Models and Cognition, by Jonathan A. Waskan, MIT press, 2006, pp. 344, cloth, $36.00.

In this adventurous book, Jonathan Waskan attempts to use results from cognitive science to show how philosophers of mind have erred in accounting for thought and philosophers of science have likewise erred in accounting for what explanation consists in. He puts forward a solution in terms of the cognitive equivalent to *scale models*, such as the cardboard models an architect makes of the house he intends to build. Chapter 1 provides a short historical account of the development of cognitive science, chapter 2 surveys the relation between cognitive science and folk psychology and chapter 3 presents an account of mental content. These chapters are aimed at the non-specialist reader, with the aim being to convince the reader that cognitive science vindicates folk psychology and hence there is no scientifically motivated objection to being realists about beliefs, desires an the like.

The main claims of the book appear in chapter 4, in which Waskan contrasts what he calls the "logic metaphor for mental representation and inference" (roughly, the language of thought hypothesis) with the "scale-model metaphor" (p. 107). He argues that approaches based on the logic metaphor all succumb to the *frame problem*, the problem of specifying the *non*-effects of a particular action. Imagine a blue ball in a bucket. If the bucket is moved, we typically infer that the ball moves as well, as the ball remains in the bucket, but not that the ball changes colour. Colour change is not (typically) one of the effects of moving the ball, but it is perfectly consistent with the description of the situation that the ball *will* change colour when moved. Non-effects of change can be added as explicit rules, known as *frame axioms*, to the effect that moving the bucket will not alter the colour of the ball. Typically, huge numbers of frame axioms are required, each of which must be qualified in innumerable ways (for example, the ball does change colour if the bucket is moved into a tub of paint). The worry is that such approaches cannot possibly scale up to realistic scenarios whilst remaining computationally tractable.

Waskan contends that the frame problem is "so serious" for the logic metaphor "as to make the search for a viable alternative appear downright mandatory" (p. 110). The alternative he provides is again given as a metaphor: the *scale model* metaphor. The thought is that human planning, for example, is best explained by supposing that "we manipulate the cognitive counterparts of scale models" (p. ix). The thought is that *actual* scale models do not succumb to the frame problem. A model of a blue ball in a (hole-free) bucket will continue to model the ball as being blue and being in the bucket when the bucket is moved. The task that Waskan undertakes is to show that the scale model metaphor can explain other aspects of thought just as successfully as the logic metaphor and, just as importantly, that it is compatible with current neuroscience. Before addressing these arguments, some interpretive issues should be addressed.

The frame problem was initially noticed by Hayes and McCarthy in the early days of (logic-based) AI. The problem posed for logical AI is that the predictions generated by the formal theory (such as Hayes and McCarthy's situation calculus) are not correct, unless a huge number of frame axioms are relied on, each having a huge number of qualifications. The resulting framework cannot scale up from toy examples to remotely realistic domains whilst remaining computationally tractable. This is an *logical* engineering problem: either the way that such theories represent situations, or the notion of inference that they reply upon, is not suited to the development of artificial agents. This does not appear to be the

problem that Waskan has in mind, however. There was a large amount of work on the logical frame problem in AI during the 1980s and a good number of workable solutions were arrived at, including Sandewall's notion of *occlusion*, Reiter's use of successor state axioms, the fluent calculus and the event calculus. Waskan does not refer to any of this logical engineering work, presumably because the problem that he envisages for 'logic metaphor' accounts is a philosophical, not purely an engineering, problem.

In epistemology, by contrast, the frame problem poses a puzzle about rationality. Dennett and Fodor both note the problem as concerning how an agent can determine which of its beliefs to alter as a consequence of change in the world. Agents typically entertain huge numbers of beliefs, most of which are accurate, mundane truths. How can such agents rationally keep track of their beliefs in a changing world? But this is not how Waskan frames the problem; in fact, rationality is hardly discussed at all in the book. So the epistemological version of the frame problem does not appear to be the problem that Waskan has in mind either. But, if the problem he envisages for 'logic metaphor' accounts is neither one of engineering logical systems that do not suffer form the frame problem, nor that of explaining how agents can *rationally* maintain and update their beliefs as the environment changes, what is the problem? I cannot find an answer to this question in the text.

Chapter 5 attempts to show that the scale model metaphor is able to explain *prima facie* problematic areas of thought: singular thoughts, negative and disjunctive thoughts, general thoughts and thoughts about abstract objects. The aim is to show that the scale model metaphor can replace the logic metaphor as a metaphor of human thought entirely, not just in the realm of our reasoning about changes in the world. The claim is that particular thoughts should be conceived as mental models, not mental sentences.

The worry about capturing singular thoughts is this. Suppose that entertaining the thought that Bob is fat just *is* having a mental model of Bob, depicting him as being fat. This model will clearly represent Bob as being many other ways as well: being human, for example. But the thought that Bob is fat does not imply that Bob is human, who might well be a rotund family pet. Waskan admits (p. 141) that "neither pictures nor scale models can, *by themselves* ... single out particular properties of particular objects". His response is to claim that "we have the ability to selectively attend to particular objects at the expense of others or to certain (local and relational) properties of objects" (p. 141) in the case of physical situations, so why not in the case of mental models? But I defy anyone to attend to Bob's (real or modelled) fatness without thereby noticing whether he is (or is modelled as being) human, feline or whatever.

Thoughts about certain abstract objects are equally hard to explain on the mental models approach. I have no idea what model would correspond to the thought that 1 is a natural number, or the thought that the power set of $\mathbb{N}$ has a larger cardinality than $\mathbb{N}$. Waskan's view is that "thoughts about such processes and entities involve some reliance upon analogies and metaphors that are rooted in domains that we *can* represent with images and models" (p. 139). Maybe so, but this is a long way from the claim that these thoughts are *themselves* constituted by models from analogous domains.

One worry that Waskan is keen to allay is that *there are no scale models in the head*! Chapter 6 argues that models, just like logical languages, can be implemented in computational systems, such as virtual reality systems and the finite-element models used in engineering. Waskan claims that there is a level of description that singles out physical scale

models in abstraction from their physical realization, and similarly a level of abstraction at which virtual reality models (say) are distinct from the syntactic programs in which they are constructed (p. 193). At a suitably high level of description, non-sentential computational models really do exist and, given the familiar computational analogy, the thought runs that "humans harbor and manipulate specific, intrinsic cognitive models". They do so "quite literally", albeit "at a high and "distinguished" level of abstraction" (p. 195).

The ontological claim here may well be overstated. Distinct physical models can represent a single situation, but this does not imply that there exists a representation of that situation somehow over and above the physical models. And nothing that Waskan says precludes instrumentalism about (physical, computational or cognitive) models. We might for explanatory purposes talk *as if* models exist (independently of particular manifestations) without this troubling us when we do ontology. (Programmers talk about the 'objects' of object-oriented programming languages, for example, but writing such programs does not obviously create new entities.) Regardless, if computational models do provide a suitable analogy for explaining our thought processes and such models are constructed by a program (i.e. lines of code), doesn't such code provide a suitable analogy for a language of thought?

Chapters 7 to 9 change the focus to theories of explanation. The thought is that models, rather than theories, provide the best account of what constitutes explanation. Chapter 7 surveys some worries for the Deductive-Nomological account of explanation, according to which to explain an event is to subsume it under a law. In its place, chapter 8 offers the *Model model*, according to which one has an explanation of an event or regularity when one has a cognitive model of the physical process underlying the event or regularity. The Model model is most comfortable in explaining the spatial manipulation of medium-sized objects in physical processes. How to accommodate philosophical or mathematical explanation in this way is not so clear. How could a formal proof, or the existence of larger and larger infinities, or a substance dualist's account of the mind, be explained by something analogous to an architect's model?

Perhaps the greatest benefit of this book will be to students curious as to how philosophy of mind and cognitive science can interact. Each chapter is summarized at the beginning and ends with a brief recap. The writing style is easy going throughout and the book makes for good reading, perhaps at the expense of detail in some of the arguments. But without a precise specification of the commitments of the logic metaphor and a more detailed account of the problem that the frame problem poses for these commitments, I think it unlikely that supporters of the language of thought hypothesis will feel compelled to abandon their approach on the back of Waskan's arguments. And whilst Waskan's book provides an entertaining account of the fascinating interplay of ideas between philosophy and cognitive science, I feel it unlikely, given these considerations, that we will witness the change of direction in cognitive science or philosophy of science that Waskan hopes for.

Mark Jago
University of Nottingham
mark.jago@nottingham.ac.uk