HUMAN-CENTERED COMPUTING

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

The Seven Deadly Myths of "Autonomous Systems"

Jeffrey M. Bradshaw, Robert R. Hoffman, Matthew Johnson, and David D. Woods

In this article, we explore some widespread misconceptions surrounding the topic of "autonomous systems." The immediate catalyst for this article, like a previous article that appeared in this department,[1] is a recent US Defense Science Board (DSB) Task Force Report on "The Role of Autonomy in DoD Systems." This report affords an opportunity to examine the concept of autonomous systems in light of the new DSB findings. This theme will continue in a future column, in which we'll outline a constructive approach to design autonomous capabilities based on a human-centered computing perspective. But to set the stage, in this essay we bust some "myths" of autonomy.

Myths of Autonomy

The reference in our title to the "seven deadly myths" of autonomous systems alludes to the seven deadly sins. The latter were so named not only because of their intrinsic seriousness but also because the commission of one of them would engender further acts of wrongdoing. As designers conceive and implement what are commonly (but mistakenly) called autonomous systems, they have succumbed to myths of autonomy that are not only damaging in their own right but are also damaging by their continued propagation—that is, because they engender a host of other serious misconceptions and consequences. Here, we provide reasons why each of these myths should be called out and cast aside.

Myth 1: "Autonomy" is unidimensional. There is a myth that autonomy is some single thing and that everyone understands what it is. However, the word is employed with different meanings and intentions.[2] "Autonomy" is straightforwardly derived from a combination of Greek terms signifying self (*auto*) governance (*nomos*), but it has two different senses in everyday usage. In the first sense, it denotes self-sufficiency—the capability of an entity to take care of itself. This sense is present in the French term *autonome* when, for example, it's applied to an individual who is capable of independent living. The second sense refers to the quality of self-directedness, or freedom from outside control, as we might say of a political district that has been identified as an "autonomous region."

The two different senses affect the way autonomy is conceptualized, and influence tacit claims about what "autonomous" machines can do. For example, in a chapter from a classic volume on agent autonomy, Sviatoslav Brainov and Henry Hexmoor[3] emphasize how varying degrees of autonomy serve as a relative measure of self-directedness—that is, independence of an agent from its physical environment or social group. On the other hand, in the same volume Michael Luck and his colleagues,[4] unsatisfied with defining autonomy in such relative terms, argue that the self-generation of goals should be the defining characteristic of autonomy. Such a perspective characterizes the machine in absolute terms that reflect the belief of these researchers in autonomy as self-sufficiency.

It should be evident that independence from outside control doesn't entail the self-sufficiency of an autonomous machine. Nor do a machine's autonomous capabilities guarantee that it will be allowed to operate in a self-directed manner. In fact, human-machine systems involve a dynamic balance of self-sufficiency and self-directedness. We will now elaborate on some of the subtleties relating to this balance.

Figure 1 illustrates some of the challenges faced by designers of machine capabilities. A major motivation for such capabilities is to reduce the burden on human operators by increasing a machine's self-sufficiency to the point that it can be trusted to operate in a self-directed manner. However, when the self-sufficiency of the machine capabilities is seen as inadequate for the circumstances, particularly in situations where the consequences

of error may be disastrous, it is common to limit the self-directedness of the machine. For example, in such circumstances a human may take control manually, or an automated policy may come into play that prevents the machine from doing harm to itself or others through faulty actions. Such a scenario brings to mind early NASA Mars rovers whose capabilities for autonomous action weren't fully exercised due to concerns about the high cost of failure. Because their advanced capabilities for autonomous action weren't fully trusted, NASA decided to micromanage the rovers through a sizeable team of engineers. This example also highlights that the capabilities machines have for autonomous action interact with the responsibility for outcomes and delegation of authority. Only people are held responsible for consequences (that is, only people can act as problem holders) and only people decide on how authority is delegated to automata.[5]

When self-directedness is reduced to the point that the machine is prevented from fully exercising its capabilities (as in the Mars rover example), the result can be described as *under-reliance* on the technology. That is, although a machine may be sufficiently competent to perform a set of actions in the current situation, human practice or policy may prevent it from doing so. The flipside of this error is to allow a machine to operate too freely in a situation that outstrips its capabilities (such as high self-directedness with low self-sufficiency). This error can be described as *over-trust*.

In light of these observations, we can characterize the primary challenge for the designers of autonomous machine capabilities as a matter of moving upward in a 45-degree diagonal on Figure 1—increasing

machine capabilities while maintaining a (dynamic) balance between self-directedness and self-sufficiency. However, even when the self-directedness and self-sufficiency of autonomous capabilities are balanced appropriately for the demands of the situation, humans and machines working together frequently encounter potentially debilitating problems relating to insufficient observability or understandability (upper right quadrant of Figure 1). When highly autonomous machine capabilities aren't well understood by people or other machines working with them, work effectiveness suffers.[7,8]

Whether human or machine, a "team player" must be able to observe, understand, and predict the state and actions of others.[9] Many examples can be found in the literature of inadequate observability and understandability as a problem in human-machine interaction.[5,10] The problem with what David Woods calls "strong silent automation" is that it fails to communicate effectively those things that would allow humans to work interdependently with it—signals that allow operators to predict, control, understand, and anticipate what the machine is or will be doing. As anyone who has wrestled with automation can attest, there's nothing worse than a so-called smart machine that can't tell you what it's doing, why it's doing something, or when it will finish. Even more frustrating—or dangerous—is a machine that's incapable of responding to human direction when something (inevitably) goes wrong.

To sum up our discussion of the first myth: First, "autonomy" isn't a unidimensional concept—it's more useful to describe autonomous systems *at least* in terms of the two dimensions of self-directedness and self-sufficiency. Second, aspects of
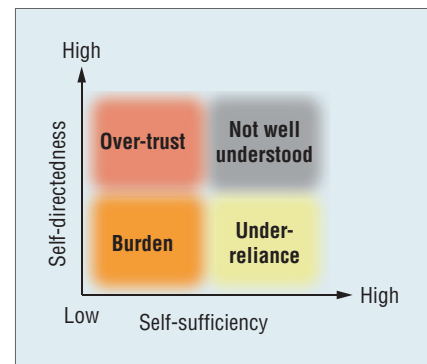
Figure 1. Challenges faced by designers of autonomous machine capabilities.[6] When striving to maintain an effective balance between self-sufficiency and self-directedness for highly capable machines, designers encounter the additional challenge of making the machine understandable.

self-directedness and self-sufficiency must be balanced appropriately. Third, to maintain desirable properties of human-machine teamwork, particularly when advanced machine capabilities exhibit a significant degree of competence and self-governance, team players must be able to communicate effectively those aspects of their behavior that allow others to understand them and to work interdependently with them.

Myth 2: The conceptualization of "levels of autonomy" is a useful scientific grounding for the development of autonomous system roadmaps. Since we've just argued for discarding the myth that autonomy is unidimensional, we shouldn't have to belabor the related myth that machine autonomy can be measured on a single ordered scale of increasing levels. However, because this second myth is so pervasive, it merits separate discussion.

A recent survey of human-robot interaction concluded that "perhaps the most strongly human-centered application of the concept of autonomy is in the notion of level of autonomy."[11] However, one of the most striking recommendations of the DSB report on the role of autonomy is its

recommendation that the Department of Defense (DoD) should abandon the debate over definitions of levels of autonomy.[12] The committee received input from multiple organizations on how some variation of definitions across levels of autonomy could guide new designs. The retired flag officers, technologists, and academics on the task force overwhelmingly and unanimously found the definitions irrelevant to the real problems, cases of success, and missed opportunities for effectively utilizing increases in autonomous capabilities for defense missions.

The two paragraphs (from pp. 23–24) summarizing the DSB's rationale for this recommendation are worth citing verbatim:

> An … unproductive course has been the numerous attempts to transform conceptualizations of autonomy made in the 1970s into developmental roadmaps. … Sheridan's taxonomy [of levels of automation] … is often incorrectly interpreted as implying that autonomy is simply a delegation of a complete task to a computer, that a vehicle operates at a single level of autonomy and that these levels are discrete and represent scaffolds of increasing difficulty. Though attractive, the conceptualization of levels of autonomy as a scientific grounding for a developmental roadmap has been unproductive. … The levels served as a tool to capture what was occurring in a system to make it autonomous; these linguistic descriptions are not suitable to describe specific milestones of an autonomous system. … Research shows that a mission consists of dynamically changing functions, many of which can be executing concurrently as well as sequentially. Each of these functions can have a different allocation scheme to the human or computer at a given time.[12]

There are additional reasons why the levels of automation notion are problematic.

First, functional differences matter. The common understanding of the levels assumes that significantly different kinds of work can be handled equivalently (such as task work and teamwork; reasoning, decisions, and actions). This reinforces the erroneous notion that "automation activities simply can be substituted for human activities without otherwise affecting the operation of the system."[13]

Second, levels aren't consistently ordinal. It isn't always clear whether a given action should be characterized as "lower" or "higher" than another on the scale of autonomy. Moreover, a given machine capability operating in a specific situation may simultaneously be "low" on self-sufficiency while being "high" on self-directedness.[7]

Third, autonomy is relative to the context of activity. Functions can't be automated effectively in isolation from an understanding of the task, the goals, and the context.

Fourth, levels of autonomy encourage reductive thinking. For example, they facilitate the perspective that activity is sequential when it's actually simultaneous.[14]

Fifth, the concept of levels of autonomy is insufficient to meet both current and future challenges. This was one of the most significant findings of the DoD report. For example, many challenges facing human-machine interaction designers involve teamwork rather than the separation of duties between the human and the machine.[9] Effective teamwork involves more than effective task distribution; it looks for ways to support and enhance each member's performance[6]—this need isn't addressed by the levels of autonomy conceptualization.

Sixth, the concept of levels of autonomy isn't "human-centered." If it were, it wouldn't force us to recapitulate the requirement that technologies be useable, useful, understandable, and observable.

Last, the levels provide insufficient guidance to the designer. The challenge of bridging the gap from cognitive engineering products to software engineering results is one of the most daunting of current challenges and the concept of levels of autonomy provides no assistance in dealing with this issue.

Myth 3: Autonomy is a widget. The DSB report points (on p. 23) to the fallacy of "treating autonomy as a widget":

> The competing definitions for autonomy have led to confusion among developers and acquisition officers, as well as among operators and commanders. The attempt to define autonomy has resulted in a waste of both time and money spent debating and reconciling different terms and may be contributing to fears of unbounded autonomy. The definitions have been unsatisfactory because they typically try to express autonomy as a widget or discrete component, rather than a capability of the larger system enabled by the integration of human and machine abilities.[12]

In other words, autonomy isn't a discrete property of a work system, nor is it a particular kind of technology; it's an idealized characterization of observed or anticipated interactions between the machine, the work to be accomplished, and the situation. To the degree that autonomy is actually realized in practice, it's through the combination of these interactions.

The myth of autonomy as a widget engenders the misunderstandings implicit in the next myth.

Myth 4: Autonomous systems are autonomous. Strictly speaking, the

term "autonomous system" is a misnomer. No entity—and, for that matter, no person—is capable enough to be able to perform competently in every task and situation. On the other hand, even the simplest machine can seem to function "autonomously" if the task and context are sufficiently constrained. A thermostat exercises an admirable degree of self-sufficiency and self-directedness with respect to the limited tasks it's designed to perform through the use of a simple form of automation (at least until it becomes miscalibrated).

The DSB report wisely observes that "… there are no fully autonomous systems just as there are no fully autonomous soldiers, sailors, airmen, or Marines. … Perhaps the most important message for commanders is that all machines are supervised by humans to some degree, and the best capabilities result from the coordination and collaboration of humans and machines" (p. 24).[12]

Machine designs are always created with respect to a context of design assumptions, task goals, and boundary conditions. At the boundaries of the operating context for which the machine was designed, maintaining adequate performance might become a challenge. For instance, a typical home thermostat isn't designed to work as an outdoor sensor in the significantly subzero climate of Antarctica. Consider also the work context of a Navy Seal whose job it is to perform highly sensitive operations that require human knowledge and reasoning skills. A Seal doing his job is usually thought of as being highly autonomous. However, a more careful examination reveals his interdependence with other members of his Seal team to conduct team functions that can't be performed by a single individual, just as the team is interdependent with the overall Navy mission

and with the operations of other co-located military or civilian units.

What's the result of belief in this fourth myth? People in positions of responsibility and authority might focus too much on autonomy-related problems and fixes while failing to understand that self-sufficiency is always relative to a situation. Sadly, in most cases machine capabilities are not only relative to a set of *predefined* tasks and goals, they are relative to a set of *fixed* tasks and goals. A software system might perform gloriously without supervision in circumstances within its competence envelope (itself a reflection of the designer's intent), but fail miserably when the context changes to some circumstance that pushes the larger work system over the edge.[15] Although some tasks might be worked with high efficiency and accuracy, the potential for disastrous fragility is ever present.[16] Speaking of autonomy without adequately characterizing assumptions about how the task is embedded in the situation is dangerously misguided.

Myth 5: Once achieved, full autonomy obviates the need for human-machine collaboration. Much of the early research on autonomous systems was motivated by situations in which autonomous systems were required to replace humans, in theory minimizing the need for considering the human aspects of such solutions. For example, one of the earliest high-consequence applications of sophisticated agent technologies was in NASA's Remote Agent Architecture, designed to direct the activities of unmanned spacecraft engaged in distant planetary exploration.[17] The Remote Agent Architecture was expressly designed for use in human-out-of-the-loop situations where response latencies in the transmission of round-trip

control sequences from earth would have impaired a spacecraft's ability to respond to urgent problems or to take advantage of unexpected science opportunities.

Since those early days, most autonomy research has been pursued in a technology-centric fashion, as if full machine autonomy—complete independence and self-sufficiency—were a holy grail. A primary, ostensible reason for the quest is to reduce manning needs, since salaries are the largest fraction of the costs of sociotechnical systems. An example is the US Navy's "Human Systems Integration" program, initially founded on a belief that an increase in autonomous machine capabilities (typically developed without adequate consideration for the complexities of interdependence in mixed human-machine teams) would enable the Navy to crew large vessels with smaller human complements. However, reflection on the nature of human work reveals the shortsightedness of such a singular and short-term focus: What could be more troublesome to a group of individuals engaged in dynamic, fast-paced, real-world collaboration coping with complex tasks and shifting goals than a colleague who is perfectly able to perform tasks alone but lacks the expertise required to coordinate his or her activities with those of others?

Of course, there are situations where the goal of minimizing human involvement with autonomous systems can be argued effectively—for example, some jobs in industrial manufacturing. However, it should be noted that virtually all of the most challenging deployments of autonomous systems to date—such as military unmanned air vehicles, NASA rovers, unmanned underwater vehicles, and disaster inspection robots—have involved people in crucial

roles where expertise is a must. Such involvement hasn't been merely to make up for the current limitations on machine capabilities, but also because their jointly coordinated efforts with humans were—or should have been—intrinsically part of the mission planning and operations itself.

What's the result of belief in this myth? Researchers and their sponsors begin to assume that "all we need is more autonomy." This kind of simplistic thinking engenders the even more grandiose myth that human factors can be avoided in the design and deployment of machines. Careful consideration will reveal that, in addition to more machine capabilities for *task work*, there's a need for the kinds of breakthroughs in human-machine *teamwork* that would enable autonomous systems not merely to do things *for* people, but also to work together *with* people and other systems. This capacity for teamwork, not merely the potential for expanded task work, is the inevitable next leap forward required for more effective design and deployment of autonomous systems operating in a world full of people.[18]

Myth 6: As machines acquire more autonomy, they will work as simple substitutes (or multipliers) of human capability. The concept of automation began with the straightforward objective of replacing whenever feasible any tedious, repetitive, dirty, or dangerous task currently performed by a human with a machine that could do the same task better, faster, or cheaper. This was a core concept of the Industrial Revolution. The entire field of human factors emerged circa World War 1 in recognition of the need to consider the human operator in industrial design. Automation became one of the first issues to attract the notice of cyberneticists and human factors researchers during and immediately after World War II. Pioneering researchers attempted to systematically characterize the general strengths and weaknesses of humans and machines. The resulting discipline of "function allocation" aimed to provide a rational means of determining which system-level functions should be carried out by humans and which by machines.

Obviously, the suitability of a particular human or machine to take on a particular task will vary over time and in different situations. Hence, the concepts of adaptive or dynamic function allocation and adjustable autonomy emerged with the hope that shifting responsibilities between humans and machines would lead to machine and work designs more appropriate for the emerging sociotechnical workplace.[2] Of course, certain tasks, such as those requiring sophisticated judgment, couldn't be shifted to machines, and other tasks, such as those requiring ultra-precise movement, couldn't be performed by humans. But with regard to tasks where human and machine capabilities overlapped—the area of variable task assignment—software-based decision-making schemes were proposed to allow tasks to be allocated according to the potential performer's availability.

Over time, it became plain to researchers that things weren't this simple. For example, many functions in complex systems are shared by humans and machines; hence, the need to consider synergies and goal conflicts among the various performers of joint actions. Function allocation isn't a simple process of transferring responsibilities from one component to another. When system designers automate a subtask, what they're really doing is performing a type of task distribution and, as such, have introduced novel elements of interdependence within the work system.[7] This is the lesson to be learned from studies of the "substitution myth,"[13] which conclude that reducing or expanding the role of automation in joint human-machine systems may change the nature of interdependent and mutually adapted activities in complex ways. To effectively exploit the capabilities that automation provides (versus merely increasing automation), the task work—and the interdependent teamwork it induces among players in a given situation—must be understood and coordinated as a whole.

It's easy to fall prey to the fallacy that automated assistance is a simple substitute or multiplier of human capability because, from the point of view of an outsider observing the assisted humans, it seems that—in successful cases, at least—the people are able to perform the task better or faster than they could without help. In reality, however, help of whatever kind doesn't simply enhance our ability to perform the task: it changes the nature of the task.[13,19] To take a simple example, the use of a computer rather than a pencil to compose a document can speed up the task of writing an essay in some respects, but sometimes can slow it down in other respects—for example, when electrical power goes out. The essential point is that it requires a different configuration of human skills. Similarly, a robot used to perform a household task might be able to do many things "on its own," but this doesn't eliminate the human's role, it changes that role. The human responsibility is now the cognitive task of goal setting, monitoring, and controlling the robot's progress (or regress).[16]

Increasing the autonomy of autonomous systems requires different kinds of human expertise and not always fewer humans. Humans and artificial

Table 1. Putative benefits of automation versus actual experience.[21]

Putative benefit	Real complexity
Increased performance is obtained from "substitution" of machine activity for human activity.	Practice is transformed; the roles of people change; old and sometimes beloved habits and familiar features are altered—the *envisioned world problem*.
Frees up human by offloading work to the machine.	Creates new kinds of cognitive work for the human, often at the wrong times; every automation advance will be exploited to require people to do more, do it faster, or in more complex ways—the *law of stretched systems*.
Frees up limited attention by focusing someone on the correct answer.	Creates more threads to track; makes it harder for people to remain aware of and integrate all of the activities and changes around them—with *coordination costs, continuously*.
Less human knowledge is required.	New knowledge and skill demands are imposed on the human and the human might no longer have a sufficient context to make decisions, because they have been left out of the loop—*automation surprise*.
Agent will function autonomously.	Team play with people and other agents is critical to success—*principles of interdependence*.
Same feedback to human will be required.	New levels and types of feedback are needed to support peoples' new roles—with *coordination costs, continuously*.
Agent enables more flexibility to the system in a generic way.	Resulting explosion of features, options, and modes creates new demands, types of errors, and paths toward failure—*automation surprises*.
Human errors are reduced.	Machines, humans, and macrocognitive work systems are fallible; errors are therefore systemic; new problems are associated with human-machine coordination breakdowns; machines now obscure information necessary for human decision making—*principles of complexity*.

agents are two disparate kinds of entities that exist in very different sorts of worlds. Humans have rich knowledge about the world that they're trying to understand and influence, while machines are much more limited in their understanding of the world that they model and affect. This isn't a matter of distinguishing ways that machines can compensate for things that humans are bad at. Rather, it's a matter of characterizing interdependence: things that machines are good at and ways in which they depend on humans (and other agents) in joint activity; and things that humans are good at and ways in which they depend on the machines (and other humans).[20]

For the foreseeable future this fundamental asymmetry, or duality, will remain. The brightest machine agents will be limited in the generality, if not the depth, of their inferential, adaptive, social, and sensory capabilities. Humans, though fallible, are functionally rich in reasoning strategies and their powers of observation, learning, and sensitivity to context. These are the things that make adaptability and resilience of work systems possible. Adapting to appropriate mutually

interdependent roles that take advantage of the respective strengths of humans and machines—and crafting natural and effective modes of interaction—are key challenges for technology—not merely the creation of increasingly capable widgets.

What's the result of belief in the myth of machines as simple multipliers of human ability? Because design approaches based on this myth don't adequately take into consideration the significant ways in which the introduction of autonomous capabilities can change the nature of the work itself; they lead to "clumsy automation." And trying to solve this problem by adding more poorly designed autonomous capabilities is, in effect, adding more clumsy automation onto clumsy automation, thereby exacerbating the problem that the increased autonomy was intended to solve.

Myth 7: "Full autonomy" is not only possible, but is always desirable. In refutation of the substitution myth, Table 1 contrasts the putative benefits of automated assistance with the empirical results. Ironically, even when

technology succeeds in making tasks more efficient, the human workload isn't reduced accordingly. David Woods and Eric Hollnagel[5] summarized this phenomenon as the law of stretched systems: "every system is stretched to operate at its capacity; as soon as there is some improvement, for example in the form of new technology, it will be exploited to achieve a new intensity and tempo of activity."

As Table 1 shows, the decision to increase the role of automation in general, and autonomous capabilities in particular, is one that should be made in light of its complex effects along a variety of dimensions. In this article, we've tried to make the case that full autonomy, the simplistic sense in which the term is usually employed, is barely *possible*. This table summarizes the reasons why increased automation isn't always *desirable*.

Although continuing research to make machines more active, adaptive, and functional is essential, the point of increasing such proficiencies

isn't merely to make the machines more *independent* during times when unsupervised activity is desirable or necessary (autonomous), but also to make them more capable of sophisticated *interdependent* activity with people and other machines when such is required (teamwork). Research in joint activity highlights the need for autonomous systems to support not only the fluid orchestration of task handoffs among people and machines, but also combined participation on shared tasks requiring continuous and close interaction (coactivity).[6,9] Indeed, in situations of simultaneous human-agent collaboration on shared tasks, people and machines might be so tightly integrated in the performance of their work that interdependence is a continuous phenomenon, and the very idea of task handoffs becomes incongruous. We see this, for example, in the design of work systems to support cyber sensemaking, that aim to combine the efforts of human analysts with software agents in understanding, anticipating, and responding to unfolding events in near real-time.[22]

The points mentioned here, like the findings of the DSB, focus on how to make effective use of the expanding power of machines. The myths we've discussed lead developers to introduce new machine capabilities in ways that predictably lead to unintended negative consequences and user-hostile technologies. We need to discard the myths and focus on developing coordination and adaptive mechanisms that turn platform capabilities into new levels of mission effectiveness—enabled through genuine human-centeredness. In complex and domains characterized by uncertainty, machines that are merely capable of performing independent work aren't

enough. Instead, we need machines that are also capable of working interdependently.[6] We commend the thoughtful work of the DSB in recognizing and exemplifying some of the significant problems caused by the seven deadly myths of autonomy, and hope these and similar efforts will lead all of us to sincere repentance and reformation.◻

References

1. R.R. Hoffman et al., "Trust in Automation," *IEEE: Intelligent Systems*, vol. 28, no. 1, pp. 84–88.
2. J.M. Bradshaw et al., "Dimensions of Adjustable Autonomy and Mixed-Initiative Interaction." *Agents and Computational Autonomy: Potential, Risks, and Solutions*, LNCS, vol. 2969, Springer-Verlag, 2004, pp. 17–39.
3. S. Brainov and H. Hexmoor, "Quantifying Autonomy," *Agent Autonomy*, Kluwer, 2002, pp. 43–56.
4. M. Luck, M. D'Inverno, and S. Munroe, "Autonomy: Variable and Generative," *Agent Autonomy*, Kluwer, 2002, pp. 9–22.
5. D.D. Woods and E. Hollnagel, *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*, Taylor & Francis, 2006, chapter 11.
6. M. Johnson et al., "Autonomy and Interdependence in Human-Agent-Robot Teams." *IEEE Intelligent Systems*, vol. 27, no. 2, 2012, pp. 43–51.
7. M. Johnson et al., "Beyond Cooperative Robotics: The Central Role of Interdependence in Coactive Design," *IEEE: Intelligent Systems*, vol. 26, no. 3, 2011, pp. 81–88.
8. D.D. Woods and E.M. Roth, "Cognitive Systems Engineering," *Handbook of Human-Computer Interaction*, North-Holland, 1988.
9. G. Klein et al., "Ten Challenges for Making Automation a 'Team Player' in Joint Human-Agent Activity," *IEEE Intelligent Systems*, vol. 19, no. 6, 2004, pp. 91–95.

10. D.A. Norman, "The 'Problem' of Automation: Inappropriate Feedback and Interaction, Not 'Over-Automation.'" *Philosophical Trans. Royal Soc. of London B*, vol. 327, 1990, pp. 585–593.
11. M.A. Goodrich and A.C. Schultz, "Human-Robot Interaction: A Survey," *Foundations and Trends in Human-Computer Interaction*, vol. 1, no. 3, 2007, pp. 203–275.
12. R. Murphy et al., *The Role of Autonomy in DoD Systems*, Defense Science Board Task Force Report, July 2012, Washington, DC.
13. K. Christofferson and DD. Woods, "How to Make Automated Systems Team Players," *Advances in Human Performance and Cognitive Engineering Research*, vol. 2, Elsevier Science, 2002, pp. 1–12.
14. P.J. Feltovich et al., "Keeping It Too Simple: How the Reductive Tendency Affects Cognitive Engineering." *IEEE Intelligent Systems*, vol. 19, no. 3, 2004, pp. 90–94.
15. 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.
16. 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*, Ashgate, 2012, pp. 3–34.
17. N. Muscettola et al.. "Remote Agent: To Boldly Go Where No AI System Has Gone Before," *Artificial Intelligence*, vol. 103, nos. 1–2, 1998, pp. 5-48.
18. J.M. Bradshaw et al., "Introduction to Special Issue on Human-Agent-Robot Teamwork (HART)," *IEEE Intelligent Systems*, vol. 27, no. 2, 2012, pp. 8–13.
19. D.A. Norman, "Cognitive Artifacts," *Designing Interaction: Psychology at the Human-Computer Interface*, Cambridge Univ. Press.1992, pp. 17–38.

20. R.R. Hoffman et al., "A Rose by Any Other Name … Would Probably Be Given an Acronym," *IEEE Intelligent Systems*, vol. 17, no. 4, 2002, pp. 72–80.

21. N. Sarter, D.D. Woods, and C.E. Billings, "Automation Surprises," *Handbook of Human Factors/Ergonomics,* 2nd ed., John Wiley, 1997.

22. J.M. Bradshaw et al., "Sol: An Agent-Based Framework for Cyber Situation Awareness," *Künstliche Intelligenz*, vol. 26, no. 2, 2012, pp. 127–140.

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

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

David D. Woods is a professor at The Ohio State University in the Institute for Ergonomics. Contact him at woods.2@osu.edu.

Matthew Johnson is a research scientist at the Florida Institute for Human and Machine Cognition. Contact him at mjohnson@ihmc.us.

cn *Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.*