

Misuse of Diagnostic Aids in Process Control: The Effects of Automation Misses on Complacency and Automation Bias

J. Elin Bahner, Monika F. Elepfandt and Dietrich Manzey
Berlin Institute of Technology
Berlin, Germany

The effects of misses of an automated alarm and fault diagnosis system on different manifestations of automation misuse were examined. 24 participants operated a complex multi-task process control simulation. During training, they either experienced automation misses or were only informed that failures might occur. The experience of misses reduced complacency towards the alarm function of the decision aid as well as omission errors but did neither affect complacency towards the aid's diagnostic function nor commission errors. Implications of this specific effect of automation misses for the design of training measures as well as the theoretical understanding of automation misuse are discussed.

INTRODUCTION

Sophisticated automation is finding its way into more and more work environments as diverse as aviation, maritime operations, and process control. Although automation exhibits a great potential to extend human performance and improve safety, it also has given rise to new sources of error and risks. One of these risks is represented by an inappropriate, e.g. too high level of trust placed in the automation by the human operator (Lee & See, 2004). Such over trust can lead to automation misuse, i.e. an uncritical reliance on the proper function of an automated system without recognizing its limitations and the possibilities of automation failures (Parasuraman & Riley, 1997). One manifestation of this misuse emerges in an inappropriate monitoring or cross-checking of automated functions, a phenomenon which commonly has been referred to as "automation induced complacency" or just "complacency" (Moray & Inagaki, 2000; Parasuraman, Molloy, & Singh, 1993). In several studies it was demonstrated that particularly high and consistent reliable systems give rise to complacency effects (e.g. Parasuraman et al. 1993; Prinzl, DeVries, Freeman, & Mikulka, 2001).

Complacency-like effects have been suggested to emerge not only in classical monitoring settings but also in other fields of human-computer interaction, notably in the use of decision aids. Such aids usually serve several functions. One of these functions involves some kind of alert, i.e. making the user aware of the fact that some action is needed. Beyond that, other functions often involve recommendations of specific actions to take. An example of such an aid might include a diagnostic aid in supervisory control which, on the one hand, provides an alert in case of critical system states and, on the other hand, recommends a sequence of appropriate actions to respond to this state. According to these different functions, two kinds of error can arise which might be related to complacency effects. The first one involves so called "omission errors", i.e. when operators rely so much on the alarm function of the aid that they do not monitor the system and fail to notice problems if the automated aid fails to alert them. The second one has been described as "commission error" which occurs when operators follow a recommendation of an automated aid even though this recommendation is

wrong. Hence, complacency, in terms of an insufficient monitoring or cross-checking of the automation, might represent a possible cause for both commission as well as omission errors. Mosier and Skitka (1996) referred to these two kinds of error as automation bias. Empirical research revealed that complacency and automation bias represent persistent and difficult to avoid problems (e.g. Bailey & Scerbo, 2007; Mosier, Skitka, Dunbar, & McDonnell, 2001).

However, one possible countermeasure against complacency and automation bias might consist in the experience of automation failures. Several studies demonstrated that even single automation failures can reduce trust in automation dramatically (e.g. Lee & Moray, 1992; Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003). Thus, over trust, and therewith the basis of both phenomena, should disappear by the experience of automation failures. Based on this rationale, Manzey, Bahner, and Hueper (2006) examined the effect of automation failures during training on complacency and commission errors in the use of a decision aid. This aid supported the operator by detecting, diagnosing, and managing occurring system faults. Automation failures during training consisted in false fault diagnoses provided by the aid. Results showed that this experience of false diagnoses during training reduced complacency compared to a control group, which was just informed, that automation failures might occur. Specifically, participants who experienced false diagnoses cross-checked the diagnoses provided by the decision aid in the subsequent test phase more thoroughly. However, exploratory data analyses revealed that the experience of false diagnoses did not increase cross-checking during "normal" system state, i.e. when the decision aid did not display any failure message. Even though the experience of diagnostic failures decreased the participants' level of complacency towards the diagnostic function, it obviously did not affect their level of complacency towards the alarm function of the decision aid. This implies that the participants perceived the two functions of the system as qualitatively different. Although this exploratory result does not allow for a clear-cut interpretation, it clearly contrasts the finding of Muir and Moray (1996) that distrust spreads between separate system functions. Yet, the result obviously bears analogy to the theoretical distinction between reliance and compliance in the

context of binary warning systems. While compliance refers to the response when an operator acts according to a warning signal, reliance represents the response when the warning system indicates that the system is intact and the user accordingly does not take precautions (Meyer, 2004). Several studies suggest that compliance and reliance are affected differently by automation failures, i.e. misses and false alarms of a warning system. Yet, it is still a matter of debate whether reliance and compliance are independent from one another (Dixon, Wickens, & McCarley, 2007; Meyer, 2001). However, whether such a differential effect of failure types found in the use of binary warning systems also holds for more complex decision aids, remains unclear. A clarification of this issue would be important with regard to operator training. More specifically, it would suggest that users might develop different levels of trust with regard to different automated functions although all of them are served by the same device. As a consequence, users would need to be familiarized with automation failures of all main automation functions during training in order to reduce automation misuse effects comprehensively.

The present study shall contribute to further elaborate the relationship between complacency, commission errors, and omission errors in interaction with decision aids. Using a similar experimental paradigm as Manzey et al. (2006) it is investigated to what extent experiences of automation failures during training affect the user's behavior with respect to different automated functions. Complementary to the study of Manzey et al. (2006), it is addressed how the experience of failures of the aid's alarm function ("automation misses") during training affects misuse towards the different aid's functions, i.e. its alarm function and diagnostic function, in the subsequent test phase. Two different experimental groups were compared. One group just got the general information that automation failures might occur but worked with a completely reliable aid during training. The other group got the same information. However, during training participants of this group were additionally exposed to sudden automation failures. These failures involved "automation misses", i.e. events in which the aid failed to alert the user in case of a critical system state. Assuming the existence of a specific failure effect on automation misuse, the following hypotheses can be derived: (1) Compared to the sole information that failures might occur, the experience of automation misses during training decreases complacency towards the alarm function of the decision aid and (2) reduces the number of omission errors in case of occurring automation misses. (3) The participants' degree of complacency towards the aid's diagnostic function remains unaffected by the experimental manipulation as does (4) the number of commission errors in case of a suggested false diagnosis.

METHOD

Participants

A total of 24 engineering students (4 female, 20 male) participated in the experiment. One male participant did not

obey the instructions regarding the preassigned procedure of fault detection and had to be excluded from the experiment. The age of the remaining participants ranged from 21 to 29 years ($M = 24.33$, $SD = 1.85$). They were paid € 40 each for completing the study. None of the participants had any prior experience with the AutoCAMS task environment used in the study.

Apparatus: AutoCAMS Task Environment

The experiment was conducted by using a modified version of the PC-based simulation of a process control task AutoCAMS (Hockey, Wastell, & Sauer, 1998; Lorenz, Di Nocera, Roettger, & Parasuraman, 2002). This simulation is based on the Cabin Air Management System (CAMS) task originally developed by Hockey et al. in order to investigate the effects of stress on complex human performance.

AutoCAMS simulates an autonomously running life support system of a spacecraft consisting of five subsystems that are critical to maintain atmospheric conditions in the space cabin with respect to different parameters (oxygen, nitrogen, carbon dioxide, temperature and pressure). By default all of these subsystems are automatically maintained within their target range. However, different faults may occur occasionally, due to a malfunction in any subsystem (e.g. leaks or blocks of a valve or defective sensors). The primary task of the operator involves supervisory control of the subsystems including diagnosis and management of system faults. The latter task is supported by an automated aid supporting fault diagnosis and management (Automated Fault Identification and Recovery Agent, AFIRA). In case of a fault, usually a general master alarm occurs. The presence of a critical system state always has to be approved by means of a mouse-click on an alarm mode icon which confirms the operator's being aware of the change of system state. Together with the alert, AFIRA displays both, a fault-diagnosis that is generated automatically, as well as a supposed sequence of actions for effective fault management which then has to be implemented manually by the operator. The proposed sequence of actions always includes hints for appropriate manual control of the defective subsystem until it works properly again, and for initiating the repair of the diagnosed fault. Manual control activities can be implemented by selecting a subsystem-specific control window from a control menu. In order to repair the fault, a maintenance menu has to be opened by a mouse-click and an appropriate repair order has to be selected and sent from this menu. The latter initiates a repair that is achieved after 60 seconds if the diagnosis has been correct. As soon as the fault has been repaired, AFIRA displays a success message. Yet, it remains part of the operator's task to verify that all system parameters are back in their target range and, if so, to deactivate the alarm mode (mouse-click on the corresponding icon).

However, in case of AFIRA failures, i.e. false diagnosis or missed system fault, manual fault diagnosis and management are required. In addition to the information provided by AFIRA, the operator has independent access to all relevant information about the state of the different subsystems that might be used to detect system faults independent of AFIRA or

to verify the fault diagnoses suggested by AFIRA. This includes information about tank-levels and gas-flow rates in different parts of the system, as well as a “history graph” displaying the time-course of different system parameters across the past four minutes. Yet, to get specific information displayed, the operator has to activate it by mouse-click on a specific field. The information then is shown for 10 seconds before it is switched off again until the participant recalls the information another time.

Besides the primary task, two secondary tasks have to be accomplished. The first one includes a prospective memory task which requires the operator to record the level of a certain parameter at fixed intervals (every 60 seconds). The second task represents a simple reaction-time task which requires clicking as fast as possible on a connection symbol which appears unpredictably (on average once a minute).

Procedure

The study consisted of two 4-hour sessions conducted on two different days. The first session included practice of manual fault identification and management. The second session included the experiment. In the first part of this latter session, participants were familiarized and trained with AutoCAMS, i.e. learned how to use AFIRA for fault diagnosis and management. As part of this training all participants were explicitly informed that failures of AFIRA may occur and warned to cross-check the system. However, only half of the subjects (“experience group”) were exposed to such failures during training. Whereas the diagnoses provided by AFIRA always were correct, two automation misses occurred, i.e. two out of ten system faults remained undetected by AFIRA. To make sure that the participants really noticed both of these two failures, they were asked to record each occurring system fault. Based on these records the experimenter provided feedback on the participants’ fault detection performance after each training unit (three units, each lasting 20 minutes).

For the other half of the participants (“information group”) AFIRA displayed and diagnosed all ten system faults correctly during practice. After this familiarization all participants had to work with AutoCAMS for 100 minutes. During this period a total of 14 system faults occurred. AFIRA detected and displayed the first nine of these faults as well as faults 11 and 12 correctly. However, AFIRA did not display faults 10 and 13 (automation misses) and provided a false diagnosis for fault 14.

Dependent Measures

Dependent measures were derived from log-file records of the mouse-clicks performed by the operators and the status of the different subsystems.

Complacency towards the alarm function. Information sampling during phases which are indicated to be fault-free by the decision aid enables participants to evaluate the factual system state. Based on this reasoning, the number of information requests (mouse-clicks) per minute during the last 120 seconds before the occurrence of a system fault was taken

as an (inverse) indicator of complacency towards the alarm function of the decision aid.

Omission error. All events where an automation miss occurred (faults 10 and 13) and participants did not activate the alarm mode before the system reached an extremely critical state, was counted as omission error. Critical system states were defined as “extreme” whenever a system parameter had exceeded the outer boundary of its target range.

Complacency towards the diagnostic function. In order to derive a direct measure for complacency towards the diagnostic function it was recorded to what extent the participants attempted to verify the automatically generated fault diagnoses before they initiated a repair order. This was done by analyzing which, if any, parameters of the different subsystems were sampled by operators after activation of the alarm mode and contrasting this with a “normative model” (Moray, 2003; Moray & Inagaki, 2000) of information sampling, i.e. which parameters should be looked at in order to verify a certain diagnosis. Based on this rationale, an automation verification score was defined as the portion of all parameters relevant to verify a certain diagnosis that were actually sampled by the participant. Note that this measure is inversely related to complacency, varying from zero (no attempt of verification at all; i.e. extreme complacent behavior) to one (perfect verification; no complacency at all).

Commission error. If a participant initiated the wrong repair order suggested by AFIRA for fault 14, a commission error was counted.

RESULTS

Complacency towards the Alarm Function

Information sampling data for the fault-free phases preceding faults 1-3, 4-6 and 7-9 were pooled in order to reduce intra-subject variability and analyzed by a 2 (group assignment) x 3 (fault blocks) analysis of variance (ANOVA) with “fault-blocks” as within-subjects factor. Participants of the “experience group” sampled significantly more information ($M = 19.06$) than participants of the “information group” ($M = 14.43$), $F(1, 21) = 4.37$, $p < .05$, (see Figure 1). Neither a main effect of fault blocks nor an interaction effect was observed.

Omission Error

At fault 10, when AFIRA failed for the first time, 80 percent of the information group, but only 18.2 percent of the experience group committed an omission error, $p < .01$ (one-tailed Fisher’s exact test). At fault 13, this group difference was not visible anymore, as again 18.2 percent of the experience group, but this time only 22.2 percent of the information group did not detect the system fault (see Figure 2). Comparison between participants who successfully detected the first automation miss and those who failed did not reveal any significant effect with respect to either kind of complacency.

Complacency towards the Diagnostic Function

Again data analysis was based on the first 9 faults and a 2 (group assignment) x 3 (fault blocks) ANOVA. No difference between groups was observed. Neither a main effect of fault blocks nor an interaction effect was observed. Participants sampled on average 60 percent of the parameters relevant to verify the suggested diagnoses (see Figure 3).

Commission Error

74 percent of the participants committed a commission error. These were distributed almost equally across the two experimental groups. Hence, no group effect emerged. Inspection of the verification behavior just before committing the error revealed that 80 percent of the participants followed the false recommendation because of varying levels of complacency towards the diagnostic function.

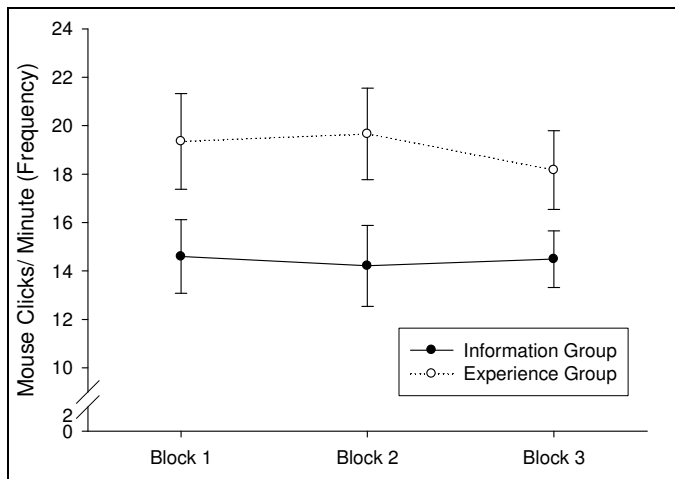


Fig. 1: Effect of failure information vs. experience on information sampling during fault-free system states (inversely related to complacency towards the aid's alarm function)

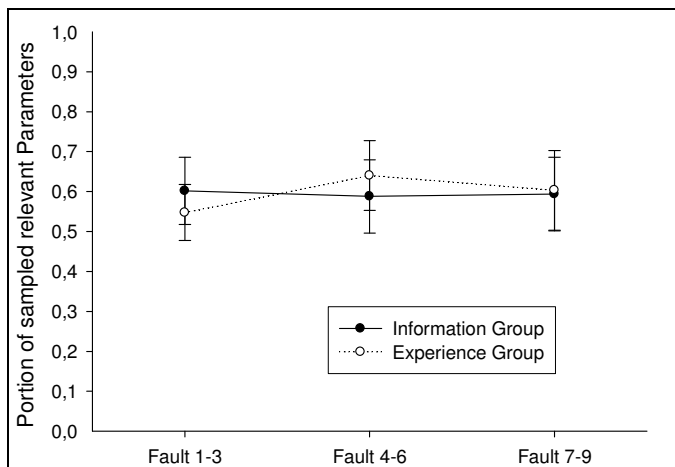


Fig. 3: Effect of failure information vs. experience on the verification of automated diagnoses (inversely related to complacency towards the aid's diagnostic function)

Yet, 20 percent of the participants followed the recommendation despite seeking out all parameters necessary to prove the automated advice wrong. Participants who committed a commission error showed a higher degree of complacency towards the diagnostic function with respect to the first nine faults, where AFIRA has worked reliably. This was revealed by a 2 (commission error yes/no) x 3 (fault blocks) ANOVA. Participants who detected the false diagnosis by AFIRA sampled a considerably higher portion of relevant parameters ($M = 0.89$) than participants who missed the failure ($M = 0.49$), main effect "commission error" $F(1, 21) = 15.01$, $p < .01$ (see Figure 4). No other effect became significant. However, participants who committed a commission error did not differ from participants who detected the false diagnosis with regard to their level of complacency towards the aid's alarm function. Furthermore, comparison regarding the number of omission errors did not reveal any difference between participants committing and avoiding commission errors.

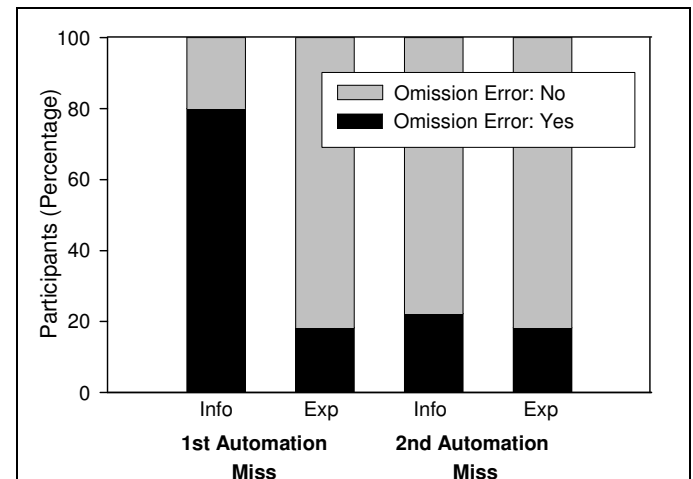


Fig. 2: Effect of failure information vs. experience on omission errors at the first and second event of automation misses

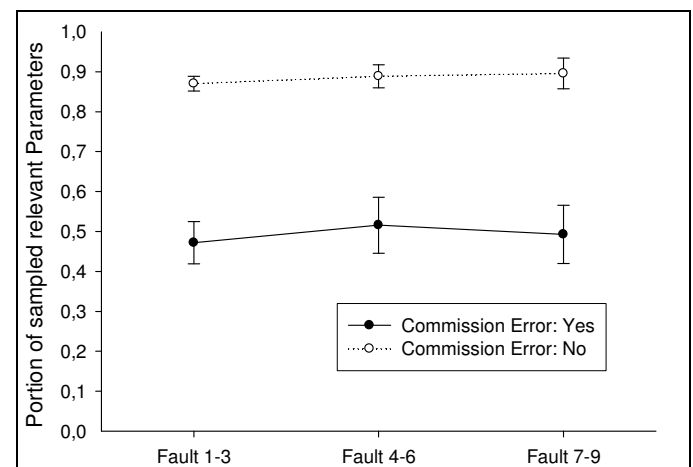


Fig. 4: Verification of automated diagnoses (inversely related to complacency towards the aid's diagnostic function) for participants making vs. not making a commission error

DISCUSSION

Four conclusions can be drawn from the data presented above: Firstly, the results provide clear evidence for a specific effect of automation failures. As expected, the experience of automation misses reduced the level of complacency towards the alarm function of the decision aid. Furthermore, the number of omission errors at the first automation miss was reduced by more than 60 percent for the experience group compared to the information group. The fact that this group effect disappeared when the second automation miss occurred, demonstrates once more the direct effect of automation failures. However, complacency towards the aid's diagnostic function and commission errors were not affected by the experimental manipulation. Similarly, Manzey et al. (2006) showed that false diagnoses during training reduce complacency towards the diagnostic function of a decision aid but not towards the aid's alarm function. Apparently, the impaired reliability of one system function does not call into question the reliability of the system as a whole. This implies that operators are very well capable of differentiating between function components with varying degrees of reliability, i.e. to exhibit what has been referred to as "high functional specificity" (Lee & See, 2004). Such a high functional specificity represents a precondition of an appropriate level of trust and accordingly a desirable effect.

Secondly, results suggest that commission and omission errors represent independent phenomena. This is revealed by the differential effects of automation misses and false diagnosis on the two different aspects of complacency and automation bias, as described above. Furthermore, committing one of the two error types was not associated with a higher risk of committing the other error type in question. In line with this effect, no link between complacency towards the diagnostic function and the alarm function of the decision aid was observed.

The third conclusion is that commission errors appear to be clearly linked to a high level of complacency towards the diagnostic function of a decision aid, as participants who committed a commission error showed a significantly higher level of complacency in previous trials. This effect is in line with results of Manzey et al. (2006) and provides further evidence for the assumption that complacency is one possible cause of commission errors. However, according to Mosier et al. (2001) commission errors might occur either because of some kind of complacency or because of a decision making problem, i.e. despite all relevant information necessary to falsify the recommendation of an automated system was sampled before. Exploratory data inspection reveals evidence for both kinds of commission errors. The majority (80 percent) of the participants committed a commission error due to some degree of complacency. Yet, about 20 percent followed the recommendation despite seeking out all information to prove the automated advice wrong.

The fourth conclusion is a practical one: Training programs which aim at a reduction of automation misuse should take the specific effect of automation failures into account. The present study shows that the experience of

automation failures during training represents an effective countermeasure to reduce automation misuse. Yet, the inhibiting effect remained failure specific. Hence, trainings which aim at a comprehensive prevention of over trust related effects should involve each automated function and the corresponding potential automation failures.

REFERENCES

- Bailey, N.R., & Scerbo, M.W. (2007). Automation-induced complacency for monitoring highly reliable systems: the role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science*, 8, 321-348.
- Dixon, S.R., Wickens, C.D., & McCarley, J.S. (2007). On the independence of compliance and reliance: are automation false alarms worse than misses? *Human Factors*, 49, 564-572.
- 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.
- Hockey, G.R.J., Wastell, D.G., & Sauer, J. (1998). Effects of sleep deprivation and user interface on complex performance: a multilevel analysis of compensatory control. *Human Factors*, 40, 233-253.
- Lee, J.D., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35, 1243-1270.
- Lee, J.D., & See, K.A. (2004). Trust in automation: designing for appropriate reliance. *Human Factors*, 46, 50-80.
- Lorenz, B., Di Nocera, F., Roettger, S., & Parasuraman, R. (2002). Automated fault-management in a simulated spaceflight micro-world. *Aviation, Space, and Environmental Medicine*, 73, 886-897.
- Manzey, D., Bahner, E.J., & Hueper, A.-D. (2006). Misuse of automated aids in process control: complacency, automation bias, and possible training interventions. In: *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society, San Francisco, 16-20 October, 2006*.
- Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors*, 43, 563-572.
- Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors*, 46, 196-204.
- Moray, N. (2003). Monitoring, complacency, scepticism and eutactic behavior. *International Journal of Industrial Ergonomics*, 31, 175-178.
- Moray, N., & Inagaki, T. (2000). Attention and complacency. *Theoretical Issues in Ergonomics Science*, 1, 354-365.
- 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.
- Mosier, K.L., Skitka, L.J., Dunbar, M., & McDonnell, L. (2001). Aircrews and automation bias: the advantages of teamwork? *International Journal of Aviation Psychology*, 11, 1-14.
- Muir, B., & Moray, N. (1996). Trust in automation. Part II: experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, 39, 429-460.
- 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., & Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Human Factors*, 39, 230-253.
- Prinzl, L.J., De Vries, H., Freeman, F.G., & Mikulka, P. (2001). *Examination of automation-induced complacency and individual difference variates (Tech. Memo. No. TM-2001-211413)*. Hampton, VA: NASA Langley Research Center.