Book Review

This review is forthcoming in *Minds and Machines.*

Steven Pinker, *How the Mind Works*, New York: W. W. Norton, 1997, xii + 660 pp.; $29.95 (cloth), ISBN 0-393-04535-8; $17.00 (paper), ISBN 0-393-31848-6.[1]

1

Steven Pinker should be well-known to readers of this journal, not only for his research in linguistics and cognitive science, but for his semi-popular book, *The Language Instinct* (1994; for the *M&M* review, see Kemmerer 1995). For those of us who liked *The Language Instinct* (as I did), Pinker now offers us the rest of the story: *How the Mind Works*, despite its inclusive-sounding title, literally complements its predecessor (cf. p. x), since it omits all discussion of language; thus, readers hoping for a user manual for the mind (and instructors hoping for a single-volume cognitive-science text) may be disappointed. One might have hoped for at least a chapter-length précis of the earlier book, though that would no doubt have made this 660-page tome even heftier. As it is, the book covers a wide range of topics: vision; computationalism; evolutionary psychology and natural selection; consciousness and qualia; intelligence; concepts and categories; the problem of other minds; reasoning; emotions; the psychology of social relations; and the psychology of the arts, humor, religion, and philosophy.

The book is aimed at "anyone who is curious about how the mind works", including "professors and students ... scholars and general readers" (p. x). Researchers in cognitive science will find its "bird's-eye view of the mind" useful and its evolutionary-psychology slant interesting, even if most of the topics covered go over familiar ground; lay readers should find it enlightening (if incomplete – see above) but perhaps a bit snide (especially on the topic of religion – see below). The writing is appealing, with many humorous asides – some to the point, others just throwaway lines that are probably funnier if heard in an undergraduate lecture than when read on paper.

2

Two themes run through Pinker's story: computationalism and evolutionary psychology:

The mind is a system of organs of computation, designed by natural selection to solve the kinds of problems our ancestors faced in their foraging way of life, in particular, understanding and outmaneuvering objects, animals, plants, and other people. The summary can be unpacked into several claims. The mind is what the brain does; specifically, the brain processes information, and thinking is a kind of computation. The mind is organized into modules or mental organs, each with a specialized design that makes it an expert in one arena of interaction with the world. The modules' basic logic is specified by our genetic program. Their operation was shaped by natural selection to solve the problems of the hunting and gathering life led by our ancestors in most of our evolutionary history. The various problems for our ancestors were subtasks of one big problem for their genes, maximizing the number of copies that made it into the next generation. (p. 21; cf. pp. x, 429.)

2.1

Pinker advocates a strong version of the computational theory of mind: What "the brain does, which makes us see, think, feel, choose, and act ... *is* information processing, or computation" (p. 24, my emphasis). Thus, mind (= what the brain does) is a species of information processing, which is equated to computation. This solves part of the mind-body problem, viz., "how to connect the ethereal world of meaning and intention, the stuff of our mental lives, with a physical hunk of matter like the brain" (p. 24). The solution is that

... beliefs and desires are *information*, incarnated as configurations of symbols ... [which] are the physical states of bits of matter, like chips in a computer or neurons in the brain. They symbolize things in the world because they are triggered by those things via our sense organs and because of what they do once they are triggered. If the bits of matter that constitute a symbol are arranged to bump into the bits of matter constituting another symbol in just the right way, the symbols corresponding to one belief can give rise to new symbols corresponding to another belief logically related to it, which can give rise to symbols corresponding to other beliefs, and so on. Eventually the bits of matter constituting a symbol bump into bits of matter connected to the muscles, and behavior happens. The computational theory of mind thus allows us to keep beliefs and desires in our explanations of behavior while planting them squarely in the physical universe. It allows meaning to cause and be caused. (p. 25.)

Although Pinker's computationalism is a strong one, he distinguishes the computational theory of mind from the "computer metaphor"; i.e., brains aren't PCs. Rather, "brains and computers embody intelligence for some of the same reasons" (pp. 26–27); e.g., birds and airplanes both fly, for the same physical reasons, but birds don't have jet engines or "complimentary beverage service" (p. 27). Moreover, the computational theory of mind is essential for "mak[ing] sense of the evolution of the mind. ... [E]volution equipped us ... with a neural computer"

(p. 27; cf. pp. 36ff).

But the computational theory of mind disappears after the careful and generally good discussion of it in the first two chapters, and little is said about the computational implications of the evolutionary analysis. This may, however, merely represent the state of the art; if Pinker (and his primary sources) are right about what the mind is, much work remains to be done on the details of how it works (despite the book's title).

2.2

The real heart of the book is evolutionary psychology. Pinker does an admirable job clarifying evolutionary theory, though occasionally he slips into teleological or anthropomorphic locutions (as he also did at the end of the "key idea" quotation from p. 21: How can genes have "problems"?). For instance, instead of saying:

> The ultimate goal that the mind *was designed* to attain is maximizing the number of copies of the genes that created it. (p. 43, my emphasis.)

he should have said that the mind has evolved in such a way that the number of genes that created it is maximized. And instead of saying:

> Natural selection *cares* only about the long-term *fate* of entities that replicate, that is, entities that retain a stable identity across many generations of copying. (p. 43, my emphasis.)

it would have been more accurate to say merely that it is replicators that natural selection applies to. Or instead of this:

> When we ask questions like "Who or what is *supposed to* benefit from an adaptation?" and "What is a design in living things a design *for*? the theory of natural selection provides the answer: the long-term stable replicators, the genes. (p. 43; my emphasis.)

it would have been better simply to say that genes benefit (not: are "supposed to" benefit) from an adaptation, and that genes give rise to and are the beneficiaries of (apparent?) design in living things.

Or consider this passage, which is an interesting mixture of what I would consider correct terminology, as well as bad teleological/anthropomorphic terminology:

> But almost everyone misunderstands the theory. Contrary to popular belief, the gene-centered theory of evolution does *not* imply that the point of all human striving is to spread our genes. ... Our goals are subgoals of **the ultimate goal of the genes**, replicating themselves. (pp. 43–44; my boldface; italics in original.)

It is misleading, especially in a book one of whose purposes is to clarify evolution, to talk about goals; I would have said simply that genes replicate themselves (not that that is their "goal"). And instead of this:

> But the two are different. As far as *we* are concerned, our goals, conscious or unconscious, are not about genes at all, but about health and lovers and

> children and friends. ... Sexual desire is *not* people's strategy to propagate their genes. It's people's strategy to attain the pleasures of sex, and **the pleasures of sex are the genes' strategy to propagate themselves**. (pp. 43–44; my boldface; italics in original.)

I would have said that genes have evolved in such a way that we humans experience the pleasure of sex and, as a result, genes propagate themselves.

3

In addition to the twin themes of computation and evolution, I detected an implicit third theme: The "mind" that Pinker endeavors to explain is not mind in general or in the abstract, but the human mind in particular. Some cognitive scientists (I count myself among them – see Rapaport 1998) are more interested in the former – in how cognition, mentality, thinking (call it what you will) works, independently of how it is implemented. Granted, other cognitive scientists (e.g., Thagard 1986) have been at pains to point out that the implementing medium may contribute a lot to the mind's functioning. And some of these cognitive scientists have suggested that (probably) only minds implemented in animal brains are worthy of the name (Searle 1980 can be read this way; for further discussion, see Rapaport, forthcoming).

Although Pinker gives lip service to the more abstract view:

> ... the mind is not the brain but what the brain does, and not even everything it does, such as metabolizing fat and giving off heat. ... Information and computation reside in patterns of data and in relations of logic that are independent of the physical medium that carries them." (p. 24; also see the quotation from p. 65, below.)

he is concerned primarily with the *human* mind, as evidenced by passages such as one on postpartum depression: "... as with all explanations of complex emotions, one must ask *why* the brain is wired so as to let hormones have their effects" (p. 444). It seems unlikely that robots will suffer from postpartum depression. Thus, there remains a certain tension between such an implementation-dependent view of (the human) mind and the computational view, for if mind *is* computational, then it can be implemented in other media besides the (human) brain, in which case, the effect of hormones is peculiar to the human mind. (Unless, of course, it turns out that mind in the abstract requires some (computational?) *analogue* of hormones in order to function. Maybe robots *would* suffer from postpartum depression?)

4

As the context of the previous quotation about hormones may have suggested, there are some topics in the book that are of questionable relevance: Postpartum

depression, parent-offspring conflict, the nature of sex differences, the human mating system, fashion, and war are not likely to be high on any cognitive scientist's list of *mental* functions. Besides the implicit theme of *human*-brain dependency that I noted, a clue to the relevance of these topics may be found in a passage on war:

> War is, to put it mildly, a major selection pressure, and since it appears to have been a recurring event in our evolutionary history, it must have shaped parts of the human psyche. (p. 510.)

So, the theme of evolutionary psychology leads Pinker to be very inclusive about the (human) mind's functions. But lots of things must have been major selectional pressures – e.g., climate and natural disasters – yet these would clearly be irrelevant topics to include in a book called "How the Mind Works". I hasten to add that the very inclusion in this book of topics such as postpartum depression, war, etc., may well cause them to *become* mainstream cognitive science, and certainly represents the fact that some cognitive scientists already consider them thus.

5

One disappointment is the overly cute, but opaque, chapter titles. Someone skimming the table of contents would be hard pressed to know what "Revenge of the Nerds" or "Hotheads" were about. So let me tell you.

Chapter 1, "Standard Equipment", describes the design of a robot with vision, motion control, reasoning and common sense, and motives. Evolutionary psychology (relying heavily on the work of John Tooby and Leda Cosmides)[2] and computational cognitive science play the central roles in this chapter, though Pinker sees the latter as a component of the former. Pinker spends the rest of the chapter clarifying evolutionary theory, and concludes with a discussion of some moral implications of evolutionary psychology.

Chapter 2, "Thinking Machines", defines 'intelligence' as "the ability to attain goals in the face of obstacles by means of decisions based on rational (truth-preserving) rules" (p. 62; Pinker's discussion is based on Dennett 1971, Newell and Simon 1972, and Pollock 1993[3]). And, following Dretske 1981, 'information' is defined as "a correlation between two things that is produced by a lawful process (as opposed to coming about by sheer chance). ... Correlation is a mathematical and logical concept; it is not defined in terms of the stuff that the correlated entities are made of" (p. 65). Pinker expands on the computational theory of mind, giving an extended and good description of one computational model of thinking (production systems), and he discusses the McCulloch-Pitts (1943) theory of mental computation. He takes an internalist, methodologically solipsistic stance (without mentioning Fodor 1980): "The information in an internal representation is all that we can know about the world" (p. 84), and he presents and tries to rebut the Chinese-Room Argument (Searle 1980) and (more briefly) the arguments of Penrose 1989, 1994.

He discusses connectionism, correctly noting that "Connectionism is not an alternative to the computational theory of mind, but a variety of it" (p. 114). As might be expected from those familiar with Pinker and Prince 1988, he takes a moderately, but detailed, anti-connectionist stand: Without modification, pure connectionism cannot handle: (1) "the concept of an individual" (p. 114), (2) compositionality, (3) variable binding and quantification, (4) recursive thoughts, or (5) rule systems. The chapter concludes with a discussion of consciousness; eventually, Pinker goes along with Colin McGinn's (1993) view that we probably are incapable of understanding it.[4]

Chapter 3, unhelpfully titled "Revenge of the Nerds", elaborates on the theory of natural selection and the origin and evolution of life, intelligence, and the brain. It contains crucial passages on a proper understanding of natural selection:

> The eye ... appears to have been designed ... with the *goal* of putting together something that sees. ... What else but the plan of God could effect the teleology ... of life on earth?
>
> Darwin showed what else. He identified a forward-causation physical process that mimics the paradoxical appearance of backward causation or teleology. The trick is *replication*. ... [State of affairs] B can't cause A if A comes first. ... But let's say that A causes B, and B in turn causes the protagonist of A to make a copy of itself[:] ... AA. AA looks just like A, so it appears as if B had caused A. But it hasn't ; it has only caused AA, the *copy* of A. Suppose there are three animals, two with a cloudy lens, one with a clear lens. Having a clear lens (A) causes an eye to see well (B); seeing well causes the animal to reproduce by helping it avoid predators and find mates. The offspring (AA) have clear lenses and can see well, too. It looks as if the offspring have eyes *so that* they can see well (bad, teleological, backward causation), but that's an illusion. The offspring have eyes because *their parents'* eyes *did* see well (good, ordinary, forward causation). Their eyes *look like* their parents' eyes, so it's easy to mistake what happened for backward causation. (pp. 156–157.)

Chapter 4, "The Mind's Eye", concerns the evolution and nature of vision, and contains a fascinating and detailed explanation of "magic eye" autostereograms and Bela Julesz's (1971) theory of cyclopean perception, as well as a discussion of mental imagery (although he spends only minimal time on the pictorial-vs.-propositional debate).

Chapter 5, "Good Ideas", concerns the evolution of reasoning and inference, categorization, and the concept of an object. As for categories, Pinker takes an anti-Lakoff (1987; see Lammens 1994) stand:

> Many anthropologists and philosophers believe that categories are arbitrary conventions that we learn along with the other cultural accidents standardized in our language. ... But categories would be useful only if they meshed with the way the world works. ... Mental boxes work because things come in clusters that fit the boxes. ... Our theories, both folk and scientific, can idealize

away from the messiness of the world and lay bare its underlying causal forces. (pp. 308, 312.)

He does, however, acknowledge the existence of fuzzy categories (e.g., *fish*). He also discusses the so-called "theory of mind" theory, but he does not mention the "theory-theory"-vs.-simulation controversy.[5] He concludes with a discussion of metaphor, concentrating on the two fundamental metaphors of force and spatial location.

Chapter 6, with another cute but opaque title ("Hotheads"), is about emotions. 'Intelligence' is re-defined (more weakly) as the "pursuit of [rather than "the ability to attain" – cf. p. 62)] goals in the face of obstacles" (p. 372), where the human goals are (1) "the Four Fs" (feeding, fighting, fleeing, and reproduction),[6] (2) "understanding the environment", and (3) "securing the cooperation of others" (p. 373). Later (p. 541), he adds survival, which, along with reproduction, is more of an overarching goal: presumably, one feeds, flees, fights, understands the environment, and cooperates in order to survive. Intelligence requires emotions because not all goals can be pursued simultaneously: "emotions ... set the brain's highest-level goals" (p. 373). Among the "emotions" Pinker discusses are disgust, fear, happiness, love and altruism, self-control, and self-deception. As I noted earlier, despite his attempts to be a good evolutionary psychologist, he does occasionally get caught in odd (but natural) linguistic traps: In a discussion of altruism, he says, "The dispersed copies of a gene *call to one another* by endowing bodies with emotions" (p. 401, my emphasis).

Chapter 7, "Family Values", is probably the least satisfying chapter for those interested in mind in the abstract (but probably the *most* interesting lectures when Pinker teaches them!). The main topic is the psychology of social relations: romantic love; the nature of families and kinship; parent-offspring conflict; the incest taboo; gender stereotypes, sex and its evolution, and the human mating system; the sense of beauty; the evolution of fashion; and war.

The final chapter, "The Meaning of Life", concerns the psychology of the arts, humor, religion, and philosophy, understood "within the theme of this book, that the mind is a naturally selected neural computer" (p. 521). Yet there is no discussion of computation in these pages, even though there easily could have been (see, e.g., Minsky 1984 on humor, or Duchan et al. 1995[7] and Ram and Moorman 1999 on narrative). The section on religion ("The Inquisitive in Pursuit of the Inconceivable", pp. 554–565) is a bit strong for the general public: I can easily imagine many readers bristling at passages such as this: "How does religion fit into a mind that one might have thought was designed [shouldn't Pinker have said "had evolved"?] to reject *the palpably not true*?" (p. 554, my emphasis). To say the least, that religion might be "palpably not true" requires an argument.[8]

6

There are useful notes and references at the end of the book, but the index leaves something to be desired, e.g., the entry for "Qualia" refers the reader to entries for "Consciousness, Sentience", but there is no entry for "Sentience", and there is an entry for "Mentalese" but not for "Language of Thought", even though that's mentioned on pp. 70 and 90. My professional proofreader's eye only caught one typo ('azslo' for 'also', p. 361) in addition to the Pollock/Pollard confusion that I mention in endnote 3 of this review.

In sum, and despite my caveats, there's a lot here, and it makes for fascinating reading.[9]

NOTES

[1]There is also an abridged, audio-cassette-tape version, read by one Grover Gardner (Audio Scholar, 1998; $24.95, ISBN 1879557509). As I write this, the paperback version may be had for $14.72 (including U.S. shipping) from http://www.kingbooks.com; see http://www.acses.com for the latest prices. Interestingly, there is a 1934 book with the same title, edited by Cyril L. Burt (New York: D. Appleton-Century), with chapters titled: "How the Mind Works in the Adult: The Conscious Mind", by Burt; "The Unconscious Mind", by Ernest Jones; "How the Mind Works in the Child: Problems in the Development of the Child", by Emanuel Miller; "Problems in the Treatment of the Child", by William Moodie; and "How the Mind Works in Society", by Burt – no doubt a very different book from Pinker's!

[2]See especially Tooby and Cosmides 1992. For an independent discussion of evolutionary psychology and the debates surrounding it, see Mitchell, forthcoming.

[3]The latter consistently misattributed as "Pollard" in the text, notes, references, and index!

[4]For useful semi-popular expositions and critiques that can supplement Pinker's presentation of this view, see McGinn 1999, Strawson 1999.

[5]For useful overviews of some of these issues, see Duchan 1995 and forthcoming.

[6]I first read this joke in an essay by Patricia Churchland, but I don't know who originated it; Pinker doesn't explain it.

[7]See Graesser and Bowers 1996 for the *M&M* review.

[8]And, in any case, Stanisław Lem's 1971 science-fiction story "Non Serviam" – which should be required reading for all researchers in artificial life (ALife) – suggests how religion can be provided an evolutionary *and* computational explanation; briefly, in the story, ALife "personoids" evolve religious beliefs about their creator (who is, in fact, their programmer). Also, Thalos 1998 argues that belief in God serves an evolutionary purpose in stimulating cooperation.

[9]I am grateful to my colleague Mariam Thalos for comments on an earlier version of this review.

References

Astington, Janet W.; Harris, Paul L.; and Olson, David R., eds. (1988), *Developing Theories of Mind*, Cambridge, U.K.: Cambridge University Press.

Dennett, Daniel C. (1971), 'Intentional Systems', *Journal of Philosophy* 68, pp. 87–106; reprinted in Daniel C. Dennett, *Brainstorms* (Montgomery, VT: Bradford Books), pp. 3–22.

Dretske, Fred I. (1981), *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.

Duchan, Judith Felson (1995), Review of Whiten 1991, *Minds and Machines* 5, pp. 115–120.

Duchan, Judith Felson (forthcoming), Review of Astington et al. 1988, Wellman 1990, and Frye and Moore 1991, *Minds and Machines*.

Duchan, Judith F.; Bruder, Gail A.; and Hewitt, Lynne E. (eds.) (1995), *Deixis in Narrative: A Cognitive Science Perspective*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Fodor, Jerry A. (1980), 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology', *Behavioral and Brain Sciences* 3, pp. 63–109.

Frye, Douglas, and Moore, Chris, eds. (1991), *Children's Theories of Mind: Mental States and Social Understanding*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Graesser, Arthur C., and Bowers, Cheryl A. (1996), Review of Duchan et al. 1995, *Minds and Machines* 6, pp. 395–399.

Julesz, Bela (1971), *Foundations of Cyclopean Perception*, Chicago: University of Chicago Press.

Kemmerer, David (1995), Review of Pinker 1994, *Minds and Machines* 5, pp. 411–417.

Lakoff, George (1987), *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, Chicago: University of Chicago Press.

Lammens, Johan M. (1994), Review of Lakoff 1987, *Minds and Machines* 4, pp. 115–122.

Lem, Stanisław (1971), 'Non Serviam', in Stanisław Lem, *A Perfect Vacuum*, trans. by M. Kandel (New York: Harcourt Brace Jovanovich, 1979).

McCulloch, Warren S., and Pitts, Walter H. (1943), 'A Logical Calculus of the Ideas Immanent in Nervous Activity', *Bulletin of Mathematical Biophysics* 7, pp. 115–133; reprinted in Warren S. McCulloch, *Embodiments of Mind*, Cambridge, MA: MIT Press, 1965), pp. 19–39.

McGinn, Colin (1993), *Problems in Philosophy: The Limits of Inquiry*, Cambridge, MA: Blackwell.

McGinn, Colin (1999), 'Can We Ever Understand Consciousness?', *New York Review of Books* (10 June 1999), pp. 44–48; available online at
http://nybooks.com/nyrev/WWWarchdisplay.cgi?19990610044R.

Minsky, Marvin (1984), 'Jokes and Their Relation to the Cognitive Unconscious', in Lucia Vaina and Jaakko Hintikka, eds., *Cognitive Constraints on Communication: Representations and Processes*, Dordrecht: D. Reidel.

Mitchell, Melanie (forthcoming), 'Can Evolution Explain How the Mind Works? A Review of the Evolutionary Psychology Debates', *Complexity*; preprint available at
http://www.santafe.edu/~mm/paper-abstracts.html#ep-essay.

Newell, Allen, and Simon, Herbert A. (1972), *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall.

Penrose, Roger (1989), *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford: Oxford University Press.

Penrose, Roger (1994), *Shadows of the Mind: A Search for the Missing Science of Consciousness*, New York: Oxford University Press.

Pinker, Steven (1994), *The Language Instinct: How the Mind Creates Language*, New York: William Morrow.

Pinker, Steven, and Prince, Alan (1988), 'On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition', *Cognition* 28, pp. 73–193; reprinted in Steven Pinker and Jacques Mehler, eds., *Connections and Symbols*, Cambridge, MA: MIT Press, pp. 73–193.

Pollock, John L. (1993), 'The Phylogeny of Rationality', *Cognitive Science* 17, pp. 563–588.

Ram, Ashwin, and Moorman, Kenneth, eds. (1999), *Understanding Language Understanding: Computational Models of Reading*, Cambridge, MA: MIT Press.

Rapaport, William J. (1998), 'How Minds Can Be Computational Systems', *Journal of Experimental and Theoretical Artificial Intelligence* 10, pp. 403–419.

Rapaport, William J. (forthcoming), 'Cognitive Science', in Anthony Ralston, Edwin D. Reilly, & David Hemmendinger (eds.), *Encyclopedia of Computer Science, 4th edition*, New York: Grove's Dictionaries.

Searle, John R. (1980), 'Minds, Brains, and Programs', *Behavioral and Brain Sciences* 3, pp. 417–457.

Strawson, Galen (1999, July 11), 'Little Gray Cells', *The New York Times Book Review*; available online at http://www.nytimes.com/books/99/07/11/reviews/990711.11strawst.html.

Thagard, Paul (1986), 'Parallel Computation and the Mind-Body Problem', *Cognitive Science* 10, pp. 301–318.

Thalos, Mariam (1998), 'The Economy of Belief or, Explaining Cooperation among the Prudent', *American Philosophical Quarterly* 35, pp. 349–363.

Tooby, John, and Cosmides, Leda (1992), 'Psychological Foundations of Culture', in Jerome H. Barkow, Leda Cosmides, and John Tooby, eds., *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, New York: Oxford University Press.

Wellman, Henry M. (1990), *The Child's Theory of Mind*, Cambridge, MA: MIT Press.

Whiten, Andrew, ed. (1991), *Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*, Oxford: Basil Blackwell.

Department of Computer Science & Engineering,   WILLIAM J. RAPAPORT
Department of Philosophy,
and Center for Cognitive Science
State University of New York at Buffalo
Buffalo, NY 14260-2000, U.S.A.
rapaport@cse.buffalo.edu
http://www.cse.buffalo.edu/~rapaport