Cognitive Science and the Twin-Earth Problem

J. A. FODOR

Introduction   There's something odd about the history of cognitive theories. On the one hand, practically all of them, from Descartes forward, have been thoroughly committed to mental representations as explanatory constructs. But, on the other hand, a continuing critical tradition in both philosophy and psychology argues that the mental representation construct is inherently defective and cannot be made scientifically respectable. This has been going on for a long time.[1] It's a bit as though physics had developed in parallel with a line of criticism which claimed that the notion of a particle is incoherent and must be dispensed with. Surely, one would think, some sort of resolution should eventually be achieved: either the criticisms should be shown to be misdirected, or we should give up the construct criticized. One would think, too, that there ought to be some way of telling whether one's theoretical commitments are incoherent, and that three hundred years or so ought to be long enough to find out.

   Anyhow, the sky is falling again. We have a cognitive science whose main tenet is that the mind is a device for the manipulation of representations. But we also have a line of philosophical criticism that goes like this: Nothing is a representation except insofar as it has representational content, and the notion *content of a mental representation* is in jeopardy. In particular, there's a new argument that is taken to show that, even if there are mental representations, and even if mental representations have contents, still the content of a mental representation is not a function of psychological variables as cognitive scientists understand such variables. So, to that extent, the notion *content of a mental representation* is not available as an explanatory construct in theories of the sort that cognitive scientists have hoped to develop.

   Now, the arguments currently fluttering the dovecotes are actually rather