The Representational Theory of Mind, by Kim Sterelny. Oxford: Basil Blackwell 1990. Pp. xiii + 252. £37.50 h/b, £12.50 p/b.

The aim of Kim Sterelny's book is to introduce a particular theoretical approach that shows how the adoption of the Representational Theory of Mind is compatible with an endorsement of physicalism. The Representational Theory holds that propositional attitudes, such as beliefs and desires, involve distinctive relations to representations. Thus, the sentence "John believes that the house opposite is on fire" is taken to have as its truth condition the fact that John is standing in a relation distinctive of belief to the representation in some way identified by the sentential clause following the "that" in the quoted sentence. The difficulty is to see how the existence of such truth conditions is compatible with the thesis that physicalism is true.

The approach that Sterelny wishes to defend was pioneered by Jerry Fodor. Sterelny characterises it by the following theses:

> Thesis 1: Propositional attitudes are realised by relations to sentences in the agent's language of thought. (p. 29)
>
> Thesis 2: The psychologically relevant causal properties of propositional attitudes are inherited from the syntactic properties of the sentence tokens that realise the attitudes. (p. 30)
>
> Thesis 3: The semantic content of propositional attitudes are explained by the semantic properties of mentalese. The semantic properties of a token of mentalese are explained by its syntactic structure, and the semantic properties of the concepts that compose it. (p. 32)

If the "language of thought" is a physical code in the brain and theses 1 to 3 are also true, then it would seem that the Representational Theory of Mind is compatible with physicalism.

The book is concerned with issues which arise if one accepts the approach recommended. The work is unashamedly interdisciplinary in character. Sterelny is no respecter of the distinction between philosophical and psychological approaches to understanding the nature of mind. In some ways this is all to the good. It would be very surprising if these two disciplines did not trespass upon each other even if those who see philosophy as continuous with science are wrong. Nevertheless, the combination of being both introductory and interdisciplinary sometimes results in a discussion that is conducted at a high level of generality leaving the reader somewhat unable to assess for him or herself the strength of the motivation for the positions adopted. One is inclined to search for the primary sources in order to understand more fully what is going on. At one level, this is a result that Sterelny should welcome. Indeed, it is inevitable, bearing in mind that the book ranges over subjects as diverse as the nature of psychological explanation, whether or not the mind has modular organisation, individualism, eliminativism, naturalist theories of content, connectionist models and the explanation of intelligence. Nevertheless, it must also be considered a flaw in any book whose aim is to provide an introduction to the field.

One area in which things might have been clearer is the treatment of theoretical frameworks and explanatory power. Sterelny endorses the familiar view that a theory of a certain type is warranted if and only if it gives us additional explanatory power, where this is understood in terms of providing "access to generalisations unstatable within other rival frameworks" (p. 206). One problem with such an approach as it stands is that it seems to imply that any theoretical framework that proposes to identify new types of entities is warranted. Even if the entities do not exist, otherwise unavailable generalisations concerning them could be stated. On the other hand, one cannot require that the generalisations be true because, first, it is surely possible that a certain type of theory is warranted even though it is not true, and second, if one has to assess whether a generalisation is true, we have not been given any further guidance as to when one should adopt a certain type of theory which, one presumes, was the aim of putting forward such a proposal.

The idea that Sterelny has in mind becomes clearer when we consider how it might work in assessing a way of understanding psychological theorising made famous by David Marr. In Sterelny's terms, there are three levels of such theorising: the *ecological*, which is concerned with what cognitive capacities the mind has; the *computational*, which is concerned with how those capacities may be achieved by processes involving the manipulation of symbols according to rules; and *physical implementation*, which is concerned with the physical nature of that which is said to possess a mind. The recommended approach to the Representational Theory of Mind depends upon the legitimacy of the computational level of psychological theorising and, Sterelny says, this is a matter of:

> 1. *Cognitive uniformity*: Our representational structures, and the procedures that access them, are near enough identical across the species.
>
> 2. *Neural diversity*: The physical implementation of these structures is enormously varied. (p. 212)

Although many have endorsed the intuition that the entities identified at the computational level are likely to be "variably realised", and, therefore, have elected to remain neutral about what the matter of mind may be, it is surely not advisable to rule out the possibility that only certain sorts of matter may realise minds. For all we presently know, the computational level may require some specification of the details of physical implementation and thereby consist of generalisations that fall within the latter theoretical framework. In spite of this one may reasonably maintain that the type of theorising that goes on at the computational level is genuinely distinct from, though not independent of, that which goes on at the level of physical implementation. One could be interested in the brain as symbol processer even if one thought such symbols could only be realised in certain sorts of matter. In contrast, Sterelny's methodology would have us conclude that there is no computational level. But why? May we not in the situation envisaged say that different terminologies and approaches alighted upon the same types, from different angles? The difference in levels would be seen as a consequence of adopt-

ing different focuses of interest rather than as a consequence of there being something else left to explain (p. 209).

It might be argued that any generalisation which abstracts away from physical detail would not be a generalisation formulated within the theoretical framework of physical implementation. However, it is hard to see why such an approach is acceptable. Talking in general terms about the character of atoms abstracts away from physical details, namely of what elements the atoms are atoms, but that would not mean that such talk is not developed within the framework of physical implementation.

There is a second reason for being sceptical about a justification for a theoretical framework that rests solely upon the accessibility of additional generalisations. Arguably, an inescapable feature of all such frameworks, other than, perhaps, the level of physical implementation, is that the generalisations hold *ceteris paribus*, or, maybe, given certain definite qualifications. The problem is that such generalisations are cheap. One can state a generalisation between two things that we would say was of no theoretical interest but which we might be right in claiming holds ceteris paribus. For instance, even though *ceteris paribus* all kitchen tables have kitchen utensils upon them at least once a day, it does not follow just from this that there is a budding science of kitchen usage. Strangely, Sterelny seems cognisant of this fact (p. 210) but unconcerned about its ramifications for his methodology.

Sterelny chooses to situate his discussion of individualism, crudely the thesis that mental states supervene solely upon a subject's internal constitution, within an assessment of the legitimacy of the ecological framework. He holds that the entities identified by this framework do not supervene upon a subject's internal constitution. So, if it turned out that the framework was legitimate, and that mental states of the relevant kind were identified by that framework, then it would follow that individualism is false.

His discussion faces two preliminary problems. The first is how one can legitimately count certain mental states, specifically those with propositional content, as part of the ecological level. The latter is said to identify capacities (p. 44). What capacity do ascriptions of such mental states attribute?

The second problem is that it is by no means clear that the ecological level is committed to identifying entities that do not supervene upon a subject's internal constitution alone. It is, of course, true that for a capacity to be manifested, the entity which possesses it must be in a certain context. It does not follow that possessing a capacity is not just a matter of the entity in question being a certain way.

To illustrate the point, consider an example of Sterelny's:

> if we had reason to attribute to bats and owls the same psychological state—say that they both perceive mice—then that state could hardly be individualistically defined. Perceptual systems vary greatly; their only common feature is that their function is the extraction of information for the adaptive control of behaviour. (p. 98)

We may concede that the internal constitution of the visual systems of bats and mice are very different but it does not follow that the function (or capacity)

described at the end of the passage is "anti-individualistic". Given the point about capacities, what we should say is that bats and owls show that certain capacities are variably realised. There are at least two ways in which an entity may have the capacity to obtain information about mice.

Sterelny may be unconvinced by this line of reflection because he holds that the perceptual systems could only really be ascribed the capacity to detect mice given that the organisms which possess them live in a certain environment and not just because of the internal constitution of these organisms. However, we could say that bats and owls do not really possess the capacity to detect mice per se, but the capacity to detect mice in certain environments. It is tempting to think that Sterelny's failure to discuss this alternative is a consequence of the emphasis he places upon the methodology we have already discussed. Scientifically interesting generalisations involving his preferred capacities are inaccessible as stated if we have adopted the alternative way of individuating capacities. This might have been thought sufficient to justify reference to the former capacities.

A second way in which Sterelny's methodology potentially distorts his discussion is revealed when he turns to examine Fodor's objection to an account of intentionality in terms of biological function. Suppose that the owl's "mouse detector" is also triggered by small rubber balls rolling along. The question is why one should take the detector to be a mouse detector as opposed to a mouse and rolling rubber ball detector. Those who appeal to biological function suggest that the answer is that the detector is only supposed to detect mice because it was that for which it was selected during the course of evolution. Fodor's objection is that one may just as well say that it was the mouse and rolling rubber ball detector which was selected because, in the circumstances, the detector in question gives the owl an evolutionary advantage even if it makes mistakes.

Sterelny rejects this objection. His reasons are: first, if one was tempted to classify the content of the detector in terms of what is phenomenally common to mice and small rubber balls one would lose the generalisations available by citing the interaction with mice; and second, if one thought that it was not clear whether in fact the regularity held between mice and owl behaviour or between mice or rolling rubber balls and owl behaviour, one is, in effect, confusing epistemic issues with metaphysical ones.

It is reasonably clear how our discussion of Sterelny's methodology throws into question the first reason offered. What of the second? Sterelny believes that the problem that Fodor has identified is akin to the issue Nelson Goodman raised in his discussion of grue. In the present case, the charge amounts to the claim that although we can't prove that nature has selected for a mouse detector rather than a mouse or rolling rubber ball detector, there is a fact of the matter as to which has been selected for.

The difficulty with this analysis of Fodor's objection is that the two alternatives that Fodor specifies seem equally able to capture what is happening in the selection process the disagreement lies in how it should be classified. In the case of grue there is a disagreement of fact. We cannot justify our preference for a law

citing green rather than grue, but we believe that the facts would be different if a law correctly cited one rather than the other. Sterelny rejects this point of disanalogy. He argues that had the environment been different so that the inability to distinguish between mice and rolling rubber balls was more detrimental to the owl's chances of survival, a better detector would have been developed. Thus, in our environment what nature is selecting for is a mouse detector. However, considering counterfactual possibilities is a questionable way of determining what the actual biological function of something is since, in effect, by considering a different environment you are considering what would be selected for by a different selection mechanism. It is the actual selective history that determines the biological function of something. So, the disanalogy seems to stand. (Thanks to Kim Sterelny for pointing out that I should try to make a previous version of this criticism clearer).

The objections that have been raised above should not obscure the fact that much of the book is interesting and, to my eyes, correct in what it says. As is so often the case in philosophy, what is primarily an introduction has much to teach those who are more familiar with the field.

Clare Hall
Cambridge CB3 9AL
UK

PAUL NOORDHOF

Analysis and Metaphysics: An Introduction to Philosophy, by P. F. Strawson. Oxford: Oxford University Press, 1992. Pp. viii+144. H/b: £ 22.50 , P/b: £ 7.95.

This is a summing-up by a distinguished elder statesman of philosophy, a brief encapsulation of Strawson's life-work, based on his lectures at Oxford. The main themes of *Individuals, The Bounds of Sense, and Freedom and Resentment* are here, distilled into an elegant quick survey, with only a few new touches.

Strawson sets out on an "analytical" approach to philosophy, as opposed to any existential reflection on the human situation which might lead to "a new and revealing vision". It soon emerges that his favoured kind of analysis is "connective" rather than reductive, yet systematic rather than piecemeal (or Wittgensteinian). Concepts are to be illuminated by showing their connections with other concepts, their places and functioning in our overall conceptual system, rather than by analysis into ultimate simples. Some concepts are shown to be basic in the sense of being pervasive, irreducible features of our ordinary thought and talk, or even perhaps in the stronger Kantian sense of being necessary features of any possible conceptual scheme.

In Strawson's approach, ontology, epistemology and logic (in a wide sense including the theory of meaning and truth) are aspects of one unified inquiry. He defends what he takes to be the central tenet of empiricism—that "concepts of the real can mean nothing to the user of them except in so far as they relate, directly or indirectly, to possible experience of the real" (p. 52). But he rejects "classical"