiving the assistance of tried and true "expert" systems.

Clearly, there is evidence to indicate that trust plays a significant role in the utilization of technology that is, in turn, influenced by preconceived notions about technology. It is also evident that utilization of automation is influenced not just by subjective trust of the user but also by the objective reliability (or accuracy) of the system in any particular task domain. No research, however, has examined the role of trust in human interaction with emergent technologies. The purpose of this research is to examine the development of human trust in emergent technologies in one particular domain – airline hand-luggage screening.

Emergent Technologies in the Domain of Airline Hand-Luggage Screening

Machine vision is a form of emergent technology that refers to the sum total of the techniques and methods used to provide imaging-based automatic inspection and analysis for such applications as automatic inspection, process control, and robot guidance in industry. In the context of airline hand-luggage screening, machine vision typically applies pattern recognition algorithms to perform template matching (or matching of an object in the luggage with a previously encountered object), finding, matching and/or counting of specific patterns, and localization of objects that may be rotated, partially occluded by other objects or vary greatly in size.

Image enhancement is very important for increasing the sensitivity of luggage screening performance at airports (Singh and Singh, 2005). Modern baggage x-ray systems are designed with inbuilt image manipulation options; however, most screeners are not conversant enough in the intricacies of their functionality to allow them to select the best or best combination of technologies to assist them in their visual search (Singh and Singh, 2005). This problem is exacerbated by the time pressure inherent in the screening task. In most cases, screeners simply enhance the luggage image once using a tried and tested technique and base their decision of this. Since the screener is expected to decide whether to stop or pass a bag within a 4-6 second window, most of the technological capabilities available to them never get used. Although the time pressure characteristic of the screening task is a major challenge, there are several questions about screeners' degree of trust in the screening technologies available to them as well. Trust is a particularly critical issue when screeners interact with novel machine-vision-based detection systems that are relatively new (or "emergent").

Trust in Emergent Luggage Screening Technologies

Presumably, trust in a particular type of screening technology would be some function of the amount of prior knowledge the screener has about the system and its capabilities vis-a-vis the screener's own ability to perform the task unaided. If such emergent systems generate "easy" errors that a screener would conceivably not have made (unaided), the screener is much less likely to trust the automated system. This has been found to be the case even when the automated system is overall statistically more accurate than an unaided human operator (Madhavan and Wiegmann, 2007). Additionally, people typically demonstrate an inclination to trust themselves over an automated aid; consequently, errors on the automated aid's part can greatly reduce trust (deVries et al., 2003) However, the negative impact of errors on trust is sometimes mitigated if an explanation regarding system functions is provided, even in circumstances where the change in trust levels is not necessarily appropriate or warranted (Dzindolet et al., 2003).

Explanations regarding system functions typically focus on the 'credibility' of the system. 'Credibility' refers to the judgment that a message and/or its source are believable and convincing (Burgoon, Bonito, Bengston, Cederberg, Lundeberg and Allspach, 2000). Corritore, Kracher and Wiedenbeck (2003) defined expertise as typified by "knowledge, experience, and competence", with an intricate relationship between trust and expertise – specifically, in the context of human-automation interaction trust is based mostly on users' perceptions of the perceived expertise of the machine.

Logically, automated systems that are perceived as 'expert' are more likely to be trusted than systems that are not. This has been found to be the case for health websites (Sillence, Briggs, Harris and Fishwick, 2006, 2007) as well as informational and transaction websites (Corriotore et al., 2003). In the context of human interaction with automation, some research has found that human partners are judged to be more credible than computer partners (Burgoon et al., 2000). However, this finding is not all pervasive. Dijkstra, Liebrand and Timminga (1998) found that if users are told that a system is credible it significantly influences their evaluation of the system's diagnoses. This might explain why users sometimes neglect the incompetence of seemingly 'expert' systems. Users, at least initially, tend to be influenced significantly by the surface characteristics of a piece of information (Burgoon et al., 2000; Corriotore et al., 2003, Sillence et al., 2006, 2007; Madhavan and Wiegmann, 2007) which can ultimately lead to human-automation interaction that is based on more on presumed credibility rather than actual expertise.

Purpose of the Present Study

Understanding the influence of perceived system credibility would help improve understanding of human interaction with emergent technologies. However, few attempts have been made to determine exactly how much background information must be given to naïve users regarding system functionality and design specifications to engender appropriate trust in and utilization of these systems. This is particularly important when the system is an emergent system whose credibility or reliability has not been widely documented.

It is possible that providing such information would improve the perceived credibility of the system and engender calibrated trust. Conversely, it is also possible that presenting a system as "new" or as one whose credibility has not been widely documented or established might negatively bias users leading to greater miscalibrations of trust and inappropriate dependence on these systems. The question arises as to whether background information would help or hurt system utilization in the case of recently developed emergent systems with relatively unknown credibility. The purpose of this research was to examine the relationship between background information (provided to users of novel emergent systems) and actual system reliability on trust and utilization. In other words, "is ignorance bliss" in terms of lesser information leading to fewer degrees of bias and consequently better calibration of trust in emergent systems?

We conducted a simulated airline luggage screening task wherein participants were assisted by an emergent system of varying reliability levels. In addition, we manipulated the amount of background information about the system that was given to participants, with some participants receiving no information and other receiving information regarding its emergent status. We examined the relationship between system reliability and background information on trust and dependence on the automated system. We hypothesized that when background information was provided about the system, individuals would (1) trust the system more, and, (2) depend on the system more, despite the system being portrayed as an emergent system whose functional credibility and reliability are not as yet established. We hypothesized that some background information would lead to more calibrated trust and higher levels of dependence on the system than no information at all since background information will provide the user a higher degree of perceived control over the task and system functions, even when the information does not directly address the system's credibility or reliability.

METHOD

Participants

Forty Old Dominion University undergraduate students from introductory psychology courses volunteered to participate in order to partially fulfill course requirements.

Procedure

Participants were asked to complete a computer simulation in which they played the role of airline luggage screeners. They were presented with 400 x-ray images of passenger luggage on a 19'' color monitor placed approximately 17'' from the edge of the desk. The computer simulation was developed using Visual Basic for Windows and presented the image, diagnosis of the automated screening system, the opportunity for participant input, and feedback. The x-ray images were created using Adobe Photoshop and were comparably cluttered with everyday items (toys, clothes, personal accessories, etc.). A subset of 20% of the images had one of eight possible knife images digitally superimposed. The task was for participants to detect these knives among the distractors in each luggage image within a limited time window. A sample x-ray image with an embedded knife is presented in Figure 1.

Figure 1. Sample x-ray image with embedded knife.

Participants were assigned to one of two groups depending on the background information they received about the automated system: (1) no information (participants were merely told that they would receive the assistance of an automated aid with no additional system details), and (2) information (participants were provided with details about the system development, and functions, and were also told that it was a recently developed emergent system whose credibility was not yet established). This information was based on the description of an automated luggage screening system used in an earlier study by Madhavan and Wiegmann (2007).

In the actual screening task, the x-ray images were grouped into two trial blocks, with each block comprising 200 trials (images). At the beginning of each trial, an x-ray image of luggage appeared on the computer screen for a duration of 4 seconds. After the trial timed out, participants received the diagnosis of the automated system regarding the presence or absence of a knife, in the form of a text message on the screen. Participants then input their own diagnosis as to whether a knife was present or not and received textual feedback on the accuracy of their diagnosis.

The automated system was, unbeknownst to participants, 70% reliable on one half of the trials and 90% reliable on the other half. The reliability was counterbalanced, with one half of the participants receiving the 70% reliable system in the first block and 90% reliable system in the second block, and vice versa. At the end of each block, participants completed a 12-item System Trust Scale (Jian, Bisantz, and Drury, 2000) to determine their trust in the automated system. Negative items in the questionnaire were reverse scored.

Since we were primarily interested in participants' subjective trust in and objective utilization of the system during the course of the screening task, the dependent variables we analyzed were (1) trust in the system, and (2) dependence on (agreement with) the system's diagnoses.

RESULTS

The data were analyzed using multiple repeated measures ANOVAS and post-hoc tests. All results with alpha value below .05 are reported as statistically significant.

Trust in the Automated Emergent System

The results for automation trust are illustrated in Figure 2. A 2 (information: no vs. yes) X 2 (reliability: 70% vs.90%) mixed ANOVA on subjective trust revealed significant main effects for information, F (1, 39) = 7.93, *p* = .008 and reliability, (1, 39) = 8.46, *p* = .006, but no interaction between the two, F (1, 39) = .72, *p* = .40. Trust was higher when the system was 90% reliable (M = 6.72, SD = .31) than when the system was 70% reliable (M = 5.77, SD = .47). As seen in Figure 2 and in keeping with hypotheses, information about the system led to higher levels of system trust (M = 6.90, SD = .39) than no information (M = 5.59, SD = .40), despite the fact that the given information portrayed the system as a newly developed system.

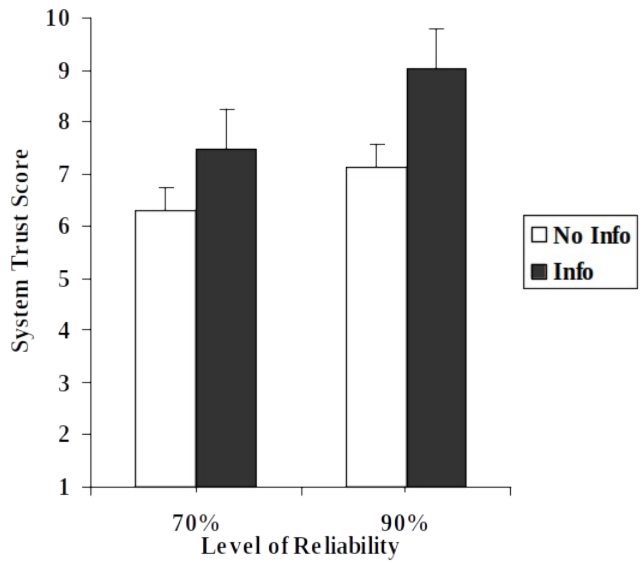

Figure 2. System trust as a function of system reliability and background information.

Dependence on the Emergent System

We used two indices of dependence: (1) compliance, which was computed as participants' agreement with the system when it said "target present", and (2) reliance, which was computed as participants' agreement with the system when it said "target absent".

Compliance. The results of the analyses for compliance are illustrated in Figure 3. A 2 (information: no vs. yes) X 2 (reliability: 70% vs. 90%) mixed ANOVA on compliance scores revealed significant main effects for information, F (1, 39) = 3.12, *p* = .07 and reliability, (1, 39) = 4.34, *p* = .032, but no interaction between the two, F (1, 39) = 2.04, *p* = .162. Compliance was higher when the system was 90% reliable (M = .65, SD = .04) relative to 70% reliable (M = .57, SD = .053). Again, in support of the hypotheses and as can be seen in Figure 3, participants who received information about the system complied with the system significantly more (M = .62, SD = .047) than those who received no information at all (M = .55, SD = .047).

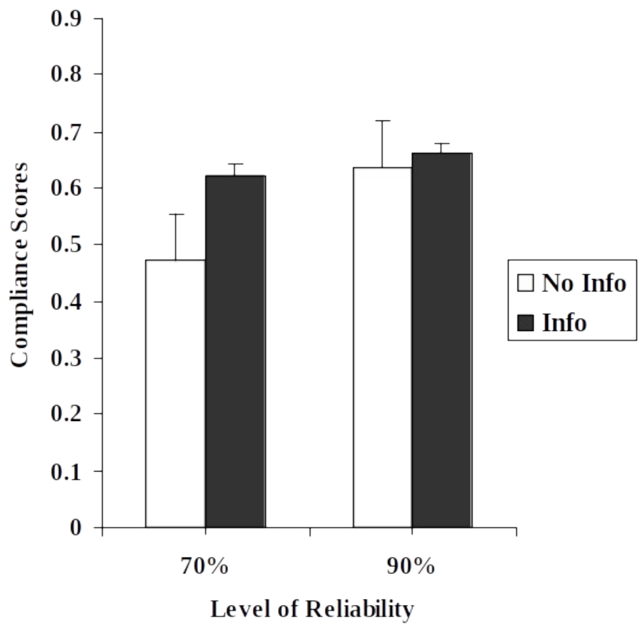

Figure 3. Compliance with the aid as a function of system reliability and background information.

Reliance. The analyses for reliance are illustrated in Figure 4. A similar 2 (information: no vs. yes) X 2 (reliability: 70% vs. 90%) mixed ANOVA on reliance scores revealed significant main effects for information, F (1, 39) = 5.52, p = .024 and reliability, (1, 39) = 55.89, p = .0001, as well as a significant interaction between the two, F (1, 39) = 4.69, p = .037. As can be seen in Figure 3, reliance was higher on trials when the system was 90% reliable (M = .85, SD = .019) relative to 70% reliable (M = .68, SD = .029). Moreover, participants who received information relied on the system significantly more (M = .74, SD = .028) than those who received no information (M = .63, SD = .029) when the system was 70% reliable. As indicated by the significant interaction, there were no significant differences in reliance as a function of background information when the system was 90% reliable.

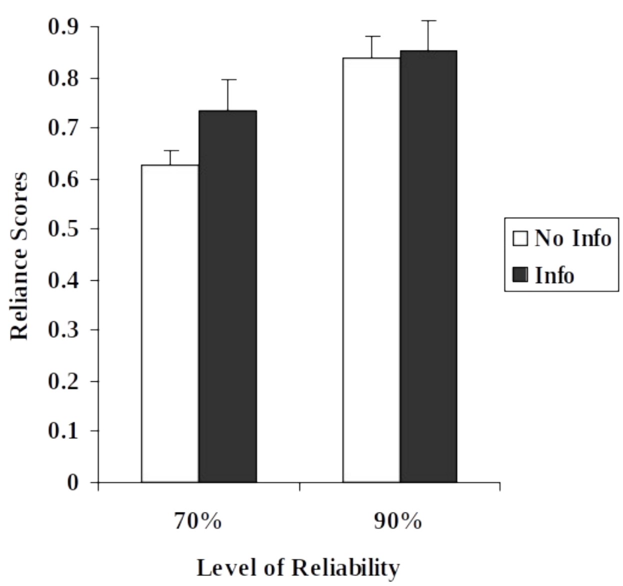

Figure 4. Reliance on the aid as a function of system reliability and background information.

DISCUSSION

The results of this research indicate that subjective trust and dependence on the emergent system were always higher when the aid was 90% reliable compared to 70% reliable, regardless of whether background information was provided or not. This suggests that the situational accuracy of the automated system had a stronger impact on trust than expertise information per se.

Within reliability levels, however, participants with background information about the system subjectively trusted the system more than those with no information, even when the information portrayed the aid as an emergent system whose credibility was not yet established. This finding can be explained by the Elaboration Likelihood Model of persuasion (Petty and Cacioppo, 1986). According to this model, the higher levels of trust that were engendered by providing background information about the system suggest that information is an important factor that influences operator trust in an emergent system. This is possibly because being aware of system features provided users with a higher degree of perceived control and understanding of system functions. This information was important even when the content did not necessarily include any specific information about system reliability or credibility.

This pattern was also observed for dependence on the system. Similar to the pattern for trust, participants who received no information about the system depended on it less than those who did receive additional information, even though the system was portrayed as one for which credibility was not established. This pattern was more salient when the system was less reliable (i.e., 70%) since the accuracy of the 90%-reliable system likely compensated for credibility information.

The results for trust and dependence observed in this study contradict the earlier findings of Dijkstra (1999), who found that when users do not have sufficient evidence to judge the advice of an automated system they are more likely to agree with rather than refute system advice. This is largely because most people implicitly associate a high level of credibility with automation, especially when they do not have sufficient background information about it (Dijkstra, 1995). This formed the basic premise for our question of whether “ignorance (about the system) is bliss” and would potentially lead to better trust and utilization than providing system specific information that presents the system as emergent.

Clearly, the results of this study refuted the above findings of Dijkstra. Contrary to existing research, providing participants no information about the system led them to evaluate the system less favorably, both subjectively and objectively. This suggests that when users have access to additional information about automated systems, they tend to use this as a cognitive anchor (Madhavan and Wiegmann, 2005) or a yardstick to judge the efficacy of the aid during the course of a task. Even when the system was presented as a recently developed system without established credibility, system specific information positively influenced both trust and dependence on the system relative to no information at all. Regardless of the actual credibility associated with the system, background information about system functions engendered more appropriate utilization than no information, possibly because of fewer biases and misconceptions regarding system capabilities.

The current study therefore indicates that ignorance is “not bliss” in situations that require humans to interact with emergent systems. Despite the lack of information about the system’s apparent credibility, providing users with practical information about system functions does help users calibrate their trust in and dependence on the system better then when they have to guess the system’s capabilities.

CONCLUSIONS

Based on the results of the present study, a potential suggestion to the designers of emergent systems is to design systems that elicit appropriate compliance or reliance strategies so as to reduce the occurrence of a ‘costly’ miss or a false alarm that might result when users interact with an unfamiliar system. An alternative, and perhaps more challenging suggestion, is to derive the “optimal” reliability level that would provide positive assistance that will help users ‘understand’ the system if sufficient background information were inaccessible.

As computer-based emergent systems are increasingly being incorporated in airline hand-luggage screening, more and more critical decisions are being influenced by these systems. From the results of this study, it follows that internal attributes of systems when combined with extraneous information on the system’s apparent credibility evidently has a stronger effect on system trust and dependence than presenting the system alone without background information. Further research needs to be conducted to delve deeper into the issue of how much knowledge is appropriate knowledge and to understand the fine line between providing 'appropriate background information' and 'damaging background information' to users of emergent technologies.

REFERENCES

Burgoon, J.K., Bonito, J.A., Bengtsson, B., Cederberg, C., Lundeberg, M., & Allspach, L. (2000). Interactivity in human-computer interaction: A study of credibility, understanding, and influence. *Computers in Human Behavior, 16*, 553-574.

Corritore, C.L., Kracher, B., Wiedenbeck, S. On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies, 58*, 737-758.

de Vries, P., Midden, C., Bouwhuis, D. (2003). The effects of errors on system trust, self-confidence, and the allocation of control in route planning. *International Journal of Human-Computer Studies, 58*, 719-735.

Dijkstra, J.J. (1995). The influence of an expert system on the user's view: How to fool a Lawyer, *New Review of Applied Expert Systems, 1*, 123-138.

Dijkstra, J.J. (1999). User agreement with incorrect expert system advice. *Behaviour and Information Technology, 18*, 399-411.

Dijkstra, J.J., Liebrand, W.B.G., Timminga, E. (1998). Persuasiveness of expert systems. *Behaviour and Information Technology, 17*, 155-163.

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P. (2003). The role of trust in automation reliance. *International Journal of Human Computer Studies, 58*,697-718.

Jian, J.Y., Bisantz, A.M., Drury, C.G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*, 53 - 71.

Madhavan, P., Wiegmann, D.A. (2007). Similarities and differences between human-human and human-automation trust: An integrative review. *Theoretical Issues in Ergonomics Science, 8*, 277-301.

Madhavan, P., Wiegmann, D.A. (2005). Cognitive anchoring on self-generated decisions reduces operator reliance on automated diagnostic aids. *Human Factors, 47*, 332-341.

Petty, R.E., Cacioppo, J.T. (1986). The Elaboration Likelihood Model of persuasion. In L. Berkowitz, Ed., *Advances in Experimental Social Psychology, vol. 19*, pp.123-205. New York: Academic Press.

Sillence, E., Briggs, P., Harris, P., Fishwick, L. (2006). A framework for understand trust factors in web-based health advice. *International Journal of Human-Computer Studies, 64*, 697-713.

Sillence, E., Briggs, P., Harris, P., Fishwick, L. (2007). Health Websites that people can trust-the case of hypertension. *Interacting with Computers, 19*, 32-42.

Singh, M., Singh, S. (2005). Image enhancement optimization for hand-luggage screening at airports. *Pattern recognition and Image Analysis, 3687*, 1-10.

Younho, S., Bisantz, A.M., Gattie, G.J. (2006). Trust, decision aiding, and feedback: An integrated approach. In A. Kirlik, Ed., *Adaptive perspectives on human-technology interaction: Methods and models for cognitive engineering and human-computer interaction*, pp.105-113. New York: Oxford University Press.