

# From Real Patterns to Prospective Quacks:

## Review of *Brainchildren: Essays on Designing Minds* by Daniel Dennett

**Joao Teixeira**

Department of Philosophy,  
University of S. Carlos,  
13.565-905, Brazil

and

Center for Cognitive Studies,  
Tufts University,  
Medford, 02155  
USA

jteixe01@emerald.tufts.edu  
djte@power.ufscar.br

Copyright (c) Joao Teixeira 1998.

PSYCHE, 4(15), November, 1998.  
<http://psyche.cs.monash.edu.au/v4/psyche-4-15-teixeira.html>

**KEYWORDS:** Daniel Dennett, Artificial Intelligence, reverse engineering, real patterns, animal mind, expert systems, medical practice, ethical issues.

**REVIEW OF:** Daniel Dennett (1998) *Brainchildren: Essays on Designing Minds*. Cambridge, MA: MIT Press, ix + 418pp. ISBN 0-262-54090-8 Price: \$US20.00 pbk.

In his new collection Dan Dennett looks over a broad range of themes, from philosophy of mind to artificial life and from ethology to animal psychology. All the essays were written in the past twelve years and have already been published in conference volumes or in specialized journals. The purpose of the collection is, as the author says, to group such essays "for the convenience of students and other readers who do not have ready access to a major university library." The book is divided into four parts. The first is devoted to philosophy of mind, the second to artificial intelligence and artificial life and the third to issues related to ethology and animal mind. There is also a final part, which

has received the title "Standing Back". This final part consists of two essays, one providing an overview of Dennett's own work over the last decades and another which points to the discussion of moral issues raised by some specific uses of expert systems in medical and professional life. This last essay was published in 1986 and it is, in my view, the most absorbing one.

The section devoted to philosophy of mind includes ten essays, some of which have become classic pieces, such as, for instance, "Can Machines Think?" (with two postscripts), "The Practical Requirements for Making a Conscious Robot" and "The Unimagined Preposterousness of Zombies: Commentary on Moody, Flanagan and Polger". The first is a lengthy commentary of the significance and prospects of the Cog project developed by Rodney Brooks and his team at the MIT. Dennett presents several philosophical implications of designing a humanoid robot as the realization of the major thought experiment of this century. The last essay of this section, "The Unimagined Preposterousness of Zombies: Commentary on Moody, Flanagan and Polger", appeared in the *Journal of Consciousness Studies* in 1995. It contains an attack on the ontological possibility of creatures such as zombies. Zombies would be behaviorally indistinguishable from human beings and, still, lack "qualia" in so far as their inner lives would be "informationally sensitive" but not "experientially sensitive" like ours. The attack is purported to be an utmost rejection of a current realism concerning "qualia" as well as sharp criticism of the claim that an account of the nature of consciousness involves extra ingredients beyond functional organization.

It is interesting to read these two essays one after the other, since they suggest an implicit order of ideas: disavowing extra ingredients for consciousness besides functional organization and "qualia" seem to be the necessary steps to conceive of the possibility of designing a conscious robot. However, the issue concerning "qualia" is still being debated. Much has been written about the nature of "qualia" and their relationship to consciousness, zombies and "zimbos" in the past few years but, still, the debate concerning such epiphenomenal ghosts is far from being settled once and for all. Such a debate reveals of a lot of misunderstandings about what we mean when we talk of "qualia", "inverted qualia", "absent qualia" and so on. This is the motivation of Dennett's essay "Instead of Qualia" in which, however, he confesses that he has lost the battle to reach at least a clear and agreed-upon meaning for the word. So viewed, the philosophical debate is likely to be endless and, most likely, not to reach even the minimal semantic consent required to answer the ontological questions surrounding the issue posed by the existence of "qualia".

No doubt the most important and controversial essay reprinted in the section devoted to philosophy of mind is "Real Patterns" where a whole reassessment of Dennett's brand of functionalism is at stake. In this paper of 1991 Dennett attempts the difficult task of steering between the Scylla of instrumentalism and the Charybdis of realism. The essay has already received critical commentaries from John Haugeland (1993) and, more recently, from William Wilkerson (1997) but several issues can still be raised about Dennett's variety of realism concerning mental states and, more particularly, folk psychology.

Haugeland bases his overall criticism of Dennett's notion of pattern in the need to disambiguate between two kinds of issues. On one hand there is the sheer ontological commitment to the existence of the elements that compound a pattern. On the other hand, there is a further and more forceful commitment to the existence of the pattern itself, that is, to something *beyond* such constitutive elements. The first can be trivially ascertained, the second involves a certain brand of ontological holism. A vacillation in distinguishing those two levels of analysis and those two notions of pattern is, according to Haugeland, pervasive in Dennett's account and also responsible for several difficulties concerning the kind of reality we are to ascribe to intentional entities. But I shall not pursue Haugeland's criticism here. Rather, I shall briefly discuss Wilkerson's claim that an account of the reality of the patterns would ultimately have to be sought by establishing some mapping between folk psychology and brain structure.

Beliefs and desires are a paradigm case for such a discussion since they are the most common elements of folk psychology involved in the intentional stance. In accordance with Dennett's intentional stance, beliefs and desires describe and predict patterns of behavior, that is, they describe an objective reality. In other words, such patterns are, according to Dennett, real, if two criteria are taken into consideration. First, a pattern is real if a "compression algorithm" can describe such a pattern, that is, if there is a description taking advantage of some basic regularity. Second, a real pattern occurs if there is predictability, that is, if predictions are successful.

Wilkerson's contention runs as follows: on the one hand folk psychology can be considered instrumentalistically, since beliefs and desires are *posits* or *abstracta* in the same way as centers of gravity or other useful constructs involved in the explanation of a specific natural phenomenon--in this case, human behavior. On the other hand, it may be the case that beliefs are something in the head, since for Dennett minds are brains. So one could ask whether or not such theoretical entities would ultimately map to brain structures and, in this case, a tension between instrumentalism and eliminativism (and realism) would be inevitable. But there is more to Wilkerson's claim than questioning the tension between instrumentalism and eliminativism. Wilkerson seems to be pressing a more forceful concern, namely, the need to find "how the various states and entities of folk psychology are realized in the human brain" (p. 558). In order to make sense of the reality of the patterns, Dennett's approach would have to be supplemented with an account meant to bridge the "gap between a mechanistic, non-intentional description of the brain's activity and the intentional description of behavior that the brain must somehow make possible" (p. 562). In other words, Wilkerson attempts to provide an account of how patterns could possibly be generated by the brain, that is, how "the pattern in human behavior is reflective of a real pattern realized in the structural features of the brain" (p. 562). He rejects the hypothesis that patterns could be generated "by a program that manipulates various symbols (tokens of a language of thought) by virtue of the symbols' syntactic features" (p. 562). In so far as Dennett would also consider such a hypothesis implausible and way ahead of empirical data, he proposes a pattern generator based on parallel processing and randomness, namely, Dennett's Joycean Machine. By transforming multiple channel information into serial processing the Joycean Machine would account for the production of patterns and would also provide the possibility of

drawing a connection between two levels of description: the description of the brain's activity (Dennett's physical stance) and the intentional description (Dennett's intentional stance). Besides overcoming the tension between instrumentalism and realism, Wilkerson suggests a further gain in his reconstruction of Dennett's notion of patterns: avoiding the possibility that Dennett's view might ultimately clash with the hopes of contemporary neuroscience, that is, the hopes for a possible reduction of folk psychology to some kind of brain structure.

I do not believe that Wilkerson's contention of an expanded account of Dennett's view of the reality of the patterns can help resolve any tension between instrumentalism and realism--if there is such a tension after all. Dennett's approach leaves open the possibility that such *abstracta* may or may not map onto something in the brain. I do not suppose that Dennett would reject a possible mapping between folk psychology and some brain structure, despite his claim that such a mapping would be subject to indeterminacy. Furthermore, the adoption of the intentional stance does not mean that any attempt to unravel the neural correlates of mental phenomena should be abandoned. The issue does not bear on the adoption of either instrumentalism or realism (or eliminativism) as a resolution for an alleged tension in Dennett's approach--such labels fall short any proper characterization of his position. For Dennett's presumption in defining his position--a position he would like to call mild, intermediate or semi-realism--is nothing over and above a commitment to *naturalism*. Naturalism does not require eliminativism--not in the sense that patterns are to be reduced to brain structure in order to be *real*. So viewed, the ontology of such posits or *abstracta* does not seem to be one of Dennett's major theoretical concerns. Folk psychology and its patterns will continue to be as *real* as centers of gravity are whether or not their possible reduction to brain structure can be accomplished. The reality of patterns is supported by their usefulness in as much as they provide a shortcut for our cognitive apparatus to deal with the unpredictability of the behavior of some agents in the environment. Folk psychology may continue to prevail in our day-to-day explanations and predictions of behavior for its pragmatic/heuristic advantages provided by compression algorithms. It is very unlikely that folk psychology will, some day, be replaced or supplemented by more sophisticated concepts derived from neuroscience. As Dennett remarks, "A truly general-purpose, robust system of pattern-description more valuable than the intentional stance is not an impossibility, but anyone who wants to bet on it might care to talk to me about the odds they'll take." (p. 120). So viewed, the alleged clash between Dennett's instrumentalism and his realism (or eliminativism) seems to arise from a misinterpretation of the purposes of his theory. For *how* patterns could be generated by the brain or by some kind of virtual machine is a subsidiary issue--one that cannot substantially alter Dennett's account.

Let us now turn to the second part of Dennett's book. This part groups together essays on Artificial Intelligence and Artificial Life. At least two of them must be mentioned: "Cognitive Science as Reverse Engineering: Several Meanings of Top-Down and Bottom-Up" and "When Philosophers Encounter Artificial Intelligence". The first draws a pathway between Artificial Intelligence and biology and evolution through combining top-down and bottom-up strategies--a task where Artificial Life plays a prominent role. The second essay focuses on the relationship between Artificial Intelligence and

Philosophy, that is, how those two disciplines can possibly be intertwined. This is not a vexed question, nor a platitude as many may suppose. Many philosophers still tend to confine themselves and look at science and at Artificial Intelligence with some form of suspicious contempt. A fearful disdain arises in so far as many questions addressed by Philosophy are also addressed by Artificial Intelligence (What is mind? What is meaning? What is rationality?). Dennett's approach to this question is to show that it ultimately boils down to the issue of whether or not we are to take mind as an object of scientific study. However, such an approach cannot be fully attained unless philosophers relinquish some of their current prejudices concerning the dignity of the ultimate mystery, the Inexplicable Mind--a kind of pride which still lingers among philosophers. The issue is summed up in a nicety: the prejudice against the study of mind as an object of science is comparable to Aristotle's prejudice against extending earthly physics to the heavens.

Besides the two essays mentioned above, section II provides at least two more papers of historical interest, both written in a fairly agreeable style. "The Logical Geography of Computational Approaches: a View from the East Pole" avails to the reader a wealth of information about the development of Cognitive Science in the eighties. Of no less importance is Dennett's detailed review of *Unified Theories of Cognition*, where a thorough appraisal of the SOAR architecture and Newell's research program for Cognitive Science is presented and discussed.

The section devoted to Ethology and Animal Mind contains one of the most provocative essays written by Dennett in this field: "Animal Consciousness: what matters and why". Most significant is the fact that this essay was published in the fall of 1995 when his book *Kinds of Minds* was probably in its penultimate draft. The paper advances several views that were to appear in the book; views that still may stir indignation and suspicion among animal lovers. But surely this is not the purpose of the essay, unless one sticks to some biased misinterpretation.

No doubt Dennett's view on animal consciousness disavows some orthodox conceptions such as those supported by Thomas Nagel (1974) in his classic paper "What is it like to be a bat?", and, more recently, by Mary Midgley (1984) and Peter Singer (1981). However, the purpose is not to disqualify animal mind, nor to downgrade their suffering so as to generate a comfortable setting for hunters, farmers and experimenters. Rather, the purpose is to dismantle the myth that we can discuss animal consciousness (if there is any) by projecting on such creatures our own current conceptions of mind, consciousness and pain. The problem with such views lies in the fact that they are deeply rooted on the presupposition of the existence of a Cartesian Theater where conscious processes occur, or, in other words, on the assumption of the further substrate of a *res cogitans* besides brain-and-body activity. Furthermore, as is emphasized in *Kinds of Minds*, one must distinguish between mental competences and mental lives, that is, resist the habitual conception that any mental activity is accompanied by a stream of consciousness. It is this natural inclination, followed by authors such as Nagel, that renders the issues involving consciousness and also animal consciousness an impenetrable mystery or an insurmountable philosophical conundrum from which several ethical dilemmas may

emerge. By dismantling the mythical presumption of the existence of such a *philosophical ego* and its projection onto either ourselves or another species, one is likely to get a better footing to overcome such conceptual difficulties. This task, however, may require several previous steps, such as, for instance, reversing some acquired conceptions according to which animals feel and perceive the world in accordance to patterns that underlie our own organization of the world's experience. Such informational unification "is not anything we are born with, not part of our innate "hard-wiring" but in surprisingly large measure an artifact of our immersion in human culture" (p. 346). Furthermore, in support of this idea, Dennett says that "My claim is not that other species lack our kind of *self-consciousness* as Nagel and others have supposed. I am claiming that what must be added to mere responsiveness, mere discrimination, to count as consciousness at all is an organization which is not ubiquitous among sentient organisms" (p. 347).

This is a fascinating and original view of the nature of animal mind--one that emerges from disentangling an item of popular philosophical mythology deeply ingrained in our current intuitions about the mental life of other species. Examples in support of such a view are elegantly chosen and presented by Dennett throughout the essay. The only noticeable theoretical difficulty with such a view is that it leads to the postulation of *unconscious pains*--an undesirable consequence which is, nonetheless, acknowledged by Dennett in his attempts to conceive of animal pain as somewhat different from human pain. Apart from that, our current intuitions about animal minds can still be fairly tricky. Those who were brought up in rural areas are likely to be familiar with the gloomy scene of a pig when it is put to death. Any such recollection will invariably include the animal's desperate bawling when it is captured in the pigpen--as if it *knew* or *was aware* beforehand of its hopeless fate. This is not saying that Dennett's view is to be rejected but rather pointing to the huge difficulty posed by any attempt to reverse our current cultural constructs regarding the notion of consciousness. How far can we undertake such a task without incurring some unacceptably counter-intuitive claims?

The last part of Dennett's book contains only two essays, but at least one of them requires a more lengthy commentary since it addresses some paramount ethical issues emerging from the relationship between human beings and computers. This piece, entitled "Information, Technology and the Virtues of Ignorance" originally appeared in *Daedalus* in 1986. The essay points to the dangers of the progressive replacement of human beings by computers in tasks that involve ethical responsibility, such as, for instance, medical diagnosis and prescription of medical treatment. Although expert systems for medical diagnosis had only an ephemeral pinnacle in the seventies, the substitution of "artful investigation" by a technology of medical diagnosis is still on the way. The traditional medical practice as relying mostly on the physician's ingenuity and capability will, sooner or later, incorporate the use of expert systems and, at the same time, pose the issue of the delegation and transfer of a great deal of human responsibility to some specific technology.

In the course of this essay, Dennett seems to alternate between fascination and concern as to the use of expert systems in medical practice. Fascination comes in so far as the author does not suppose that the partial replacement of the personal (and personable) physician

can constitute a good reason to forestall the advance of technology--we cannot bet on the "virtues of ignorance". Such a replacement can, at worst, be viewed as a change of social roles, as happened in the case of calligraphers, potters and tailors who enjoyed a much more eminent status in their communities many years ago. Nostalgic feelings that may arise in the future are too flimsy a reason to prevent the use of new technologies. However, concern crops up in so far as in the use of new medical technology there is the implicit compliance to ceding responsibilities: to what extent can we rely on expert systems for medical diagnosis and treatment? Who are we going to blame in the case of fatal errors? There is more to this issue: the replacement of the traditional physician--illustrated by the figure of the rural doctor--involves the substitution of the traditional medical interview by the use of computer designed questionnaires. Doctors will tend not to deal personally with their patients. A lot of excitement and respectability, which is normally associated to the physician's social role will diminish. Furthermore, a career opportunity may lose its traditional enticement. In this case, says Dennett, "information technology, which has been a great boon in the past, is today poised to ruin our lives--unless we are able to think up some fairly radical departures from the traditions that have so far sustained us" (p. 368).

The issue posed by the advent of expert systems in our societies does not affect only the physicians, but all careers and all aspects of human life. We are seeing not only of the replacement of manual work, but also of mental work. Besides, the common sense expectation that technology would increase our leisure time had also to be abandoned. Surely this is not what is going on in our societies. Still, Dennett is right in suggesting that the solution for this issue does not lie in a radical forestallment in the advance of technology. No one could reasonably agree to such a view nowadays. The problem does not seem to be how to conceive of technology, but rather how we can reasonably conceive of ourselves before the technology we produce--a technology that forces on us the need of a thorough and quick revision of many philosophical conceptions, including ethical ones. Philosophy seems to have disappointed us as far as those issues are concerned. This is echoed in Dennett's essay when he points to the need for "moral first aid" and asks "What to do until the Doctor of Philosophy arrives"? But will the arrival of the Doctor of Philosophy make all that much difference? How quickly can she or he point to solutions? In other words, can human reason provide some solution to the ethical problems stemming from the production of human technologies?

An alternative to the Doctor of Philosophy is to expect that technology can contribute to pointing out solutions to problems generated by its own use. That would also include possible new emerging ethical issues. But such a hope seems to entail a confrontation of human reason with itself. Curiously enough, contemporary technology seems to have been prolific in devising experiments to test the proclivities and power of human reason. This is the case, for instance, of the celebrated chess-playing computer program Deep Blue. Most of the time, when we speak of a confrontation between human intelligence and machine intelligence we forget we are the ones who designed machine intelligence. We want to make sure that human intelligence can ultimately supersede itself in the form of a technology that, nevertheless, *we* have invented. Now, to what extent can human technology offer alternatives to the problems posed by its own development? Can an

answer to the ethical problems posed by the medical expert systems be supplied by Artificial Intelligence? Can we overcome ethical and moral dilemmas by designing an expert system to deal with such matters--a "moral first aid kit"--while we wait for the Doctor of Philosophy to arrive?

This is not specifically the question Dennett addresses in his essay, but I do not suppose he would be totally averse to discussing this latter alternative. But he could point to several obstacles to the realization of such a project. The most difficult to overcome is that it is not possible to foresee all the consequences of human action, let alone its ethical consequences. Dennett's view of ethics as a possible scientific discipline--at least as far as this matter goes--would model it as sharing all the uncertainty of meteorology. Like weather forecasting, the possibility of predicting the consequences of human action and its possible ethical consequences may be an almost insurmountable problem for the design of an expert system devoted to moral decision-making.

There is also a further question not addressed by Dennett in his essay. On one hand, expert systems may improve medical practice by avoiding possible malpractice suits. However, on the other hand such a technology may lead to a surreptitious downgrading of the standards that would normally be demanded in a medical school. This is not to say that medical schools will collapse because of expert systems. However, as a natural trend, we tend to spend less time studying what we know a machine can do for us--sometimes much better than we do. This is a possible perverse side effect of the use of such a technology--a side effect whose consequences can affect medical education. If this is the case, doctors will be less prepared. Not *all* doctors, though. There will be an elite of expert systems designers who will keep mastering almost all of medical knowledge. This is an even more dangerous risk stemming from the use of this technology, that is, the risk of widening up even more the present educational differences. The melancholic pre-technological "virtues of ignorance" would turn out to be a real danger of ignorance.

Will the use of expert systems in medicine imply that our future doctors (except for the elite we referred to above) will be quacks? This is likely to be true, at least as far as a first generation of doctors is concerned. However, we do not know exactly how such a relation between doctors and expert systems may develop in the future. It might be the case that such an elite itself will also disappear if medical knowledge expands to a point where no human doctors can have the mastery of all medical information. New kinds of expert systems will have to be designed; in this case expert systems to design other expert systems and, in so doing, summarize and keep up-to-date knowledge that would normally be attained by human specialists. If this is the prospect, even the elite will disappear and a new way of conceiving of knowledge and of knowledgeable people will have to replace our ordinary views.

There is much more I could say about Dennett's book, a collection which expresses twelve years of his intellectual trajectory and whose variegated approach to so many subject matters makes for fascinating and entertaining reading.



## Acknowledgments

The author is supported by FAPESP grant # 97-03518-6.

## References

Dennett, D. (1996). *Kinds of minds*. New York: Basic Books.

Haugeland, J. (1993). Pattern and being. In B. Dahlbom (Ed.) *Dennett and his critics* (pp. 53-69). Cambridge: Blackwell.

Midgley, M. (1984). *Animals and why they matter*. Athens: University of Georgia Press.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, 435-450.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Singer, P. (1981). *The expanding circle: Ethics and sociobiology*. Oxford: Clarendon.

Wilkerson, W. (1997). Real patterns and real problems: Making Dennett respectable on patterns and beliefs. *The Southern Journal of Philosophy*, 35, 557-570.