HUMAN SERGEANTS, ROBOT SOLDIERS: SOLDIER-MACHINE INTEGRATION ISSUES WITH SEMI-AUTONOMOUS SYSTEMS

John S. Barnett, U.S. Army Research Institute
John P. Holmquist, University of Central Florida
Orlando, Florida

ABSTRACT

The Army of the future may use future combat systems (FCS) consisting of human soldiers directing semi-autonomous systems – essentially robot soldiers. This integration of human flexibility and adaptability with the power of technology promises to forge a team with remarkable combat power, provided the team can work together efficiently. However, previous experience with automation in aviation has shown human-machine teams may experience unique problems. Automation affects human workload, decision-making, crew interaction and team performance. People's confidence in the automated systems affects their performance, and they often have difficulty dealing with automation failure. It is conceivable that similar problems, as well as FCS-specific problems, may manifest themselves in future combat systems. Therefore, research is necessary to identify areas where the soldier and the automation do not get along, so that the design process can eliminate, or at least mitigate, these problems. This paper presents some known automation concerns and suggests how research can be used to address these issues prior to fielding.

ABOUT THE AUTHORS

Dr John Barnett is a research psychologist with the U.S. Army Research Institute where he conducts research in training and evaluation. He holds a Ph.D. in Applied Experimental and Human Factors Psychology from the University of Central Florida. Some of his previous research has included investigating how people develop trust in automated systems. He is a former U.S. Air Force officer with a background in aviation, operations planning, and command and control.

John P. Holmquist holds a Masters of Business Administration and has a background in computer and network design. He has participated in research concerning human expectations and reliability in automated systems and warning systems. He is currently pursuing a Ph.D. in Applied Experimental and Human Factors Psychology at the University of Central Florida and is a Research Fellow with the U.S. Army Research Institute.

HUMAN SERGEANTS, ROBOT SOLDIERS: SOLDIER-MACHINE INTEGRATION ISSUES WITH SEMI-AUTONOMOUS SYSTEMS

John S. Barnett, U.S. Army Research Institute
John P. Holmquist, University of Central Florida
Orlando, Florida

Recent changes in the world's political structure have precipitated a review of U.S. national strategy. Evolving strategic issues have led the U.S. military to extrapolate how future conflict might be conducted. Consequently, the U.S. Army is investigating how best to modernize its forces to meet this view of the future. A study by the National Defense Panel (1997) concluded that in the future, U.S. military forces will need vastly different capabilities from those they currently possess. This spurred the U.S. Army to investigate what future warfare might entail and how technological systems might be used to support combat in the future.

One of the elements under investigation by the Army is a future ground combat system. This extrapolated technological system has been dubbed Future Combat Systems (FCS). FCS is expected to be a multi-mission ground combat platform incorporating advanced technology components, including automation, that enable a significant increase in combat effectiveness (U.S. Army Armor Center, 2001).

Including semi-autonomous or robotic systems as elements of FCS has the potential to significantly increase the capabilities of the system. Applied correctly, automation can amplify human capabilities and significantly increase combat power.

However, there is the possibility for unexpected problems with human soldiers controlling an automated or semi-autonomous weapons system. Research from other domains, such as aviation, indicates applying automation or robotic systems to complex tasks often creates unique problems which must be addressed before applying automation to weapons of war.

Fortunately, there exists a window of opportunity early in the design process for research to address potential problems. Human factors research, particularly in the area of human-automation interaction, is important to uncover soldier-FCS mismatch issues which might interfere with FCS combat capability. Luckily, there are a number of exceptional tools available, such as modeling and simulation, to assist in this research.

Research using modeling and simulation presents a unique opportunity to build a robust virtual system before any component exists in physical space. This may be particularly propitious with a system such as the FCS. Virtual models of FCS components tested in a virtual environment may be indistinguishable from the real components, at least from the soldier-operator's standpoint. Such transparence to the operator suggests testing these components in virtual environments may have high external validity.

The purpose of the present work is to suggest that research is necessary to identify areas of concern in FCS where the soldier and the automation do not get along, so that the design process can eliminate, or at least mitigate, these problems.

The next section provides some background on known automation concerns from the aviation field, followed by a discussion of the concept of future combat systems and how automation problems may manifest themselves in FCS. Finally, some recommendations are made about how research might be used to find solutions to these potential problems before any future system is fielded.

KNOWN AUTOMATION CONCERNS

The use of autonomous or semi-autonomous systems in highly technical domains is not new. Automation has been a factor in the field of aviation since the 1980s. Since its introduction, considerable research has focused on human-machine integration in aviation. In his book on automation in aviation, Billings (1997) details the history of introducing automated systems into the aviation domain, including the numerous unexpected problems that have surfaced. These problems and concerns with automation fall into several categories, including poor interface design, automation affecting team performance, workload concerns, human confidence in automation, automation affecting human decision-making, and dealing with automation failure (Barnett, 2001). A brief description of each of these categories follows.

Poor Interface Design. The interface between the human and automation, if not properly designed, can often lead the human astray. For example, in 1985 a Boeing 747 flying at 41,000 feet developed an increasingly severe loss of power in one of its engines.

The crew was unaware of the severity of the problem because the autopilot compensated for the power loss as long as it was able. Eventually the autopilot was unable to compensate for the power loss, and the aircraft fell 30,000 feet until the pilots regained control. Without proper feedback on the autopilot's actions, the impending loss-of-control was hidden from the flight crew (Wiener & Nagel, 1988). Incidents such as this one suggest that in addition to human error and automation malfunctions, there may be a new class of errors which comes from the interaction of people with automation (Barnett, 2001).

One oft-repeated complaint from aviation was the automation often was too complex; it offered too may ways to perform the same task. Also, there was often too many different modes of operation, a problem known as *mode proliferation* (Sarter & Woods, 1997). Mode proliferation tends to precipitate *mode errors*. Mode errors are those in which an action is appropriate in one mode, but an error in another (Wickens, 1992). For example, raising the landing gear is appropriate if the aircraft is airborne, but a serious error if the aircraft is still on the runway.

The complexity of automated systems in aviation often confused pilots as to what mode they were operating in Even worse, the automation would change modes based on rules which weren't clear to pilots. This often led pilots to distrust the automation.

Automation Affecting Team Interaction. Automation affects the way teams and crews interact. The automated system often becomes part of the team and can be assigned tasks like a crewmember (Bowers, Oser, Salas, & Cannon-Bowers, 1996). However, automation also changes communication patterns and the way team members help each other.

In aviation, crewmembers are encouraged to cross-check each other to reduce the probability of errors. This is an important element of aviation safety. However, on some aircraft, the placement of automation control panels makes it impossible for crewmembers to see what the other members are doing, which makes it impossible to perform effective cross checks. For example, when a first officer enters information into a terminal, the captain may not be able to catch input errors, and vice versa.

Automation may disrupt normal crew coordination. For this reason, Morgan, Hershler Weiner, and Salas (1993) found crew coordination more important in automated aircraft. Automation may also affect traditional team roles and alter the authority gradient. For example, when junior team members are assigned the role of operating automated systems, team leaders are sometimes left out of the loop. Thus, significant responsibilities may shift to less experienced team members.

Workload Concerns. Frequently, the purpose of introducing automation onto the flight deck was to reduce pilot's workload. However, experience with automated systems indicates the effect of automation on operator workload is fairly complex. Studies indicate automation decreases workload at the wrong time (Billings, 1997). It often decreases workload when workload is already low, but may actually increase workload during periods of more intense activity (Sarter, 1996).

In commercial aviation, autopilots and Flight Management Systems (FMS) are automated flight control systems. They are most useful during the cruise portion of a flight, which typically is a low workload time for pilots. However, during the approach and landing phase, which often requires frequent changes of flight path in high traffic areas, the FMS requires extensive reprogramming to keep up with changes relayed by Air Traffic Control. Thus the automation exacerbates an already high workload environment.

Automation may also change the type and complexity of tasks. Users may trade systems operation for automation monitoring (Weiner & Nagel, 1988; Tsang & Vidulich, 1989, Bowers, Oser, Salas, & Cannon-Bowers, 1996). Thus, automation may reduce the physical workload, yet increase the cognitive workload (Endsley, 1997) and problem solving requirements (Idaszak, 1989).

Human Confidence in Automation. Human trust in automation runs the gamut from too much trust, which fosters complacency, to distrust, which often defeats the advantages of automated systems. Typically, human interaction with automation follows a common pattern. Initially operators implicitly trust automated systems until they experience a failure of the automation, at which time they develop almost complete distrust in the system.

Unquestioning trust in automation has been termed *automation bias*. An experiment by Mosier, Skitka, Heers, and Burdick (1998) illustrates this phenomenon. Pilots in a simulator were presented information from an automated system which indicated an engine fire, although other, non-automated instruments did not confirm the fire. Regardless, the pilots reacted to the reported engine fire without verifying the information presented by the automation. Their uncritical acceptance of the information presented by the automated system led them to react to an emergency which did not exist. Thus, the user's attitude towards automation can have real consequences.

On the other hand, users often have experiences with automation, which undermines their trust in the systems. Several experiments have demonstrated that people's trust in automation deteriorates after the system fails (Lee & Moray, 1994; Eidelkind & Papatonopoulos, 1997). Although disregarding automation in favor of manual operation may be considered a "fail-safe" alternative by distrustful operators, it can also mean forgoing some of the advantages afforded by automated systems (Barnett, 2001).

Automation Affecting Human Decision-Making. People's confidence level in automated systems often affects the decisions they make. As noted above, people tend to over rely on computer decision aids. When this occurs, they often curtail their information gathering and situation assessment and base their decisions on the automated information without question.

On the other hand, research has also shown that automation aids decision-making when the automation presents accurate status or recommendation information (Bowers, Oser, Salas, & Cannon-Bowers, 1996). This suggests that automation can have a positive influence on human decision-making, provided the information is accurate and the automation is reliable.

Dealing with Automation Failure. Responding to automation failure became a major concern in commercial aviation. The two primary issues were the problems with detecting automation failures, and the difficulty in recovering from those failures and reverting to manual modes.

Pilots complained that failures were frequently difficult to detect. Such failures were known as *silent* or *invisible* failures (Sarter & Woods, 1997). For example, a display may update information infrequently, and when it fails, the information on the screen remains the same. The failure might not be detected until the pilots realize the information displayed is obviously incorrect.

Unfortunately, monitoring systems can be a cognitively effortful task (Wickens, 1992). For this reason, it is often the case that operators do not regularly check the accuracy or status of automated systems (Mosier, Heers, Skitka & Burdick, 1997), and there is frequently a lower failure detection rate under passive monitoring (Parasuraman, Mouloua, Molloy, and Hillburn, 1996).

Ironically, the more reliable the automation, the less the probability that operators will detect failures. Operators who believe the probability of failure is low, will monitor the system less often, a classic example of expectancy theory in vigilance (Wickens, 1992).

In addition, once automation failures are identified, it is often difficult to recover from them (Endsley, 1996). When automation is assigned a task, there is often little feedback to the operators about the status of task accomplishment, thus pilots often reported they felt "out of the loop" and had little control or knowledge of automation actions. Therefore, when the automation fails, the operators have to first determine how much of the task was accomplished before they can complete the task manually. In cognitive terms, they have to update their mental models of mission status before they can switch to manual operation.

Recovery from automation failures is most difficult under stressful or non-routine conditions. Sarter (1991) found pilots were able to revert to manual modes with little difficulty during routine operations, but committed numerous errors during abnormal operations. Unfortunately, it appears that when automation is needed the most, it not only may not be useful, but may often add to the problem.

An appreciation of these issues from aviation coupled with an understanding of the direction envisioned for FCS may indicate how such problems may manifest themselves in the future.

FUTURE COMBAT SYSTEMS

The FCS is designed to prepare the U.S. Army for the combat needs of the 2012-2020 time frame. Analysts believe that conflicts in this time frame will best be met by forces that are lighter and more transportable (Defense Advanced Research Projects Agency [DARPA], 2000). In addition, as conflicts move away from large-scale battles defending home and country, to peace keeping and containment missions in far away lands, public willingness to accept military casualties has declined. Therefore, deployability and reduction of casualties are the two leading reasons for the research into FCS (Richardson, 2001).

The FCS proposal is seen as a "clean sheet of paper," allowing contractors to explore new systems and technologies. In light of this strategy, some designers propose the introduction of semi-autonomous or robotic systems into a distributed combat architecture consisting of targeting, weapons, and command and control components.

In this view, the FCS would be a distributed, network centric system with all of the functionality necessary to be successful on the modern battlefield. It would be distributed among multiple vehicle elements whose capabilities sum to the capabilities necessary to achieve supremacy in combat (Federation of American Scientists [FAS], 2001).

Thus, FCS is envisioned to consist of highly automated components directed by human soldiers.

Expected Automation Concerns with FCS

Could problems similar to those seen in aviation manifest themselves with future combat systems? It is entirely possible that FCS will encounter some of these same tribulations, and more besides. It is difficult to extrapolate how people will perform with a theoretical future combat system whose design has not been finalized. However, experience from aviation indicates human-automation interaction is a complex relationship and a potential source of problems. For this reason, it is important to conduct research to establish how soldiers will interact with the FCS interface. Such research is vital to uncover interaction issues so they may be addressed in FCS design. Experience from aviation may provide a starting point for soldier-FCS interaction research. Failure to confront these concerns prior to fielding may doom FCS to failure.

RESEARCH FOCUS

Determining how soldiers will interact with FCS and using this information to perfect FCS design will require a judicious research program. Often, with innovative concepts like FCS, many of the research questions are not known initially, and preliminary research is required to indicate the direction of subsequent experimentation. However, a knowledge of human factors principles can suggest broad areas for initial investigation.

One suggestion for the FCS research strategy might be to base it on the systems design model and consider the input, throughput, and output of the soldier-automation system. The input would consider how the soldier acquires information from the system interface through the displays, the throughput would examine how the soldier processes the information and makes decisions, and the output would be concerned with how the soldier implements those decisions and controls the FCS.

In addition to this broad architecture, there are several areas which should receive special attention when considering how soldiers will operate FCS, including interface design and issues related to workload.

Interface Design

Interface design must take into account more than the controls and displays. It is also vital to consider how the human user will operate the controls and displays. Human-centered design is essential for a well designed interface. Human-centered (or user-centered) design is a design philosophy where technology is designed to accommodate the strengths and weaknesses of the human user, rather than forcing the user to adapt to the design (Wickens, Gordon & Liu, 1998). Part of the rationale behind this design philosophy is that human "design" is fixed; that is, the way people acquire, interpret, and act on information follows relatively predictable patterns. On the other hand, technology can often be constructed to compliment the user's fortes and foibles. As an additional bonus, human-centered designs, which conform to natural human performance patterns may minimize the amount and frequency of required training.

Testing The User Interface. One of the most important aspects of interface design is user testing (Wickens, Gordon & Liu, 1998). User testing is research aimed at testing the "fit" between the user and the technology. Such testing often uncovers unexpected problems with the system interface early enough in the design process that solutions can be built in to the design.

Modeling and simulation techniques are ideal for testing high technology systems such as the theoretical FCS. The highly automated nature of the FCS with distributed, robotic components means it would naturally fit in a virtual world. Whether users interact with the physical entities or merely computer models of the robots would probably be transparent to them. Not only the interface, but the distributed components themselves could be tested in virtual form before the actual hardware is constructed.

Issues Related to Workload

Workload, particularly the soldier's cognitive workload, may be a significant concern for the design of the FCS. One might expect that the tempo of combat would often keep soldiers operating near the limit of their cognitive capacity. Once the cognitive capacity of the soldier is exceeded, the human-automation system breaks down.

The distributed nature of FCS means that soldiers will have to monitor a number of separate entities. This will require soldier-operators to divide their attention between each of the FCS components. The more entities the soldiers monitor, the less attentional resources they will have available to direct any one component.

Conversely, previous research from aviation suggests a significant portion of the soldier's workload in FCS will most likely consist of monitoring the automation. Monitoring requires vigilance on the part of the soldier-user.

Unfortunately, vigilance necessitates sustained attention, which requires considerable cognitive resources (Wickens, 1992).

These conflicting requirements present FCS operators with a dilemma; the requirement for sustained attention divided over several entities – an impossible task considering human attentional resource limits. Therefore, designers may have to develop processes to help soldier-users manage the automation, such as using automated aids.

Therefore, one workload-related question which may be important to consider is; how many semi-autonomous entities will a soldier be able to control? The answer requires considering a surprising number of variables, such as the design of the interface, the amount of attentional resources required to control each entity, the number and type of concurrent tasks, the reliability of the systems, environmental stressors, and so forth.

Another factor which may affect workload is the level of automation. Complex automated systems may operate at several levels of automation (Billings, 1997) (See Table 1). The level of automation chosen by the operator may be a function of the status of the various automated components, that is, it may be based on partial system failures.

Table 1
Levels of Automation

Autonomous operation
Management by exception
Management by consent
Management by delegation
Shared control
Assisted manual control
Direct manual control

Theoretically, the level of automation of the various FCS components would affect not only the workload, but the type of workload, physical or cognitive. As the components shift from less automated to more automated, the soldier's task would shift from performing tasks personally to monitoring tasks performed by the automation.

Additional Research Considerations

Issues from aviation such as team performance, user confidence in automation, and dealing with automation failure may also have a significant impact on the usability of FCS. In addition, there are a number of other questions which might require investigation. The following list includes examples of possible additional research questions.

Input questions:

- What information does the soldier need to make good tactical decisions?
- How will the soldier acquire this information?
- Does the volume of information needed exceed the soldier's ability to process the information (information overload)?
- Can input be diversified across human sensory modalities, e.g. visual, auditory, and possibly tactile?
- How can displays be designed to complement human sensory and information processing?
- How well can soldiers integrate multiple sensor feeds from multiple positions and orientations?
- How could displays be designed to help the soldier appropriately integrate information to form an accurate visualization of the battlespace?

Throughput questions:

- Does the workload under expected combat conditions exceed human capabilities?
- How would workload be affected by partial system failures?
- What factors inhibit the soldier's ability to make good tactical decisions? Lack of time? Task overload? Stress?
- How can automation facilitate the soldier's decision-making?
- What decisions can be made by the automation (if any) and what must be made by the human? How does this relate to levels of automation?

Output questions:

- What methods can the soldier use to control off-board entities?
- Are multiple control methods necessary? Are they linked to levels of automation?
- What is the best human-human and human-automation authority structure?
- How will automation affect the human leadership structure?

CONCLUSION

As this paper indicates, a considerable body of knowledge exists concerning human-automation interaction. This information suggests human-automation interaction is a complex relationship which may influence the effectiveness of FCS. However, this existing knowledge can be used to develop models of soldier-FCS interaction. These models can be validated through user testing to address problems before the systems are fielded.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the contributions of Larry Meliza, U.S. Army Research Institute, whose valuable suggestions significantly enhanced this work. The ideas expressed in this paper are those of the authors and are not necessarily the official view of ARI or the U.S. Army.

REFERENCES

Barnett, J. S. (2001), Affects of training on user confidence in automation. *Dissertation Abstracts International-B, 61* (7), 3884 (University Microfilms International No. AAT 9977805).

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach*. Mahwah, NJ: Lawrence Ehrlbaum.

Bowers, C. A., Oser, R. L., Salas, E. & Cannon-Bowers, J. A. (1996). Team performance in automated systems. In R. Parasuraman and M. Mouloua (Eds.) *Automation and Human Performance: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum.

Defense Advanced Research Projects Agency [DARPA], 2000) *News Release, Office of Assistant Secretary of Defense May 9, 2000.* Retrieved from http://www.darpa.mil/fcs/news/news_release.htm

Eidelkind, M. A. & Papatonopoulos, S. A. (1997). Operator trust and task delegation: Strategies in semi-autonomous agent system. In M. Mouloua and J. Koonce (Eds.), *Human-Automation Interaction: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum.

Endsley, M. (1996). Automation and situation awareness. In R. Parasuraman and M. Mouloua (Eds.) *Automation and Human Performance: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum.

Endsley, M. (1997). Level of automation: Integrating humans and automated systems. *Proceedings of the Human Factors 41st Annual Meeting.* Santa Monica, CA: The Human Factors and Ergonomics Society.

Federation of American Scientists (2001). *Future combat system.* Retrieved from http://www.fas.org/man/dod-101/sys/land/fcs.htm.

Itaszek, J. R. (1989). Human operators in automated systems: The impact of active participation and communication. *Proceedings of the Human Factors and Ergonomics Society 33rd Annual Meeting*. Santa Monica, CA: The Human Factors Society.

Lee, J. D. & Moray, N. 1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies 40*, 153-184.

Morgan, B. B., Jr., Herschler, D. A., Weiner, E. L. & Salas, E. (1993). Implications of automation technology for aircrew coordination and performance. *Human/Technology Interaction in Complex Systems 6*, 105-136.

Mosier, K. L., Heers, S., Skitka, L. J. & Burdick, M. (1997). Patterns in the use of cockpit automation. In M. Mouloua and J. Koonce (Eds.), *Human-Automation Interaction: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum.

Mosier, K. L., Skitka, L. J. Heers, S., & Burdick, M. (1998). Automation bias: Decision-making and performance in high-tech cockpits. *International Journal of Aviation Psychology 8* (1), 47-63.

National Defense Panel (1997). *Transforming defense: National security in the 21st century.* Retrieved from http://www.dtic.mil/ndp/FullDoc2.pdf.

Parasuraman, R., Mouloua, M., Molloy, R. & Hilburn, B. (1996). Monitoring automated systems. In R. Parasuraman and M. Mouloua (Eds.) *Automation and Human Performance: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum.

Richardson, Doug (April, 2001) *UAV's and battlefield robots.* Retrieved from: http://www.atalink.co.uk/DSR/CLIENT/richards.htm

Sarter, N. B. (1991). The flight management system: Pilot's interaction with cockpit automation. *Proceedings of the 1991 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 1307-1310). Piscataway, NJ: IEEE.

Sarter, N. B. (1996). Cockpit automation: From quantity to quality, from individual pilots to multiple agents. In R. Parasuraman and M. Mouloua (Eds.) *Automation and Human Performance: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum.

Sarter, N. B. & Woods, D. D. (1997). Team play with a powerful and independent agent: Operational experiences and automation surprises on the Airbus A-320. *Human Factors 39*(4), 553-569.

Tsang, P. S. & Vidulich, M. A. (1989). Cognitive demands of automation in aviation. In R. S. Jenson, *Aviation Psychology*. Brookfield, VT: Gower Publishing.

U.S. Army Armor Center (2001). *Future combat systems* Retrieved from http:// 147.238.100.101/center/dfd/FCV.htm.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance* (2nd Ed.). New York: Harper Collins.

Wickens, C. D., Gordon, S. E. & Liu, Y. (1998). *An Introduction to Human Factors Engineering*. New York: Longman.

Wiener, E. L & Nagel, D. C. (1988). *Human Factors in Aviation*. San Diego, CA: Academic Press.