MTR190017

MITRE TECHNICAL REPORT

Nothing Can Go Wrong—A Review of Automation-Induced Complacency Research

Dr. Valerie Gawron

January 2019

McLean, VA

Center for Advanced Aviation System Development

Table of Contents

List of Figures

List of Tables

Nothing Can Go Wrong—A Review of Automation-Induced Complacency Research

The ASRS defines complacency as “self-satisfaction, which may result in non-vigilance based on an unjustified assumption of satisfactory system state.” Ropelewski (1996) suggested that pilots suffer from complacency when they become too reliant on and confident of the automation. This can lead to accidents. He gives the example of the Air Inter A320 accident near Strasbourg France on January 20, 1992. The A320-pilot entered a trap involving low airspeed, low altitude, low power, and rising terrain into which the aircraft crashed without enough power to climb. System prevented pilot from raising the nose of the aircraft to avoid the terrain since this would have stalled the aircraft.

To help study automation complacency, Singh, Molloy, and Parasuraman (1992; 1993) developed and then validated a scale for measuring the potential for complacency. They tied the concept of complacency potential to “premature cognitive commitment.” The scale has 20 items that are rated on a five-point Likert scale. 139 undergraduate students completed the original scale. Their responses were submitted to a factor analysis that identified the following factors: general automation, confidence, reliance, trust, and safety-related complacency. One example of an item on the scale is: “People save time by using automatic teller machines (ATMs) rather than a bank teller for banking transactions” (p. 23). An additional 175 undergraduates completed the scale. The results were similar.

In a more recent study, Prinzel, DeVries, Freeman, and Mikulka (2001) examined the effects of three personal traits on performance in an automated task. The traits were:

1. Complacency potential as measured by the Complacency Potential Rating Scale (CPRS) in which subjects rate each of 20 items on a scale from 1 (strongly agree) to 5 (strongly disagree). This was the same scale as used by Singh, Molloy, and Parasuraman (1992). In addition, individuals were classified as high or low in complacency if they scored above or below the group median on this scale.
2. Boredom proneness as measured by the Boredom Proneness Scale (BPS) in which subjects answer each of 28 items with a yes or a no. A sample item is: “It is easy for me to concentrate on my activities.”
3. Cognitive failure as measured by the Cognitive Failure Questionnaire (CFQ) in which subjects identify how often 25 items occurred to them in the last six months from very often (0) to very often (4). A sample item is: “Do you fail to notice sign posts in the road?”

The subjects were 40 undergraduate students. The tasks were monitoring an automated system for failures, managing fuel, and tracking. The last two tasks were performed manually. Subjects completed two 40-minute sessions. The sessions were scheduled one day apart. The results indicated that high complacency subjects did worse on both the system monitoring task and on the tracking task than those in the low complacency group. High complacency subjects also scored higher on the BPS than low complacency subjects. All three scales were significantly correlated: CPRS x BPS = +0.596, CPRS x CFQ = +0.709, and BPS x CFQ = +0.717. The authors concluded, “there are personality individual differences that are related to whether an individual will succumb to automation-induced complacency” (p. 35).

A related issue is over reliance on automation. Riley (1996b) proposed a theory to identify the factors associated with this phenomenon (see Figure 1). He then ran a series of simple computer-based experiments to test his theory. There were two tasks: 1) categorize a character as either a letter or a number and 2) correct random disturbances of a marker from a target location. Riley manipulated the reliability (50 versus 90% correct) of the automated system that could be selected to perform the categorization task. Workload and uncertainty were also varied. The dependent variable was the proportion of subjects who used automation. The overall use of automation was low: 35% under normal conditions to 50% in high workload. Further, there was no decrease in the proportion of subjects using automation after the automation failed. His subjects were undergraduate students. In a follow-on experiment with pilots, one third of the pilots continued to use the automation throughout its failure period.

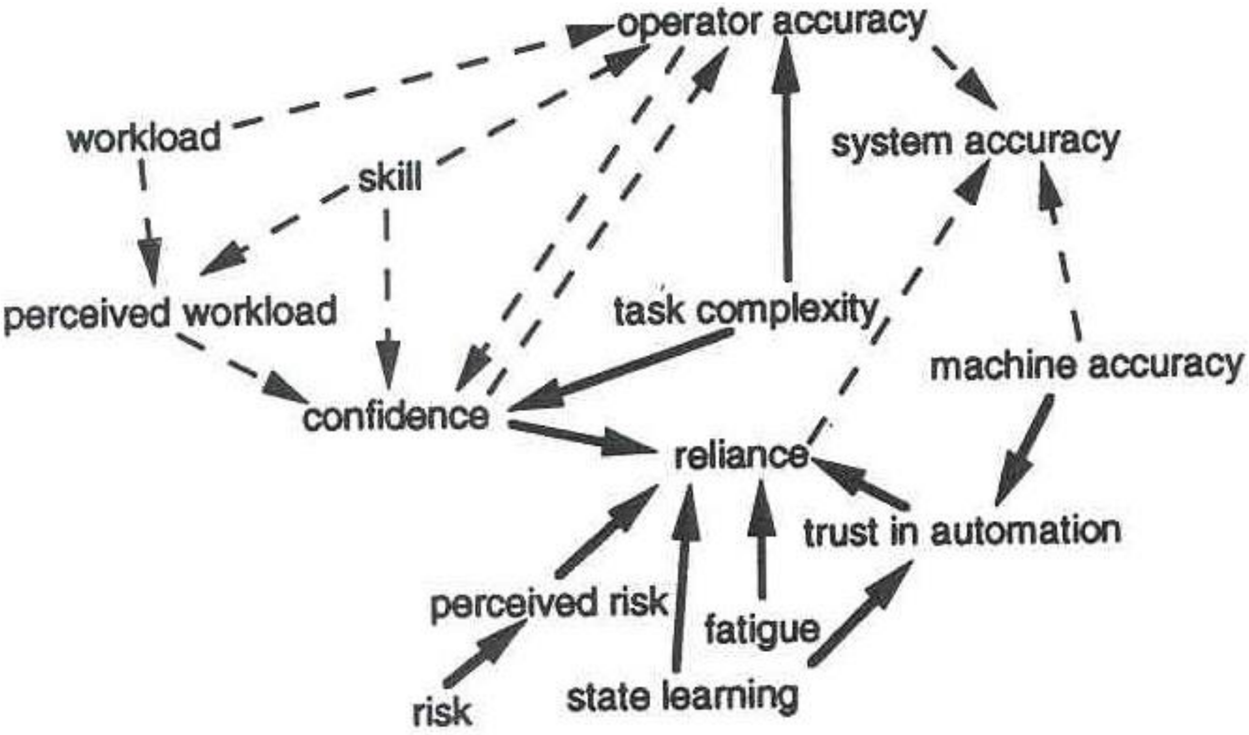

Figure 1. Revised Theory of Automation Use. Dotted Arrows Show Hypothesized Relationships that Have Not Been Confirmed by Experimental Evidence, Whereas Solid Lines Represent Those Relationships Supported from Evidence (Riley, 1996b, p. 33).

The Knowing—Or Not—A Review of the SA Research

SA is knowledge relevant to the task being performed. For example, pilots must know the state of their aircraft, the environment through which they are flying, and relationships between them, such as thunderstorms are associated with turbulence. It is a critical component of decision making and has been included in several models of decision making (e.g., Dorfel & Distelmaier model, 1997; see Figure 2). SA has three levels (Endsley, 1991): Level 1, perception of the elements in the environment; Level 2, comprehension of the current situation; and Level 3, projection of future status. Endsley (1995) analyzed aviation accidents from 1989 to 1992 to identify SA problems. The results are presented in Table 1. As can be seen from the table, the predominant errors were related to vigilance and failure detection although errors occurred at all three levels of SA. Jones and Endsley (1996) applied the same methodology to ASRS incidents (see Figure 3). Clearly the major problem is failure to monitor. The authors went farther and explored why pilots failed to monitor automated systems. The results are presented in Figure 4. Not unexpectedly the most frequent reason was task distraction. One example is the Eastern L-

1011 accident near Miami Florida on December 29, 1972. This result directly impacts the design of automated systems.

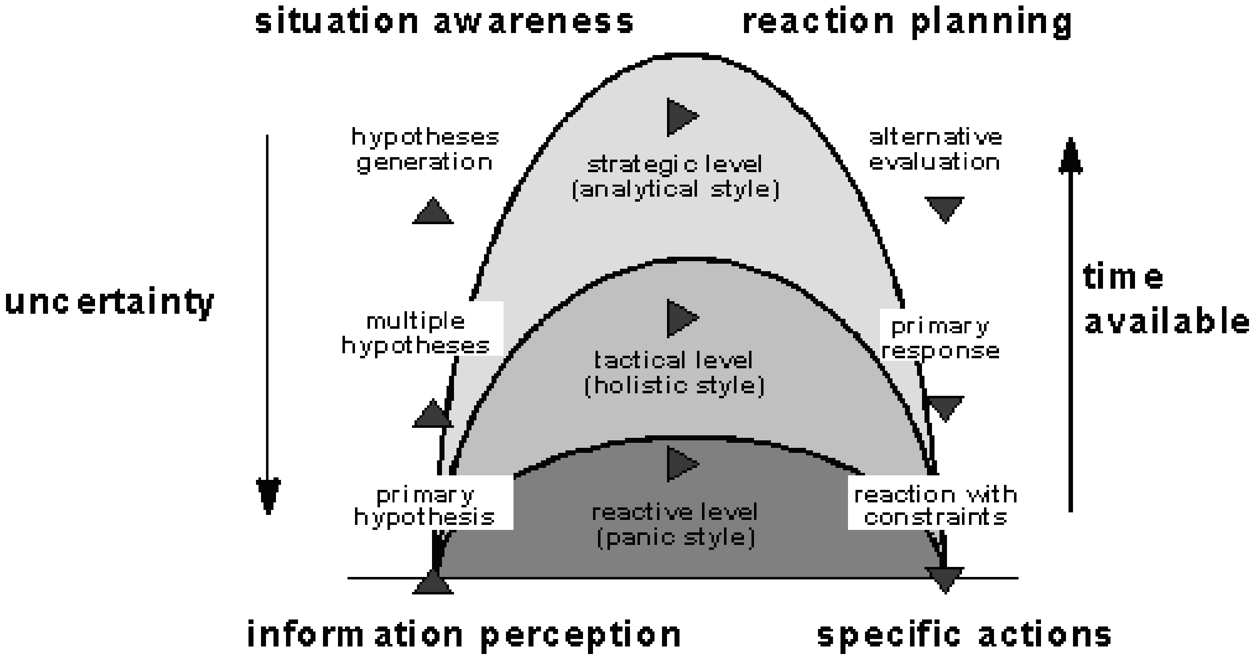

Figure 2. Decision Making Under Uncertainty and Time Pressure (Dorfel & Distelmaier, 1997, p. 2)

Table 1. ***Aviation Accidents in Which SA was a Factor (Endsley, 1995, p.290)***

Accident	Description	Position	Causal Factors	Other Factors
NTSB/AAR-92/05	Spatial disorientation	• Captain	• Level 1 - misperception	Physiological
NTSB/AAR-92/01	Crash on approach	• Captain	• Level 2 - other (significance)	Weather, Decision
NTSB/AAR-91/09	Crash on take-off (icing)	• Captain	• Level 2 - no/poor model	Weather, Procedure, Physiological
NTSB/AAR-91/08	Landing aircraft on occupied runway	• Controller	• Level 1 - difficult to detect, memory failure/task load, distraction	
NTSB/AAR-91/05	Runway collision (fog)	• Crew • Controller	• Level 1 - difficult to detect • Level 1 - difficult to detect • Level 2 - over-reliance on defaults	Weather, Procedure

Accident	Description	Position	Causal Factors	Other Factors
NTSB/AAR-91/04	Ran out of fuel	• Crew • Controller	• Level 2 - other (significance) • Level 3 - failure to project (time) • Level 1 - no data	Weather
NTSB/AAR-91/03	Landed aircraft on occupied runway	• Controller 1 • Controller 2	• Level 2 - other (integration) • Level 1 - failure/distraction	
NTSB/AAR-91/01	Loss of control - landing	• First Officer	• Level 1 - failure/omission	Psycho-motor
NTSB/AAR-90/05	Crash into mountain	• Crew	• Level 1 - difficult to detect • Level 2 - wrong model	Weather, Decision Physiological
NTSB/AAR-90/04	Struck power lines	• Crew	• Level 1 - misperception	Weather, Procedure
NTSB/AAR-90/03	Crash on take-off (Mistrimmed rudder)	• Captain	• Level 1 - failure/task load • Level 2 - wrong model	Procedure
NTSB/AAR-90/02	Loss of control (cargo door open)	• First Officer • Captain • Captain	• Level 1 - difficult to detect • Level 1 - no data (auto. Failure) • Level 3 - no/poor model	Mechanical Psycho-motor
NTSB/AAR-89/04	Crash on take-off (Miss-set flaps & slats)	• First Officer • Captain	• Level 1 - failure/task load • Level 1 - no data (auto. Failure)	Procedure
NTSB/AAR-89/01	Crash on approach	• First Officer • Captain	• Level 1 - failure/task load, misperception • Level 1 - failure/attn. Narrowing • Level 1 - failure/misperception	Decision

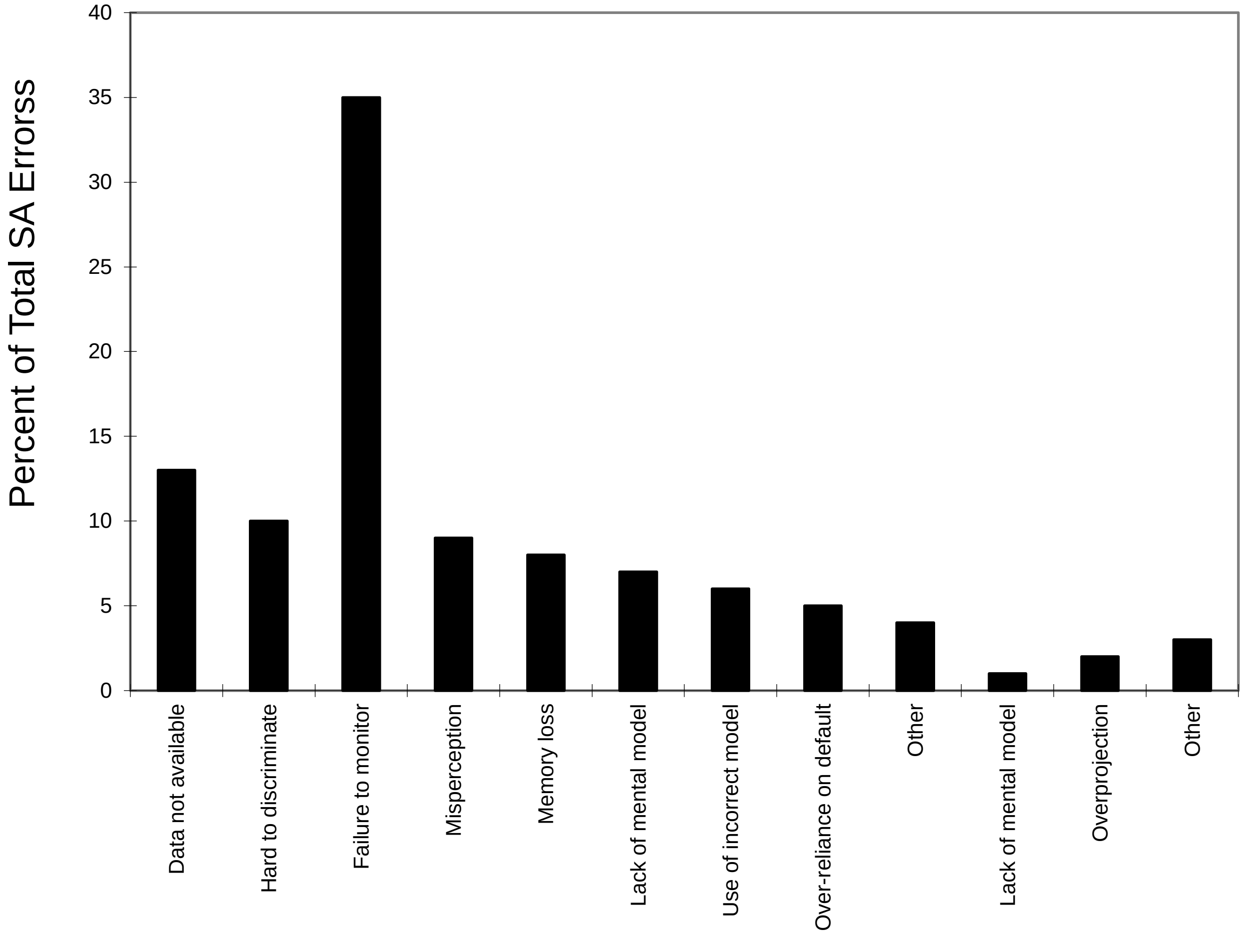

Figure 3. Based on 143 ASRS Incidents January 1986–May 1992 (Jones & Endsley, 1996, p. 509)

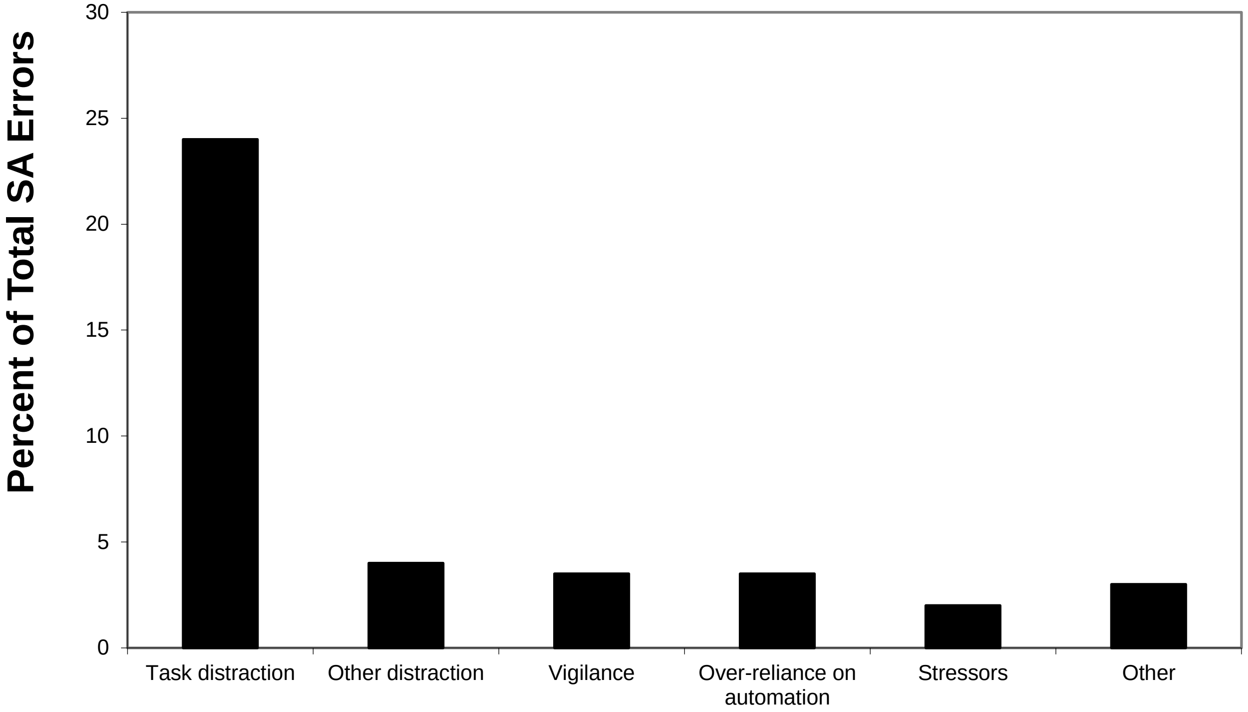

Figure 4. SA Errors Associated with Failure to Monitor (Jones & Endsley, 1996, p. 510)

In follow on work, Kaber and Endsley (1997) classified human supervisory control and monitoring of automated systems as forms of out-of-the-loop (OOTL) performance. As such they argue it is associated with the following negative consequences: “operator failure to observe system parameter changes and intervene when necessary (vigilance decrement), human over-trust in computers (complacency), operator loss of system of SA, and operator direct/manual control skill decay” (p. 126). The authors then suggest using LOA to counter the decrements in OOTL performance. Their LOA taxonomy is presented in Table 2.

Table 2. ***LOA Taxonomy (Kaber & Endsley, 1997, p. 129)***

LOA	Monitoring	Generating	Selecting	Implementing
1. Manual control	Human	Human	Human	Human
2. Action support	Human/computer	Human	Human	Human/computer
3. Batch processing	Human/computer	Human	Human	Computer
4. Shared control	Human/computer	Human/computer	Human	Human/computer
5. Decision support	Human/computer	Human/computer	Human	Computer
6. Blended decision making	Human/computer	Human/computer	Human/computer	Computer
7. Rigid system	Human/computer	Computer	Human	Computer
8. Automated decision making	Human/computer	Human/computer	Computer	Computer
9. Supervisory control	Human/computer	Computer	Computer	Computer
10. Full automation	Computer	Computer	Computer	Computer

Lessons Learned

A general lesson learned in automation design is that automated systems are not black and white but rather multiple shades of gray. There have been many attempts at naming these shades of gray. Kaber and Endsley (1997) presented a taxonomy (see Table 2). Roberts and Gawron (1986) took a different approach. In Figure 5 is their flow chart comparing operator and manager responsibilities. Note human and machine are not the terms used and either could be operator or manager.

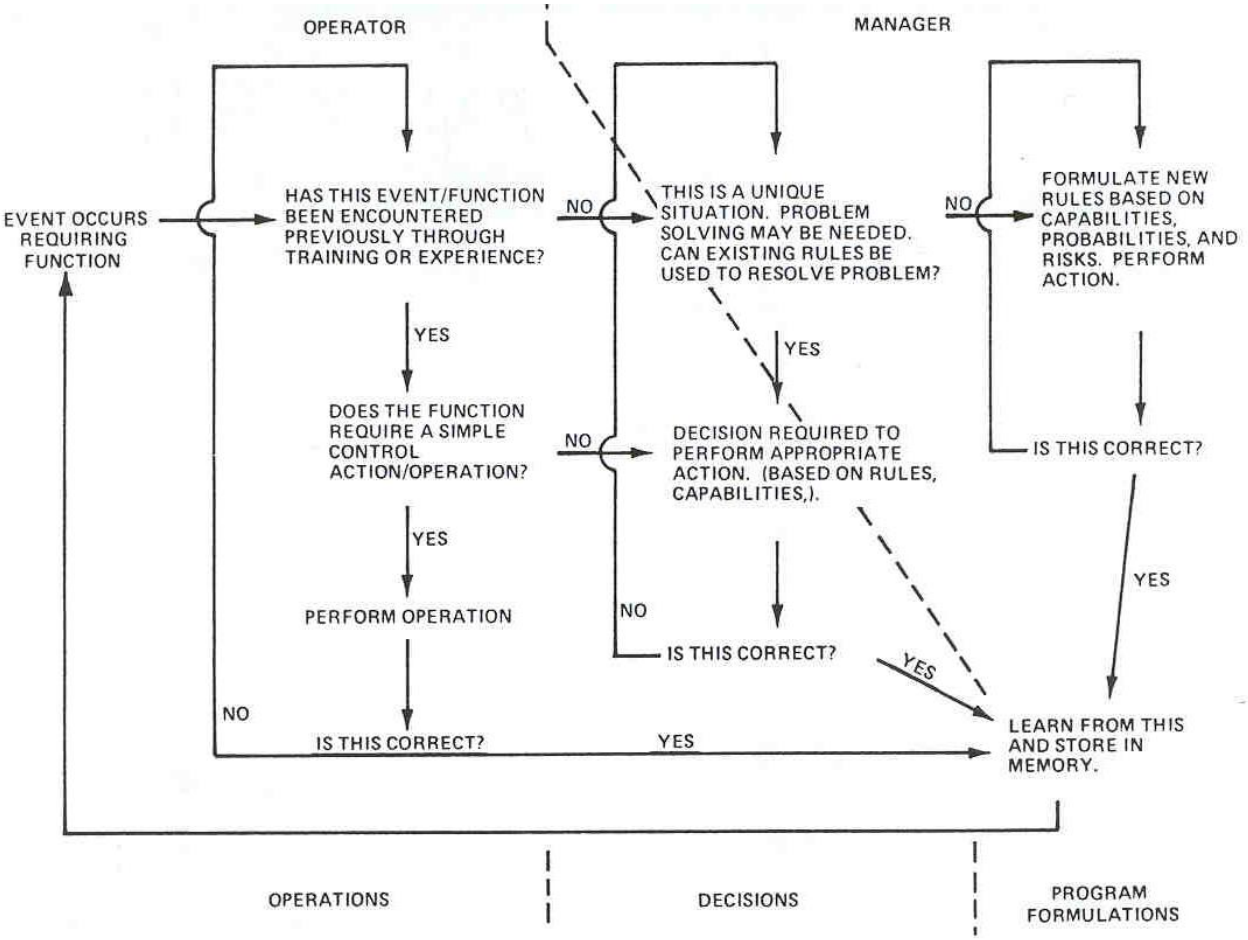

Figure 5. Operator-Manager Flow Chart (Roberts & Gawron, 1986, p. 858)

Designing for the many shades of gray is the first lesson learned in automation. Others include the need for human centered design, designing for failure, and designing for automation use, misuse, disuse, and abuse.

References

Dorfel, G. & Distelmaier, H. (1997). Enhancing situational awareness by knowledge-based user interfaces. In K. Garner (Ed.), *Proceedings for the Second Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment.* Patuxent River, MD: Naval Air Warfare Center Aircraft Division.

Endsley, M. R. (1991). Situation Awareness in dynamic systems. In R. M. Taylor (Ed.), *Situational awareness in dynamic systems (IAM Report 708).* Farnborough, United Kingdom: Royal Air Force Institute of Aviation Medicine.

Endsley, M. R. (1995). *A taxonomy of situation awareness errors in human factors in aviation operations,* (Vol. 3, pp. 287–292). Proceedings of the 21st Conference of the European Association for Aviation Psychology, Brookfield, Vermont: Avebury Aviation.

Jones, D. G. & Endsley, M. R. (1996). Sources of situational awareness errors in aviation. *Aviation, Space, and Environmental Medicine, 67(6),* 507–512.

Kaber, D. B. & Endsley, M. R. (1997). Out-of-the-loop performance problems and use of intermediate levels of automation for improved control system functioning and safety. *Process Safety Progress, 16(3),* 126–131.

Prinzel, L. J., DeVries, H., Freeman, F. G., & Mikulka, P. (December 2001). *Examination of automation-induced complacency and individual difference variates* (NASA/TM-2001-211413). Hampton, Virginia: National Aeronautics and Space Administration.

Riley, V. (1996b). Operator reliance on automation: Theory and data. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: theory and applications,* (pp. 19–36). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Roberts, D. P. & Gawron, V. J. (1986). *Operator/Manager Classification as a Precursor to Function Allocation,* (pp. 857–860). Proceedings of the 30th Annual Meeting of the Human Factors Society.

Ropelewski, R. (August 1996). Control in the cockpit. *Aerospace America,* 28–33.

Singh, I. L., Molloy, R., & Parasuraman, R. (1993). Automation-induced “complacency”: Development of the Complacency-Potential Rating Scale. *International Journal of Aviation Psychology, 3(2),* 111–122.

Singh, I. L., Molloy, R., & Parasuraman, R. (1992). *Development and validation of a scale of automation-induced “complacency.”* (pp. 22–25). Proceedings of the Human Factors Society 36th Annual Meeting.