

Current Concepts and Trends in Human-Automation Interaction

Session Chair: Ernesto A. Bustamante, *University of Idaho*

Session Co-Chair: Poornima Madhavan, *Old Dominion University*

PANELISTS

Christopher D. Wickens, *Alion Science & Technology, Inc.*

Raja Parasuraman, *George Mason University*

Dietrich Manzey & J. Elin Bahner-Heyne, *Berlin Institute of Technology*

Joachim Meyer, *Ben Gurion University of the Negev*

James P. Bliss, *Old Dominion University*

John D. Lee, *University of Iowa*

Stephen Rice, *New Mexico State University*

The purpose of this panel was to provide a general overview and discussion of some of the most current and controversial concepts and trends in human-automation interaction. The panel was composed of eight researchers and practitioners. The panelists are well-known experts in the area and offered differing views on a variety of different human-automation topics. The range of concepts and trends discussed in this panel include: general taxonomies regarding stages and levels of automation and function allocation, individualized adaptive automation, automation-induced complacency, economic rationality and the use of automation, the potential utility of false alarms, the influence of different types of false alarms on trust and reliance, and a system-wide theory of trust in multiple automated aids.

SUMMARY

Human-automation interaction is a very complex domain that encompasses several concepts and trends. Most of these concepts and trends have evolved through the development of new theories and the support of recent empirical findings. Christopher D. Wickens presented an underlying framework of the stages and levels of automation and its use for function allocation. Raja Parasuraman discussed the advantages of individualized vs. group based adaptive automation. Dietrich Manzey and J. Elin Bahner-Heyne addressed the controversy surrounding the conceptualization of automation-induced complacency. Joachim Meyer discussed the notion of economic rationality as it applies to the use of automation. James P. Bliss emphasized the potential beneficial applications of false alarms for high-consequence task environments. John D. Lee presented a framework regarding the different types of automation failures and their potential influence on trust and reliance. Stephen Rice developed a

theory of a system-wide trust in multiple automated aids. The compilation of all these contributions was intended to raise awareness and promote discussion of the current concepts and trends in human-automation interaction.

ABSTRACTS

Automation Stages and Levels Taxonomy and Function Allocation (Christopher D. Wickens)

The stages and levels taxonomy of human-automation interaction (Parasuraman, Sheridan & Wickens, 2000, 2008) has been presented as a tool which we believe is of value for establishing optimal function allocation between human and machine. It enables consideration of this allocation, not as an all-or-none process as in the traditional "Fitts List," but rather as a matter of degree. According to this framework, increasing levels and later stages are progressively more problematic, as the negative consequences of imperfect automation increase (i.e., automation "failures"). Reasons for

this prescriptive assertion are presented, and are illustrated with an example that contrasts synthetic vision displays (earlier stage automation) with highway in the sky displays (later stage automation). The ubiquity of the early vs. late contrast is illustrated with examples from medicine, aviation, and highway safety.

Individualized Adaptive Automation (Raja Parasuraman)

A growing body of evidence points to the efficacy of adaptive automation for supporting human-system performance (Inagaki, 2003; Parasuraman, 2000; Scerbo, 2007). For the most part, previous investigations have focused on implementing adaptive automation at the *group* level (Parasuraman & Wickens, 2008; for exceptions, see Parasuraman, Mouloua, & Molloy, 1996; Wilson & Russell, 2007). However, given that one size may not always fit all, system adaptation at the *individual* level arguably could yield greater benefits. I describe supporting evidence from recent studies examining the effects of individualized adaptive automation for supporting human supervisory control of multiple unmanned vehicles (Parasuraman et al., 2009). Adaptive aiding linked to assessment of individual levels of situation awareness led to superior overall performance compared to static automation or group-based adaptive automation. Individualized adaptive automation may also help resolve the debate between system-driven and user-driven adaptation by pointing to the combined benefits of both.

Automation-Induced Complacency: A Controversial Concept of Human-Automation Interaction (Dietrich Manzey & J. Elin Bahner-Heyne)

The concept of “automation-induced complacency” has been introduced in order to describe a certain risk involved in human-automation interaction, i.e. an uncritical reliance on its proper function resulting from an inappropriate high trust in the system’s reliability. On the behavioral level this effect is reflected in an

inappropriate monitoring of automated functions which directly involves the risk of missing automation failures. Originally complacency has been identified to represent an issue of monitoring automated dynamic processes in the area of cockpit automation (Billings, 1976; Parasuraman Molloy, & Singh, 1993). However, complacency-like effects can also emerge in other fields of human-automation interaction, like interaction with alarms or decision-aids. For example, Mosier & Skitka (1996) have described two types of errors which may occur in interaction with decision-aids. The first kind of error (error of omission) is reflected in failures to detect and respond to critical system states if they are not indicated by an automated aid. The second one includes errors of commission, i.e. accepting recommendation of the automated aid without any attempt to verify them against other available information or even despite contradicting information is available from other sources (e.g. displays). It is obvious that both kinds of errors reflect some kind of over-reliance on automation with errors of omission and the first type of commission errors related essentially to the same issue that has been termed complacency in supervisory process control.

From the very beginning of its introduction, complacency has been a controversial concept (see for example the current discussion of Dekker & Hollnagel, 2004 vs. Parasuraman, Sheridan, & Wickens, 2000, 2008). Three different points of criticism have been raised. The first one challenges the usual operational definition of complacency in terms of missing of automation failures and raises the question how complacency can be distinguished from a rational adaptive strategy in dealing with highly reliable automated systems. The second one criticizes the concept because it contains a pejorative connotation which accuses the operator of an erratic behavior, and, thus, flattens the whole issue to another example of human error. And the third line of criticism questions the scientific value of the concept in general by arguing that just naming a certain effect does not help in understanding it.

The presentation will evaluate the current status of the concept. It will be argued that complacency in terms of a behavioral phenomenon indeed

represents an important issue of human-automation interaction, and does not just represent a rational strategy of humans dealing with highly reliable systems. However the underlying determinants of this effect are not very well understood, yet. Based on an analysis of the relevant literature as well as own studies, a framework model of complacency will be presented which might be helpful in identifying current gaps of knowledge and in guiding further research. This model assumes that complacency cannot be reduced to an issue of erratic human behavior but represents the result of a complex interplay of the human, the system and the situation. More specifically, it is argued that human-system interactions result in certain levels of trust in automation which in turn can lead to complacent behavior in interaction with a specific system. Whether or not this occurs seems to depend on the situational context, the performance state of operators, and dynamic adaptive processes in interaction with the system. Questions which need to be addressed in this context include, e.g., how complacency as a behavioral construct relates to trust in automation, what situational characteristics are relevant in supporting/preventing the occurrence of complacency, and how the adaptive processes involved in the development of complacency can be modeled.

Economic Rationality and the Use of Automation (Joachim Meyer)

How should rational beings, acting according to their preferences, use automation? How should we design automation for rational beings? I demonstrate on the example of an operator in a semi-conductor production clean room that operators' may use automation in complex ways, beyond simple trust or mistrust. They may act differently from what we might expect, if we believe they should use automation whenever it can be trusted. Even a perfectly trusted information source may not be followed by a rational operator, acting according to her preferences. Some implications for the design of automation at the device and the overall system level are discussed.

The Utility of Close Calls: Beneficial Applications of False Alarms for High-Consequence Task Environments (James P. Bliss)

The deleterious effects of false alarms have been discussed by researchers at length. Demonstrated performance effects have included increased reaction times, degraded response frequencies, and poor reaction appropriateness levels. Some theories concerning automation trust have focused on the idea that operators practice "trust calibration" as they interact with signaling systems. Such an idea suggests that trust is a malleable commodity that may be adjusted cognitively to coincide with the perceived reliability level of the automation. If true, complex task operators may be expected to adjust their trust levels in real time to reflect experiences with an automated signaling system. However, discussions of trust calibration rarely include the progression of alarm reactions across time, and make no mention of the possible benefits of false alarms. Some researchers have noted that presenting false alarms may in fact ultimately lead to heightened operator awareness, lower net workload, and resistance to the startle effect. This presentation will focus on these possible benefits. The presenter will address the possible benefits of false alarms, relying on literature from established cognitive theories and case studies from industries where signaling system failures are common. The topic area will also be explored as it could apply to variations in the operator perception and response processes, including shared response environments and extremely low and high reliability signaling systems. Particular attention will be devoted to the relationship between operator training strategies and subsequent perception of false alarms.

Different Types of False Alarms Influence Trust and Reliance (John D. Lee)

The study of operator response to automation that warns and alerts has a long history grounded in signal detection theory. Such a perspective draws on an engineering-centric definition of warnings that defines discrete categories of hits, misses, false

alarms and correct rejections. Acceptance tends to decline with systems that generate many false alarms relative to hits, which becomes problematic when the base rate of true events is low (Parasuraman, Hancock, & Olofinboba, 1997). While useful, signal detection theory may not be sufficient to explain operator acceptance and reliance on imperfect warning systems (Allendoerfer, Pai, & Friedman-Berg, 2008; Barnes, Gruntfest, Hayden, Schultz, & Benight, 2007; Lees & Lee, 2007).

A more system-centric definition of warnings is needed. Lees and Lee (2008) identified different types of “false alarms” based on the dimensions of purpose, process, and performance that underlie trust in automation. Differences between these types of false alarms had a strong influence on the trust and reliance on the warning system. Not all failures are detrimental to trust and compliance. Understanding how the performance of a warning system affects trust and reliance demands a more complete description of warning types than that afforded by signal detection theory. A description of warning types should reflect the actions and expectations of the operator.

Theories of Trust in Multiple Aids (Stephen Rice)

The objective of this discussion is to determine how operator trust is affected by the presence of multiple auditory diagnostic aids. Prior research on diagnostic automation has focused primarily on single aids. In the current study, two competing theories of trust in multiple aids are presented: a) Component-Specific Trust Theory (CST) predicts that operators will differentially place trust in automation aids of varying reliability; b) System-Wide Trust Theory (SWT) predicts that operators will merge their trust across the aids despite varying reliability levels. Participants flew a simulated UAV mission, performing a pursuit tracking task while monitoring two system gauges that were augmented by diagnostic automation. The data provided evidence for a SWT model, as operators merged their trust across multiple aids. Operator dependence in a perfectly reliable aid suffered as much as their dependence in paired unreliable aid.

When designing multiple automation aids for a single operator, designers need to take into consideration the impact of merging trust across systems.

REFERENCES

- Allendoerfer, K. R., Pai, S., & Friedman-Berg, F. J. (2008). *The complexity of signal detection in air traffic control alert situations*. Paper presented at the Human Factors and Ergonomics Society 52nd Annual Meeting.
- Barnes, L. R., Gruntfest, E. C., Hayden, M. H., Schultz, D. M., & Benight, C. (2007). False alarms and close calls: A conceptual model of warning accuracy *Weather and Forecasting*, 22, 1140-1147.
- Billings, C. E., Lauber, J. K., Funkhouser, H., Lyman, G., & Huff, E. M. (1976). *Aviation safety reporting system* (Tech. Rep. TM-X-3445). Moffett Field, CA: NASA Ames Research Center.
- Dekker, S. & Hollnagel, E. (2004). Human factors and folk models. *Cognition, Technology, and Work*, 6, 79–86.
- Inagaki, T. (2003). Adaptive automation: Sharing and trading of control. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp.147–169). Mahwah, NJ: Erlbaum.
- Lees, M. N., & Lee, J. D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics*, 50(8), 1264-1286.
- Mosier, K. L. & Skitka, L. J. (1996). Human decision makers and automated decision aids: Made for each other? In R. Parasuraman, & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications* (pp. 201-220). Mahwah, NJ: Lawrence Erlbaum Associates.
- Parasuraman, R. (2000). Designing automation for human use: Empirical studies and quantitative models. *Ergonomics*, 43, 931-951.
- Parasuraman, R., Cosenzo, K., & de Visser, E. (in press). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation

awareness, and mental workload. *Military Psychology*.

Parasuraman, R., Hancock, P. A., & Olofinboba, O. (1997). Alarm effectiveness in driver-centred collision-warning systems. *Ergonomics*, 40(3), 390-399.

Parasuraman, R., Molloy, R. & Singh, I. L. (1993). Performance consequences of automation induced "complacency". *The International Journal of Aviation Psychology*, 2, 1-23.

Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human Factors*, 38, 665-679.

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, 286-297.

Parasuraman, R., Sheridan, T. B. & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, 2(2), 140-160.

Parasuraman, R., & Wickens, C. D. (2008). Humans: Still vital after all these years of automation. *Human Factors*, 50, 511-520.

Scerbo, M. S. (2007). Adaptive automation. In R. Parasuraman and M. Rizzo (Eds.) *Neuroergonomics: The brain at work*. New York: Oxford University Press.

Wilson, G. F., & Russell, C. A. (2007). Performance enhancement in a UAV task using psychophysiological determined adaptive aiding. *Human Factors*, 49, 1005-1018.