**Shadows of the Mind**

**Roger Penrose**

**Oxford University Press**

**David Papineau**

[written for a new supplement to *The Economist* in 1994 but not published as the supplement was never launched]

Roger Penrose is the Rouse Ball Professor of Mathematics at Oxford and one of our leading mathematical physicists. His first non-specialist book, The Emperor's New Mind, published five years ago, was a masterpiece of popularization spun along a thread of highly dubious argument. The sequel, Shadows of the Mind, is also a good read, but it magnifies the weaknesses and dilutes the strengths of its illustrious predecessor.

 As a writer of popular science, Penrose is a throwback to the age of authors who respected their readers enough to explain things properly. When he mentions complex numbers, say, or quantum mechanics, or Turing machines, he doesn't just wave his pen at their mysteries, but stops to go through the details, using equations where necessary, but also pictures, metaphors and the clear language of understanding. On this level, The Emperor's New Mind was a triumph, a compendium of everything intelligent people ought to know about modern physics and mathematics but never had time to study.

 Unfortunately, Penrose also used the book to articulate a theory linking up human consciousness, quantum mechanics and Gödel's famous theorem in mathematical logic. It would be generous to call this theory speculative. Both Gödel's theorem and quantum mechanics are certainly philosophically puzzling. But Penrose's suggestion that Gödel's theorem provides evidence for mysterious quantum powers in the human mind struck most informed commentators as moonshine.

 Even worse, they pointed out, it wasn't even original moonshine. Penrose's central claim that Gödel's theorem shows the human mind is not a computer was originally put forward by the Oxford philosopher John Lucas thirty years ago. Since then this thesis has been thoroughly thrashed out in hundreds of books and articles by logicians and philosophers, and the overwhelming consensus is that no such conclusion follows.

 In Shadows of the Mind Penrose seeks to respond to this criticism and other objections to his original thesis. Once more he ranges far beyond the confines of his argument, and there are many fascinating digressions down the by-ways of contemporary science. But overall the tone is rather less generous than in the earlier book, and sometimes this volume reads like an essay by a bright but obstinate student who hopes that if he keeps arguing long enough his opponents will give in.

 More than half the book is on the Gödel-Lucas issue. The essence of Gödel's theorem is that no system of proof procedures can capture all the truths of mathematics. This is surprising enough, but even more curious is the fact that Gödel's theorem itself seems to prove some of the unprovable truths. Penrose, following Lucas, infers that Gödel's theorem gives human consciousness some kind of non-systematic access to these unprovable truths. The alternative view, however, is that there are systems and systems, and that Gödel's theorem simply shows that no system can consistently identify and endorse its own proof procedures, whether it be a human or a machine. Penrose explores these issues in great detail, and insists that this alternative view is implausible. But even after 200 pages I remained unpersuaded, and I wondered whether Penrose had sufficiently considered the obstacles facing self-referential acts like endorsing your own proof procedures.

 The other half of Penrose's original theory was that quantum mechanical processes in the brain are responsible for the special mathematical powers of human consciousness. Quantum mechanics is mysterious because it represents the microscopic world as essentially fuzzy, yet fails to explain why this fuzziness never shows itself in the world of human experience. In *The Emperor's New Mind* Penrose hypothesized that it might be gravity at the quantum level that is responsible for removing the fuzziness.

 In this new volume he is much more specific on how this gravitational reduction of the fuzziness is supposed to work. As he admits, there is as yet no empirical support for his hypothesis, and the equations he proposes seem arbitrary, but in defence it could be said that many other contemporary attempts to solve the puzzle of quantum mechanics are just as bad.

 Where Penrose seems to go right off the rails, however, is in trying to explain how the gravitational reduction of quantum fuzziness helps the human mind to do mathematics. In the earlier book this was left mercifully obscure. But here he speculates that the microtubles in brain cells may somehow house a distinctive style of quantum computation. As always, his account of the scientific facts themselves is fascinating. But he never really explains where the mathematical powers come from, and the longer he continues waving his hands the less rational the overall theory seems.

 Readers who enjoyed *The Emperor's New Mind* will no doubt be pleased to have more of the same. But I hope that in his next book Penrose tries something different, and gives up his theory of consciousness as a bad job.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

David Papineau is Professor of Philosophy of Science at King's College London. His last book was Philosophical Naturalism (Basil Blackwell, 1993).