cal structures and functions (*9–11*) . Instead, he besmirches the entire enterprise of political psychology, perpetuating canards from the right-wing blogosphere and lazy, empirically unsubstantiated accusations of "liberal bias." For example, Shermer writes:

> Why are people conservative? Why do people vote Republican? The questions are typically posed without even a whiff of awareness of the inherent bias in asking it in this manner—that because Democrats are so indisputably right and Republicans so unquestionably wrong, conservatism must be a mental disease, a flaw in the brain, a personality disorder that leads to cognitive malfunctioning. Much as medical scientists study cancer in order to cure the disease, liberal political scientists study political attitudes and voting behavior in order to cure people of the cancer of conservatism.

In passages such as this, Shermer is not merely hyperbolic, inflammatory, and wrong about the specifics of the scientific articles he purports to critique. (One doubts he even read them.) By resorting to ideological deconstruction and essentially ad hominem forms of attack, Shermer violates his own intellectual standards—succumbing to the tendency, which he scorns in others, to reject out of hand scientific findings that might be experienced as disagreeable. Belief-dependent realism, indeed.

Shermer ought to know better, but he is enabled (and led considerably astray) by Jonathan Haidt, whose non–peer-reviewed Internet provocation "What Makes People Vote Republican?" (*12*) provides the only data Shermer considers and, at the same time, a title to which he can object. What happened to the relentless thirst for empirical evidence and the evaluation of such evidence according to rigorous, established scientific criteria? When push comes to shove—as it often does with politics—Shermer sets the evidence aside and trades in stereotypical assumptions about the ideologies and personal backgrounds of the investigators. Consequently, the origins and dynamics of political beliefs shall remain an unsolved mystery to the book's readers.

The broader point, which is crucial to the future success of the social and behavioral sciences, is not that scientists themselves are somehow immune to cognitive or other sources of bias. It is that the scientific community is and should be ruthlessly committed to evaluating claims and settling disputes through the inspection and analysis of empirical data and through meaningful discussion and debate about how to properly interpret those data, using agreed-upon methodological standards—and not through ideological deconstruction or all too convenient allegations of bias. The politics chapter is therefore not only unscientific, it is anti-scientific.

Let us end on a more upbeat note. Shermer has done much to raise public awareness of the importance of scientific research and to confront Holocaust deniers, 9/11 "truthers," and others who stubbornly resist logic and evidence. His general commitment to science is appreciated by many—and rightly so. In tackling "why people believe things, full stop," Shermer has bitten off more than even he can chew. Nevertheless, many readers will learn something from the material that he has taken the time to actually digest. The challenge posed, both directly and indirectly, by this uneven book is to discern staunch self-confidence—whether it belongs to the author or the believers he is in search of—that is appropriate and justified from that which is not. In *The Believing Brain*, Shermer does not really try to explain why some people hold truer beliefs than others. But the difference between science and other human pursuits suggests that there may be more than one way of believing.

References

1. C. K. Morewedge, J. Preston, D. M. Wegner, *J. Pers. Soc. Psychol.* **93**, 1 (2007).
2. K. Gray, D. M. Wegner, *Pers. Soc. Psychol. Rev.* **14**, 7 (2010).
3. A. H. Eagly, S. Chaiken, *The Psychology of Attitudes* (Harcourt Brace Jovanovich, Fort Worth, TX, 1993).
4. J. C. Turner, *Social Influence* (Open Univ. Press, Buckingham, UK, 1991)
5. Z. Kunda, *Psychol. Bull.* **108**, 480 (1990).
6. J. T. Jost, J. Glaser, A. W. Kruglanski, F. J. Sulloway, *Psychol. Bull.* **129**, 339 (2003).
7. A. S. Gerber, G. A. Huber, D. Doherty, C. M. Dowling, S. E. Ha, *Am. Polit. Sci. Rev.* **104**, 111 (2010).
8. J. T. Jost, C. M. Federico, J. L. Napier, *Annu. Rev. Psychol.* **60**, 307 (2009).
9. D. M. Amodio, J. T. Jost, S. L. Master, C. M. Yee, *Nat. Neurosci.* **10**, 1246 (2007).
10. R. Kanai, T. Feilden, C. Firth, G. Rees, *Curr. Biol.* **21**, 677 (2011).
11. D. R. Oxley *et al.*, *Science* **321**, 1667 (2008).
12. www.edge.org/3rd_culture/haidt08/haidt08_index.html.

10.1126/science.1209161

COMPUTER SCIENCE

Canny Minds and Uncanny Questions

Greg Wayne[1] and Alex Pasternack[2]

The humans didn't stand much of a chance. Even before the Man-Made Minds discussion got under way at New York's World Science Festival in early June, it was clear that the show-stealer wouldn't be the panel of leading artificial intelligence (AI) researchers or even the doe-eyed robot Kismet, which graced the event poster. All eyes were on the flat-screen television perched on one corner of the stage, its animation pulsing in apparent anticipation.

Watson, IBM's question-answering machine that in March triumphed over human Jeopardy! champions, is actually housed in a set of computers that take up the space of about 10 refrigerators and guzzle some 80

[1]Department of Neuroscience, Columbia University, New York, NY 10032, USA. E-mail: gdw2104@columbia.edu [2]www.motherboard.tv. E-mail: pasternack@post.harvard.edu

kilowatts of electricity. For principal investigator David Ferrucci, it would have been impossible to bring the machine in anything but a stripped-down version. But any doubts about Watson's influence were terminated when the panelists gripped their buzzers for a pick-up game of Jeopardy! Hod Lipson, a roboticist at Cornell University, demurred, "This is the first time I'm feeling the fear of AI."

What was there to be afraid of, especially for a scientist who himself builds "self-aware" robots? The question lingered beneath the discussion, guided by television and radio correspondent Faith Salie across a smattering of long-standing philosophical questions. If a robot simulates human behavior, is it sentient? Should robots have human rights? Could an artificial intelligence know its own intellectual limitations? Will a technological singularity bring about the end of the human race? Could machines attain consciousness, and if so, would that trivialize what it means to be human? Why are some robots cute and others downright creepy?

Since its christening at a conference at Dartmouth in 1956, the field of AI has deployed a variety of methodologies in its struggle against two more basic questions: How can a computer approximate intelligence? And to what end? The panelists reflected the field's ongoing divisions: symbolic AI, which seeks to reason about natural language and logical expressions; machine learning, which largely uses statistical reasoning to find patterns in clouds of data; and biologically inspired AI, which creates reactive robots that model simple animal behavior. Over and above these divisions is a larger tension between those seeking merely intelligent machines ("weak AI") and those who believe machines could rival the sentience of humans ("strong AI").

Man-Made Minds
Living with Thinking Machines

Faith Salie (moderator), Hod Lipson, David Ferrucci, Eric Horvitz, and Rodney Brooks

A panel discussion. World Science Festival, New York. 4 June 2011.

One research program aiming to assemble more convincing humanoid agents is "integrative AI," which combines machine vision, machine learning, and language processing to make a system that works closely with a human actor. Think of these agents as descendants of Clippy, the old Microsoft Office virtual assistant, now capable of reading your e-mails, listening to your phone calls, or greeting you at the airport. For Microsoft researcher Eric Horvitz, integrative AI is already an everyday presence.

One response. Ken Jennings's postscript to his final Jeopardy! answer.

A female avatar sits outside his office on a computer screen, handling questions from visitors, debriefing him about appointments, and recording his behavior and priorities. "She can even predict which meetings I won't go to," he reported cheerily. There are grander applications, too: a similar system has been used as a triage assistant in medical care, capable of interviewing patients about their symptoms, drafting diagnoses, and scheduling appointments. Motivating his efforts to build a stronger AI is the belief that the human mind can ultimately be computed. "There's something astounding and magical about that," Horvitz remarked.

While Horvitz creates machines that enable human beings, Lipson has been building machines that think for themselves. By analyzing the behavior of a two-link pendulum, one machine managed to "discover" Newton's law $F = ma$. Another biologically inspired robot exhibited the rudiments of self-awareness. This starfish-shaped crawler learned a three-dimensional representation of its body the same way the human brain might—by experimenting with different forms of self-motion and inferring what body structure would generate the data culled from its sensors. Exploiting this mental model, the robot managed to learn a walking gait and even adapted to a hobble after one of its legs was removed. "It was pretty lame," Lipson lamented. "We were hoping it would develop an evil, spidery walk."

Other researchers, meanwhile, are instead trying to mitigate the creepiness of robotic behavior. Rodney Brooks, the Massachusetts Institute of Technology professor emeritus who founded iRobot, warned of the so-called "uncanny valley," that psychic discomfort zone induced by "robots that look a little too human but don't deliver on the promise." Brooks, who recently launched a startup focused on manufacturing, demonstrated a panoply of social robots that can track eye movements, make faces, or (in the case of a baby doll) burp. Endowing robots with the simple capacity to follow and make eye movements goes a long way toward creating a sympathetic humanoid, provided humans can fill in the blanks. "I think we give each other souls," Brooks said, gesturing at Horvitz. "He's a bag of skin, but I interact with him and anthropomorphize him."

The struggle to comprehend man-made minds indicates not only how wide the gulf can be between different branches of AI research but also how our public conversation about artificial intelligence tends to remain fairly artificial itself. If Star Trek's computer serves as one bookend in the popular imagination (and, Ferrucci says, an inspiration for his pursuits at IBM), then HAL of Stanley Kubrick's *2001* is the other—portending a dystopian future where machines work not with but against us. Notwithstanding Watson's hilarious mistakes (which delighted the humans in the audience), the Jeopardy! winner suggests both futures: the computer that humans have built to do our "thinking" for us and the kind that can instill dread in us, too.

But while Watson's success in language processing and guesswork is certainly impressive, the computer itself doesn't signal progress in reverse-engineering the brain or imply that reconstructing human intelligence is even necessary for AI. On the contrary, Ferrucci suggested, Watson's greatest application may be as a physician's assistant (able to navigate the growing body of medical knowledge) or as a nimble processor of the information that humans now produce at accelerating rates. "We are overwhelmed by data," added Lipson. "Computers can help."

Ferrucci noted that IBM designed Watson to be faceless, so humans would judge it on the merits of its software. The discussion, however, reminded us that by pitting machines against humans in high-publicity events or showcasing them in glossy stage demos, we risk underselling both our computers and our minds. And by drawing our concerns about AI from the pages of science fiction, we overlook other, more plausible threats, such as AI's effects on labor and financial markets and on digital security. Many technologies, from the atomic bomb to particle accelerators to synthetic biology, have raised all kinds of public specters. The progress of AI requires public understanding as much as scientific research. Like Watson's "thinking" process, our own relies on learning from failures. Advancements in the field will depend not just on making machines better able to generate answers but on contemplating the many deep questions they ask of us.

10.1126/science.1209941