Editors: **Robert R. Hoffman, Jeffrey M. Bradshaw, and Ken Ford,** Florida Institute for Human and Machine Cognition, rhoffman@ihmc.us

Myths of Automation, Part 2: Some Very Human Consequences

Robert R. Hoffman, *Institute for Human and Machine Cognition*
John K. Hawley, *Army Research Laboratory*
Jeffrey M. Bradshaw, *Institute for Human and Machine Cognition*

During the major combat operations phase of the second Gulf War (Operation Iraqi Freedom—OIF, March and April 2003), US Army Patriot air and missile defense units were involved in two fratricide incidents. In the first, a British GR-4 Tornado was misclassified as an antiradiation missile and was subsequently engaged and destroyed. The second fratricide incident involved a Navy F/A-18 Hornet that was misclassified as a tactical ballistic missile and also was engaged and destroyed. Three flight crewmembers lost their lives in these incidents. OIF involved a total of 11 Patriot engagements by US units. Of these 11, nine resulted in successful tactical ballistic missile engagements; the other two were fratricides. Although significant in and of themselves, these fratricides opened the door for a unique look at the human performance problems introduced by increasing the complexity of technology and operations in a major weapons system.

A team from the Army Research Laboratory (ARL) began looking into the OIF Patriot fratricides and the more general issue of Patriot human-system performance, at the invitation of the then Ft. Bliss, Texas, Commanding General, Major General (MG) Michael A. Vane. Ft. Bliss was the site of the Army's Air Defense Artillery Center and School. In his own words, Vane was interested in operator vigilance and situation awareness as they relate to the performance using automated air defense battle management systems. He was particularly concerned by what he termed a "lack of vigilance" on the part of Patriot air battle management crews, along with an apparent "lack of cognizance" of what was being presented to them on situation displays, with a resulting "unwarranted trust in automation."

Framed in this way, the explanation for the fratricides involves blaming the human operators.

The ARL's Assessment

When ARL began looking at contributing factors, the first was the undisciplined insertion of more and more automation without regard for the downstream consequences for human performance, or any clear notion of how to anticipate and evaluate the downstream consequences. Undisciplined automation tends to define the operators' activities as responses to the operations of the automation. Otherwise, every function that can be automated is automated. Operators are left in the control loop merely to monitor the engagement process and to override that process only when it's determined that the weapon system's engagement logic isn't accurate. Research and operational experience indicate that this is a difficult role for operators to perform adequately.[1]

In the Patriot OIF fratricide case, undisciplined automation involved the following factors:

- *Unacknowledged system fallibilities*. This is a deficiency in the technology that's known but hasn't been satisfactorily resolved. For example, a series of Patriot operational tests going back to the 1980s indicated that the weapon system's automated engagement logic was subject to track misclassification problems. A misclassification occurred when the weapon system's category designation didn't match the track's actual status. The sources of automation unreliability weren't satisfactorily addressed during weapon system software upgrades, nor did information about these failure points find its way into operator training, air battle management doctrine, crew procedures, or unit standard operating procedures.

1541-1672/14/$31.00 © 2014 IEEE
Published by the IEEE Computer Society

- *Adherence to a technology-centric design philosophy*. System developers continued to pursue technology-centric solutions to automation reliability problems—such as increased use of artificial intelligence, automated non-cooperative target recognition, and improved Identification Friend or Foe (IFF) query systems.
- *Failure to keep the operators informed*. What made matters worse was that operators weren't informed about the problems in track classification and identification, or if they were informed, little if any effective responsive action was identified for them.
- *Blind faith*. Emboldened by Patriot's seeming success in engaging the Iraqi tactical ballistic missile threat during the First Gulf War, Patriot's organizational culture and command structure during OIF emphasized reacting quickly, engaging early, and trusting the weapon system without question.
- *Failure to train for expertise*. The cultural norm of blindly trusting the weapon system was exacerbated by the air defense community's traditional training practices, which were criticized in the Army's post-fratricide review as emphasizing rote drills rather than the "exercise of high-level judgment." The Patriot user community approached training for air battle operations in much the same manner as less knowledge- and skill-intensive tasks, such as system movement and setup. The emphasis during training was on mastering routines (crew drills) rather than critical thinking and adaptive problem solving. The existing training was inadequate because it was too short to produce necessary levels of operator competence. It was ill-focused because the training content didn't address critical operator or crew skills. And it was inappropriate because the instructional methods weren't suited to the job's skill content.
- *Inappropriate job assignment*. Negative results from the inadequacy of individual and crew training were made worse by the branch's methods of assigning personnel to air battle management crews. Traditional personnel assignment practices tend to place inexperienced personnel in key air battle management crew positions. Moreover, routine personnel administration practices usually rotate crewmembers out of battle staff positions and on to other jobs rather quickly. The result was that tactical crews were generally formed from a unit's newest and least-experienced personnel.

In sum, the ARL's assessment did not place blame for the OIF fratricides on the Patriot crews themselves.

Let's Not Blame the Operator

The ARL review argued that the roots of the crews' apparent performance shortcomings could be traced back to systemic problems resulting from decisions made years earlier by concept developers, software engineers, procedures developers, and others. Indeed, the OIF Patriot air battle management crews did exactly what they had been trained to do and what Patriot's command structure and culture had emphasized and reinforced: Trust the weapon system. Crews were actively discouraged from questioning the machine's recommendations. That strategy presumes that they were qualified to do so, but that was part of the problem.

In a weapon system such as Patriot, effective control means that operators—not the machine—are the ultimate decision makers in engagement decisions. Decisions about whether to shoot must be made by crews having:

- the technical potential for adequate situation understanding;
- the background to understand the significance of the information available to them; and
- the expertise to notice anomalies and adapt according to context, especially when this necessitates a departure from standardized—and even "mandated"—procedures.

The Army's post-fratricide board of inquiry remarked that the Patriot weapon system is too lethal to be placed in the hands of crews trained to a limited standard. The effects of lack of expertise continued to be manifest in subsequent operational tests involving Patriot software upgrades.

Painting Yourself into a Corner

The Patriot Missile weapon system is lethal. Hence, in the early days of Patriot, the Army opted against using the weapon system in an automatic mode. However, using Patriot against tactical ballistic missile threats necessitates the use of the automatic mode, because in-the-loop control is too slow. The operating concepts and training methods weren't adjusted to compensate for this major change.

Driven by technological opportunity and mission expansion, the role of the Patriot air battle management crew changed from that of a traditional operator to that of a supervisory controller whose primary role was to manage the subordinate automated engagement routines controlling the system. That is, the subordinate systems automatically close a control loop on the task or process. The crew is expected to monitor the technology for correct performance, and intermittently respond to system cues when necessary. As a result, control is indirectly exercised through automated engagement rather than directly exercised through traditional manual control. The supervising crew is thus "on" the control loop rather than "in" it, as is the case in traditional manual control. The term "on-the-loop" versus "in-the-loop" is quickly becoming the standard description to reflect this role change.

This change in terminology might appear minor, but it's significant for design, training, and operational practices. One of the myths of autonomy discussed in a previous article in the "Human-Centered Computing" department[2] is that when technology succeeds, human workload is reduced accordingly. However, the supervisory control approach that's increasingly used to manage "advanced" technology requires *more* human expertise, and training to a much *higher level of proficiency*.

We must consider abandoning the word "automation" entirely, as it has come to be a code word for "fewer humans" (seemingly less cost) and "fewer experts" (seemingly much less cost).

So, What Do We Do?

The ARL team's investigation report recommended that the Army reexamine the level of expertise required to operate a weapon system such as Patriot on the modern battlefield. In phrasing this recommendation, the ARL team deliberately chose to rely on the concept of "training for expertise" rather than just "training." Crews need to achieve a high level of proficiency to be able to think critically about emerging tactical situations, thus making use of extensive technical, tactical, and experiential knowledge. The Army's post-fratricide board of inquiry criticized Patriot training for emphasizing rote battle drills over critical thinking and problem solving. The lesson learned is that the extensive use of automation in battle command doesn't eliminate the need for operator expertise (see p. 2 of related article by John K. Hawley[3]).

To address the expertise issue, the ARL team partnered with the Air Defense School on a project intended to demonstrate what expertise-focused training for Patriot operators and crews would look like, how well it would work, and how it would be received. Expert job performers served as instructors, and the demonstration implemented a deliberate practice instructional model. This engages the trainees in work at "tough cases."[4] The demonstration was conducted using participants from an operational Patriot unit. Results from the demonstration project were positive with respect to the evaluation criteria.

Moving Forward?

The focus of training development within the Army must shift away from a preoccupation with training technology and equipment—analogous to the technology preoccupation in the acquisition arena. This shift will require a renewed focus on designing human-centered training systems that support the efficient acquisition of essential skills to high proficiency. So far, the Army has focused on the use of simulators—not a bad approach if it weren't for the fact that it's using the new technology to implement old training methods (see p. 3 of Hawley's other article[3]). In a 2001 report on training, the Defense Science Board warned of an increasing risk that "training failure will negate hardware promise" (p. 2).[5] The future will require more of our people to do more, new, and more complicated things. Unfortunately, old thinking tends to dominate training concept formulation, development, and conduct. We keep doing what we've always done. This is a frustrating reality, because the whys and hows underlying the development of expertise are generally known. The challenge going forward is applying this theory and practice in the contemporary training environment.

Patriot and other technology-intensive systems require considerable operator expertise for effective use. Research indicates that developing these levels of expertise requires several years of full-time effort just to reach the journeyman level. Much of this training must take place on the job in tactical units, but despite the best intentions, the Army's formal personnel system makes it difficult to keep operators and crews in one position long enough to reach the necessary levels of on-the-job competence. Inappropriate personnel practices can undo the best-laid training plans and practices.[4]

After Operation Desert Storm, the Army actually reduced the experience level of Patriot crews and cut back on training. To paraphrase what one officer explained to the ARL researchers, "We thought that automation did not require well-trained crews. I guess we were wrong."

And a consequence is that that once you have "drained the expertise out of the force" it can take a long time to replace it. This illustrates another of the myths of autonomy[2]: the erroneous idea that once achieved, full autonomy obviates the need for human-machine collaboration. Of course, it's true that when used correctly, technology can help increase overall work system performance. However, technology—at least our current technology—can't substitute for human expertise and adaptive capacity.

Even Broader Lessons

A good step forward would be to require that the subject of analysis, testing, and costing be the work system in its organizational context. Looking at the technology alone without crews and outside its dynamic organizational context paints a biased picture of a work system's eventual cost and performance.[6] This assertion comes not only from the tenets of human-centered computing but from within the subject matter itself. In their history of technological innovation in warfare, MacGregor Knox and Williamson Murray (pp. 176–177) stated that, "Technology did not simplify war, as contemporary superstition now claims: it made it exponentially more complex. Each new scientific development, each new weapons system, demanded fresh thought and ever-greater tactical, technical, and logistic expertise."[7]

Technology, now commonly referred to as "automation"—and in its more advanced forms, "autonomy"—has made contemporary work more cognitively complex. Complex systems theory and research[8–11] shows a broad consensus that intrinsic complexity can't be reduced. It's necessary to accept complexity—and increasing complexity—as a persistent and pervasive fact. And then deal with it. It can be deadly to attempt to avoid complexity by making reductive assumptions and attempting to implement simple, quick-fix solutions. These merely hide or ignore the root problem of complexity, resulting in deadly consequences, as the Patriot Missile case illustrates.

The increasing use of information processing technology in the workplace changes the nature of work and the skill, knowledge, and experience requirements of the people who perform that work.[12] Unit modernization involves far more than simply giving a unit new equipment and assuming that they somehow will use it according to the script.

Significant challenges are posed for humans by technology-intensive intelligent systems in complex macrocognitive work systems. Current concepts for the creation and procurement of large-scale intelligent systems still rely on training and personnel administration processes that were created in an earlier time for use with simpler technologies and less-complex application contexts. They must be modified to provide the highly skilled personnel that contemporary and emerging technologies now require.

Going forward, we must escape the myth that the complexity of cognitive work systems is reduced by injecting more automation (Myth 7).[2] We also must escape the myth that more automation guarantees reduced cognitive workload (Myth 6).[2] It's in the interest of governments to fully embrace human-centering considerations in the processes that are mandated for large-scale procurement of intelligent systems. It's crucial to fully embrace the notion of human-machine interdependence and acknowledge the necessary role of human expertise in the implementation and operation of intelligent systems. As Hawley[3] (p. 4) noted, *"It is a mistake to believe that desired performance levels will be achieved by widgets alone"* (Myth 3).[2] ■

Acknowledgment

The views expressed here are the authors' own and don't necessarily reflect those of the Army Research Laboratory.

References

1. D.D. Woods, and E. Hollnagel, *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering,* Taylor & Francis, 2006.
2. J.M. Bradshaw et al., "The Seven Deadly Myths of "Autonomous Systems," *IEEE Intelligent Systems*, vol. 28, no. 3, 2013, pp. 54–61.
3. J.K. Hawley, "Not by Widgets alone," *Armed Forces J.,* Feb. 2011; www.armedforcesjournal.com/not-by-widgets-alone.
4. R.R. Hoffman et al., *Accelerated Expertise: Training for High Proficiency in a Complex World,* Taylor and Francis/CRC Press, 2014.
5. J. Braddock and R. Chatham, *Training Superiority and Training Surprise*, final report of the Defense Science Board (DSB) Task Force on Training Superiority and Training Surprise, Office of the Undersecretary of Defense for Acquisition, Technology, and Logistics, 2001; www.acq.osd.mil/dsb/reports/ADA386579.pdf.
6. W. Zachary et al., "Human Total Cost of Ownership: The Penny Foolish Principle at Work," *IEEE Intelligent Systems*, vol. 22, no. 2, 2007, pp. 88–92.
7. M. Knox and W. Murray, "Conclusion: The Future behind Us," *The Dynamics of Military Revolution, 1300–2050*, Cambridge Univ. Press.
8. J.M. Carlson and J. Doyle, "Complexity and Robustness," *Proc. National Academy of Sciences*, vol. 99, supplement 1, 2002, pp. 2538–2545.
9. J.K. Hawley and A.L. Mares, "Human Performance Challenges for the Future Force: Lessons from Patriot after the Second Gulf War," *Designing Soldier Systems: Current Issues in Human Factors*, P. Savage-Knepshield, ed., 2012, pp. 3–34.
10. R.R. Hoffman and D.D. Woods, "Beyond Simon's Slice: Five Fundamental Tradeoffs that Bound the Performance of Macrocognitive Work Systems," *IEEE Intelligent Systems*, vol. 26, no. 6, 2011, pp. 67–71.
11. E. Hollnagel and D.D. Woods, *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*, Taylor & Francis, 2005.
12. F. Levy and R.J. Murnane, *The New Division of Labor: How Computers Are Creating the Next Job Market*, Princeton Univ. Press, 2012.

Robert R. Hoffman is a senior research scientist at the Institute for Human and Machine Cognition. Contact him at rhoffman@ihmc.us.

John K. Hawley is an engineering psychologist at the Army Research Laboratory Human Research and Engineering Directorate at Fort Bliss, Texas. Contact him at john.k.hawley.civ@mail.mil.

Jeffrey M. Bradshaw is a senior research scientist at the Institute for Human and Machine Cognition. Contact him at jbradshaw@ihmc.us.