

## CORRESPONDENCE

---

### A Note to Our Readers

With this issue, we launch two new sections in the journal: a “Reviews and Reconsiderations” section that will consider books and authors, both old and new; and a “Correspondence” section to provide space for friends and critics to continue the conversations that start in our pages. We urge readers to send letters via e-mail to [letters@thenewatlantis.com](mailto:letters@thenewatlantis.com) or by post to our editorial offices: 1015 Fifteenth St. N.W., Suite 900, Washington, D.C. 20005.

ERIC COHEN  
Editor

### Visions of the Future

Thank you for your kind review of my book *Radical Evolution* [“The Rhetoric of Extinction,” Winter 2006].

One point in the piece bears further reflection. If I am indeed the model of probity you say I am (thanks), and if you accept that I respect and value the arguments of people who value human nature as it exists, you may want to publish a vision of the future you desire that looks positive and achievable to open-minded people like me. As I say in *Radical Evolution*:

Right now, the argument is usually cast rather fruitlessly between the proponents of the Hell and Heaven Scenarios. One side sees the dangers and wants everything stopped. The other side sees the promise and serves as cheerleaders. They talk past each other.

I do not see myself as a rhetorician on either side of this sterile debate. I am a reporter. In *Radical Evolution* I saw my

task as investigating where technology is taking human nature. I did my best to follow the facts where they took me. Indeed, many thoughtful readers find my report on the “Heaven” scenario, which lays out the logic for the “extinction” of “Version 1.0 humans” in favor of supposedly higher forms of humanity, as even scarier than the “Hell” scenario, which describes the possible technological destruction of all life on earth.

While I can see how either of these scenarios might in fact occur, I can also see how both of them are simplistic and techno-deterministic. Indeed, they are mirror images, binary, black and white. This is why I continued my reporting, to see if there was a third scenario that a pragmatist might view as optimistic. This is how I uncovered “Prevail”—the scenario in which human nature indeed continues its ornery and herky-jerky rise to glory, without becoming the tool of its tools.

By lumping me in the same camp as the forthright advocates of the Heaven scenario he abhors, I fear Mr. Rubin does himself and your journal few favors. Just because I’m not persuaded that the Hell prediction is revealed truth doesn’t make me a “rhetorician of extinction.” It does, however, mean that you’ve got some work to do to give any fair-minded reader a reason to embrace your position.

As I wrote:

Even assuming that in our role as architects of the future we pick and choose the elements that we implement, I have yet to encounter a persuasive argument that the advantages conveyed by the GRIN technologies [genetics, robot-

ics, info-tech, and nanotech] are likely to be stopped worldwide—short of a cataclysm. In fact, the more problems we face, the more rapidly we probably will reach for an ingenious and seemingly miraculous fix. Nor have I seen a case made that convinces me I'd like to live in a world in which human imagination were so entirely blocked.

I do not wish to be cast as an opponent or a debunker of the social critics of technology. I hope I have presented them and their scenarios fairly. Readers should examine their arguments carefully. They offer important reasoning regarding the cautions we should consider....In the absence of an attractive alternative, however, I elect to light out for the Territory. I choose to examine the possibility that human nature might continue to evolve and be improvable, and to consider what transformation might actually look like and what it might mean.

It is that lack of an attractive alternative that I speak to here. After all my work, I did not encounter the case that “convinces me I'd like to live in a world in which human imagination were so entirely blocked.”

So here is my challenge: Craft that case in a positive fashion, one that might resonate with a careful and fair-minded pragmatist like me who did not enter this topic with any preconceived agenda. I'm not sure that simply scaring the hell out of people by yelling “extinction” does the job.

JOEL GARREAU  
Broad Run, Va.

CHARLES T. RUBIN responds: I'm glad to be among the “thoughtful readers” of Joel Garreau's book who have pointed out that the “Heaven” scenario he describes could be considered hellish for its promise

of replacing human beings with “supposedly higher forms of humanity.” That highlighting this possibility he considers “yelling” suggests to me the essential sympathy with transhumanist goals that I observed in his book; if the emperor is naked then any voice so saying is presumably too loud.

Because Mr. Garreau cannot imagine that anything short of catastrophe could slow the accelerating pace of technological development, all of his scenarios—even “Prevail”—involve fundamental alterations in the shape of human life. Anything less than a “rise to glory” he regards as blocking human imagination. For this variety of open-mindedness, which implicitly seems to equate imagining with doing, the only way to meet one grand project for the future is with another grand project for the future.

A thoughtful understanding of technological change requires something that is at the same time much easier and much harder than creating a grand vision: an imaginative sympathy with what William Dean Howells called, with some irony, “poor Real Life.” It is easier because even with vast technological change, we all still live Real Life; “the fundamental things apply as time goes by.” People love, form families, face triumph and tragedy, fight and get along and get by. One could aspire to a world where our descendants will still be able to make most of our mistakes, even as conditions which compromise the chance for a fully human life are ameliorated. Such a sympathy is harder than grand visions, however, because what is closest to us is often most difficult to see. And we are likely to be further distracted when “pragmatism” or “open-mindedness” are defined, as Garreau defines them, in terms of techno-evolutionary utopianism

and dystopianism. If Mr. Garreau lights out for the Territory, he will likely find it just another way of failing to come to grips with what we are, and failing to see that being “positive” about human life hardly requires believing in man’s progressive “rise to glory.” There are other genuinely human alternatives, and one need look no further than Yuval Levin’s fine article in these pages, aptly called “Imagining the Future” [Winter 2004], to see and understand them.

### The Turing Test

As the sponsor of the Loebner Prize, I wish to clarify some points made by Mark Halpern in his article “The Trouble with the Turing Test” [Winter 2006].

It is poor scholarship to rely on obsolete and inappropriate data in a scholarly article, which is exactly what Mr. Halpern has done. He states that “only the first [Loebner Prize contest], held in 1991, was well documented and widely reported on in the press, making that inaugural event our best case study.” I am astonished that Mr. Halpern would consider the amount of press coverage as the measure of what constitutes a “best case study.” The first contest *may* have been the most widely reported, but that certainly does not make it the best case study.

In fact, the first contest is among the *worst* case studies of the Loebner Prize as a Turing Test, because it was rife with errors. The greatest objection to the 1991 contest is that it required the participants to restrict their conversations to a single topic. This is not at all what Turing described, and we eliminated that restriction starting in 1995. (I did, however, introduce a minor restriction in 2004 and 2005 that each entity had to begin each conversation with the sentence, “Hello,

my name is [X] and I am the human.” Although Turing did not describe this, I introduced the requirement to focus the attention of the participants on the essential task of the enterprise.)

Another objection to the first contest—and to some of the other early contests that followed it—is that it did not conform to the structure of the test that Alan Turing described. His “Imitation Game” can be characterized as a “paired comparison” in which a judge must decide which of *exactly two* entities is the human and which is the computer. But in the 1991 contest, and every contest until 2004, there were several humans and several computers which had to be compared with each other simultaneously. It wasn’t until 2004 that this error was corrected and an *actual* Turing Test was conducted. In both the 2004 and 2005 contests each judge had to compare one of four pairs of entities, each comprising a human and a computer program. In effect, the 2004 and 2005 competitions were four separate Turing Tests held simultaneously.

This is not to say that the 2004 and 2005 competitions were without faults. There were at least two serious faults in the 2004 contest, one of which was corrected in the 2005 contest, and the other of which will be corrected this year, in the 2006 competition.

The first fault in 2004 related to the scoring. Each judge was required to divide 100 points between the two entities being compared based upon “how human” each of the entities appeared. The winner was the computer with the highest sum of the scores of the four judges. Some judges assigned scores with a high variance (e.g., 95-05) and others with a low variance (e.g., 45-55). Each judge’s relative influence in determining the winner was therefore directly proportional to the variance of his scores. This problem was corrected

in the 2005 contest, when raw scores were converted to rankings, and the computer entity with the highest mean rank was declared the winner.

The second fault was related to the way the humans and judges communicated. In order to conduct a Turing Test it is necessary that there be some means for the human confederate to communicate with the judge. In both 2004 and 2005, the creators of each competing computer program had to provide a separate communications program so that the judge-human conversation would have the same on-screen appearance as the judge-computer conversation. Unfortunately, one competitor submitted a communications program that slightly garbled the judge-human conversation. This difficulty was not discovered until the competition was held. Even though that rule didn't change in 2005—I did not want to impose a major change without a two-year lead-time to accommodate it—there were no apparent communications errors last year. In 2006, I will provide the communications programs.

I have one other nit to pick with Mr. Halpern. He states: “Turing’s paper claimed that suitably programmed digital computers would be generally accepted as thinking by around the year 2000.” That is not what Turing actually wrote. Here are Turing’s words:

I believe that in about fifty years’ time it will be possible, to program computers, with a storage capacity of about  $10^9$  [binary digits], to make them play the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning.

This is just a benchmark that Alan Turing set, and not, I think, his criterion

for passing the test. Turing might just have well used “in ten years’ time,” “5 percent chance,” and “after 30 seconds.” In fact, as University of Canterbury professor Jack Copeland has pointed out, Turing was not wedded to the year 2000; he made other predictions about how long it would take to pass his test, including a 1952 prediction on a BBC broadcast that it would take “at least 100 years.”

HUGH LOEBNER  
New York City

MARK HALPERN responds: The letter in which Hugh Loebner complains of “poor scholarship” needs some correcting itself.

The first count in Loebner’s indictment is that in choosing the first Loebner Prize contest, that of 1991, as the example to analyze at length, I chose the worst possible case. But it is the worst only in the sense that it was the least successful of all Loebner’s attempts at realizing the test as Turing described it—and that is not a consideration that played a significant part in my criticism of the contests. Of course, I noted in passing that the conditions of the first contest were hardly those that Turing specified, but the fact that later contests approximated those conditions better and better as their designers learned from their mistakes and failures does not affect my criticisms in any way, since I regard the Turing Test itself—Loebner’s model and the ideal which he attempted to realize—as invalid and unperformable. In a sense, therefore, the Loebner Prize contests got worse, if anything, as they more closely approximated what their designers took to be Turing’s intentions.

Loebner also expresses astonishment at my reason for concentrating on the contest of 1991—namely, that it was the best documented by the press and other observers. But it was not, as he implies, the sheer

quantity of words written about it that made me regard it as the best documented, it was the quality, the variety, and the authority of the various reports and studies that were made of that event—and not of the later ones. Just because it was the first contest, it was written about at length, and from a wide variety of viewpoints: Besides substantial American print coverage of the 1991 event, stories appeared in the presses of Britain, the USSR, Italy, Brazil, China, Japan, Portugal, and the Arabic-speaking Middle East—all this in addition to extensive television coverage.

And that first contest was written about not just in the popular media, but in technical journals, where it was reviewed by AI workers, computer scientists, and philosophers. Later contests attracted nothing like this degree of coverage; for many of them, the only documentation is the official transcript of the exchanges between terminals and judges, which appears only on the contest's website. I have read extensively in all the transcripts I've been able to find of the later contests—I think I am missing the records of only one year since the inception of the contests—and I see nothing in them to cause me to retract or significantly modify any of the conclusions that my study of the 1991 contest led me to.

Finally, Loebner says that my statement "Turing's paper claimed that suitably programmed digital computers would be generally accepted as thinking by around the year 2000" misrepresents what Turing

actually wrote, and he quotes a passage of Turing's that indeed does not support my words—but it is not the passage to which I was alluding. The one I had in mind was this: "...I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted." This is the next sentence but one after the passage that Loebner quotes and claims I misrepresented. But of course the value of the Turing Test is not to be measured by the accuracy of Turing's offhand predictions and conjectures; while it is worth remarking as an aside that this prediction of his has failed, that is not why I or any of the other serious critics of his views on computer intelligence believe that those views are fundamentally and incorrigibly flawed.

Perhaps I devoted an inordinate number of pages to the Loebner Prize contest; it offers such vivid illustrations of human gullibility and self-deception that one gets caught up in it and follows it a little beyond what its value for an investigation of AI strictly warrants. But the fundamental divide between Loebner and me is that he sees the problem with his eponymous contest as its failure to fully realize Turing's thought-experiment, while I see it as getting worse as it more closely approximates that protocol; the main thesis of my article is the faults of Turing's Imitation Game, not those of Loebner's imitation of it.