A Mark of the Mental, by Karen Neander. Cambridge, MA: MIT Press, 2017. Pp. xv + 327.

1. Neander's project

Philosophers have long pondered over the relationship between folk psychology and the conception of our mind developed by the cognitive sciences. How do beliefs, desires, wishes, and the like relate to the account of cognition given by neuroscience, cognitive neuroscience, and cognitive psychology? These propositional attitudes have remained recalcitrant to scientific absorption due to their 'aboutness' or 'intentionality'. Dave's belief that there is a lion by the waterhole is only a *belief* insofar as it is *about* something; to wit, lions and waterholes. That is to say, beliefs have *content*. Moreover, the content of Dave's belief has a key normative component: it is *true* or it is *false*, depending on the external world. But it is about lions and waterholes even when there are no lions or waterholes around. Propositional attitudes represent the world, and the project of naturalising mental representation has been a central philosophical project for the last forty years or so. It is the project of explaining aboutness in terms of the mechanisms and interactions recognised by the cognitive sciences. Neander's book is an important, though deliberately modest, contribution to this project. She aims to give a theory of non-conceptual, perceptual representation. She discusses some features of human vision. But her main working examples are the prey-recognition system of anurans, such as T5-2 cells located in the optic tectum of the toad. This structure is complex enough to be a plausible example of representation rather than a mere causal relay; simple enough to be tractable; and obviously non-conceptual. Her further aim is to give an account of the mental representations involved in the propositional attitudes (or whatever person-level cognitive states there may be, if talk of propositional attitudes is discarded), using her account of non-conceptual content plus the other resources of the cognitive sciences. There are some brief hints about this further project, but the central aim of the book is to give a naturalised account of perceptual representation. This is a fine book: clear, restrained in its claims, measured, and lucid in argument. We think Neander *does* give a defensible account of perceptual content. But we do see places where pressure can be applied to her framework.

We begin by sketching that framework. Beginning with Stampe and Dretske in the late 1970s and early 1980s (Stampe 1977, Dretske 1981), one popular approach was to link beliefs to what they are about via causal-informational relations. Dave's belief that there is a lion by the waterhole is about a lion and a waterhole, because his belief correlates with, or is caused by, a lion and a waterhole. This view is too simple, as it makes misrepresentation impossible. An important improvement was to combine the causal notion with a functional one: Dave's belief is caused by a stimulus processed by mechanisms which have the function of detecting lions and waterholes. Since

mechanisms can fail or malfunction, this proposal explains the possibility of misrepresentation. Function-invoking approaches to the naturalisation project are known as teleosemantic theories, and Neander's book is a defence and elaboration of one version of teleosemantics: causal-informational teleosemantics.

Accordingly, Neander takes the content of a (non-conceptual) mental representation to be the state of affairs that is supposed to cause it. Consequently, she defends a notion of proper biological function in Chapter 3, and applies it to cognitive organisation. It is common in science to explain the operation of a complex system by compositional analysis: to show how the behaviour of the system as a whole results from the interactions amongst its components. Neander clearly sees that she needs to make a stronger claim about cognitive systems. Climate scientists explain the earth's climate system in terms of interactions between its components (incoming sunlight, the different gases of the atmosphere, the layout of the continents and oceans, and so forth). But extreme weather events are not malfunctions. She needs the idea that cognitive systems are organised systems in a richer sense than the climate system: the components play roles that they are *supposed to play*, not just that they typically do play. Cognitive components have proper functions: jobs that natural selection (or developmental equivalents like conditioned learning) has designed them to do. So she needs to defend a notion of proper function, and to show that the cognitive sciences rely on proper functional specifications of cognitive systems in explaining how they work, not just why they work. It also requires her, in Chapter 6, to defend a particular *kind* of proper function: a *response function*. This is in order to rebut output-orientated teleosemanticists who argue that proper biological functions can't be backward looking: functions are what things do, not the causes of those doings. Neander shows that a perfectly legitimate proper biological function of some systems is to activate in response to some upstream state of affairs, and being thus activated, to produce some response. Indeed, we share her puzzlement that this needed to be argued: surely all homeostatic functions are response functions.

So when some neural device designed to track something in the world fires — prey, in the visual system of a hungry toad; or a line of particular orientation, in the visual system of a human — that pattern of neural activation has both aboutness and normativity. The former represents the triggering stimulus as prey, or some more proximate indicator of prey, because that is the target it is designed to track. A similar thing holds true of the latter. The representation can be accurate or inaccurate (hence the normativity), because the mechanism can function properly or misfire. As we have noted, the representations Neander seeks to ground in causal-informational relations and biological function are not the states of folk psychology. She herself characterises the theory advanced in the book as 'a causal-informational version of teleosemantics for non-conceptual representations' (p. 142). That

project makes sense. Neander rightly points out that cognitive scientists commonly speak about representational states at the subpersonal level: for instance, when explaining early visual processing in humans. Neural states in early visual processing are taken by cognitive scientists to represent; they represent edges, shapes, faces and the like. Moreover, these representations can be veridical or erroneous: for example, task failure is explained by inaccurate representation. So there is a good case for the view that the cognitive sciences do treat subpersonal brain states as possessing both aboutness and normativity. Chapters 2–4 maintain that causal-informational teleosemantics, in particular, is implicitly supported by mainstream cognitive science. These chapters assume that cognition involves processing information, and that information grounds the aboutness of representational states. They assume that the biological functions of these representations specify their accuracy conditions. However, the project is not *only* to vindicate a certain style of scientific explanation in the cognitive sciences. One of the crucial assumptions of this book is that these representations are genuine intentional states. Neander's hope is that by successfully grounding the intentionality of non-conceptual, sensory-perceptual representations we make modest, but genuine, progress in the quest to make a home for the propositional attitudes (perhaps with some revisions to folk psychology) in the scientific image. We return to that hope at the end of this notice; we now turn to some pressure points for the modest project.

2. Do the neurosciences need proper functions?

As we have just seen, Neander defends the idea that cognitive theories of visual processing, including early visual processing, depend on a notion of function that includes the possibility of malfunction, and likewise depend on a notion of representation that leaves open the possibility of misrepresentation. That is so even though these are theories of proximate causation, and so whether a representation is veridical and whether a perceptual mechanism has responded to its design target make no difference to the upstream cognitive mechanisms which use the representation. It is clear why appeals to selection and to the proper function of a component help to explain why the system is as it is. But given that these external factors are screened off from causal interactions within the toad's perceptual system, why are they relevant to explaining how that system works? Neander's response is to point out how variable, idiosyncratic and messy real biological systems are. Even in an organism as (relatively) simple as a toad, the wiring diagram of each toad eye will be unique. The point is general: organisms are not just extremely complex, they are endlessly variable, even within the same species. However, despite that variability, selection makes it possible to build a composite or collective portrait of a typical member of a species (or of a system within the organism). For selection typically builds something that approximates a single design. We use that to construct our composite portrait. While

pathologies and other departures from that design can be endlessly varied, a specification of organic design in terms of the proper functional organisation of a system enables us to explain how the system works when it is working properly, and to explain and predict departures from functionality, in terms of the malfunction of specific components. A design stance specification of an organic system is a sweet spot; it abstracts away from enough of the detail and idiosyncratic variation to be tractable, but there is enough detail about the causal mechanisms to make prediction, explanation and intervention possible. (See especially pp. 59-65.)

Ecology poses a problem for this analysis. Ecosystems and communities often have apparent organised complexity. As with Neander's poster cases of cognitive neuroscience, they pose issues of massive and idiosyncratic variation across tokens of a type (wetlands, semi-arid savannahs, and so on). Yet prediction, explanation, and generalisation are not impossible. Think, for example, of the concept of a keystone species in a community, and the role of such species in promoting both diversity and stability. Think also of the various versions of hypotheses that link the stability of an ecosystem to its biotic diversity. We face the same challenge in constructing a collective portrait of, say, a grassland community: one that captures enough of the causal machinery to be a suitable basis for intervention and prediction, but is not swamped by idiosyncratic detail. Yet etiological notions of function do not apply. Neither ecosystems nor communities have selective histories of their own. (See, for example, Sterelny (2006).) So in constructing a collective portrait, it seems that we can get by with a minimal functional notion, perhaps of the kind used to describe the climate system. If this is true of ecosystems, why not toad vision? So Neander shows that the neurosciences do use a notion of proper function in analysing their systems of interest, and that they do so in response to a real challenge. But she does not show that they need such a rich notion of function to respond to that challenge.

The problem of idiosyncratic detail, and of finding a sweet spot between capturing too little detail and being overwhelmed by idiosyncratic variation and complexity, has been much discussed in the life sciences. In different versions, Levins (1966), Dennett (1991), Wimsatt (2007) all suggest a strategy that differs from the strategy employed by Neander. Construct a variety of collective portraits, with the choice depending on the specific explanatory or predictive interests of the investigator. These will often specify functional organisation, and hence, presumably, content in different ways. We would get some form of indeterminacy or pluralism about content. Neander is averse to this move. She points out that the naturalisation project fails if the content of some target system's representation depends on the beliefs or intentions of the investigator. It is true that if the naturalisation project is to be viable, the investigator's thoughts cannot play any role in determining the representational content of states in the system under study. (See pp. 57, 157.) However, the pluralist strategy does not require that. We think Neander

can afford to be more relaxed about content pluralism and the availability of multiple functional specifications of the same system, and hence can afford to explore a wider range of options. Instead of saying that a particular functional attribution is made true by investigator intentions, say instead that the objective facts of selective history (and the other relevant factors) make many functional specifications true of some organism system. One of those may be more salient to the predictive and explanatory projects of any particular investigation. So, to take one famous indeterminacy problem in this literature, it is plausibly true that a particular structure of the toad's visual system has the function of detecting prey, by detecting a characteristic configuration of shape and movement, by responding to a particular pattern of reflected light and changes in that pattern, and by detecting and transforming a particular pattern of activation in the toad's retinal cells. Perhaps these all support a content attribution to the activity of that structure. What is wrong with saying all these content attributions are true of the toad? The naturalisation project is not threatened by pluralism about content. For the objective facts of selection and evolution ground all of them. In the next section, we focus on these issues of indeterminacy and on Neander's solutions to them.

3. Is teleosemantic content determinate?

Chapters 7–9 of the book offer a theory of content for non-conceptual perceptual representations like those used early on in human visual processing and in the prey capture behaviour of anurans (frogs and toads). The theory is a causal-informational version of teleosemantics (CT for short). Neander's CT 'starter theory' is as follows:

A sensory-perceptual representation, *R*, which is an (*R*-type) event in a sensory-perceptual system (*S*), has the content *there's C* if and only if *S* has the function to produce *R*-type events in response to *C*-type events (in virtue of their *C*-ness) (p. 151).

Put more simply still, 'a sensory-perceptual representation refers to what is supposed to cause it'. Neander also seeks to show that CT avoids various indeterminacy problems that have plagued reductive theories of intentionality. To address indeterminacy, Neander turns to anurans and their prey-capture behaviour. Neander's use of the anuran as a model system carries real weight because, unlike many in the literature she has come to grips with the neurobiology of toad vision. In Chapter 5, a strong case is made that T5-2 cells located in the optic tectum of the toad host a representation of a visual target used by the toad in hunting prey. These cells are not mere causal relays in the toad's nervous system. Above and beyond this they are mental *representations* of targets. But what is the content of this kind of non-conceptual perceptual representation? Is it FOOD? WORM? Or is it WORM-LIKE MOTION? All are co-extensional with T5-2 activation when T5-2 activation occurs as it should. Many will recognise this well-trodden

territory. Adding another breath of fresh air to the debate, Neander goes beyond intuition in Chapter 5 to advance a pre-theoretic constraint on a theory of content for non-conceptual perceptual representations. She appeals to the explanatory aims of computational vision science. According to mainstream, computational (that is, non-Gibsonian) vision science, the visual system initially represents the surface features of an object it comes in contact with. Only by doing this can subsequent processing infer non-observed features of the object. This consideration suggests that T5-2 cells represent WORM-LIKE MOTION as opposed to WORM or FOOD.

Thus it is no surprise that Neander seeks to show that CT takes perceptual states to refer to the surface features of objects, taking as her model the prey-detector in the visual system of the frog. In Chapters 7 and 9, Neander argues that the frog's prey-detector represents SDMO (small, dark, moving objects) and not FLY. The argument relies on the claim that the prey-detector possesses the proper function of responding to SDMO and not FLY. This claim is defended by focusing on the alleged causal sensitivity of the representation producer (S). The idea is that we can only establish the proper function of S after first establishing what S is causally sensitive to. If S is causally sensitive to SDMOs but not flies, then S cannot possess the proper function of responding to flies. As a result, it can't be the case that R refers to the presence of flies as opposed to SDMOs.

Neander analyses the causal sensitivity of S using both Woodward's counterfactual account of causation (Woodward 2003) and the account offered by Lewis (1973). We will focus on Woodward's account. Neander argues that the following two statements are true:

- (a) If the visual target were changed from being SDMO to not being SDMO (but nothing else were changed), then the frog's visual system would change from producing Rs to not producing Rs;
- (b) If the visual target were changed from being a FLY (F) to not being F (but nothing else were changed), then the frog's visual system would produce Rs.

As a result, R refers to SDMOs and not flies. The point is well taken, but unfortunately it suggests a further set of counterfactuals rendering content more proximal than either FLY or SDMO (both of which are located in the same position along the causal chain from distal object to representation). Consider the particular proximal pattern of light (PPL) reflected by a FLY *qua* SDMO. Both of the following sentences seem true:

- (a) If the visual target were changed from emitting PPL to not emitting PPL (but nothing else were changed), then the frog's visual system would change from producing Rs to not producing Rs;
- (b) If the visual target were changed from being SDMO to not being SDMO (but nothing else were changed), then the frog's visual system would still produce Rs.

As a result, R now seems to be causally sensitive to PPL, as opposed to SDMO. And on Neander's logic, R would now represent PPL instead of SDMO. In response to this general kind of problem, the author proposes the *distality principle*. This principle is intended to push content back out the causal chain from proximal stimulation to distal object:

R refers to C [that is, SDMO] rather than the more proximal Q [that is, PPL] if the system responsible for producing Rs was adapted for responding to Qs...by producing Rs as a means of responding to Cs...by producing Rs, but it was not adapted for responding to Cs as a means to responding to Qs. (p. 222).

The idea is that the frog's prey-detector refers to SDMO rather than PPL because of the following fact: the sensory-perceptual system (S) that produces prey-detector activation was adapted for responding to PPL by activating the prey-detector as a means of responding to SDMO, but S was *not* adapted for responding to SDMO by activating the prey-detector as a means of responding to PPL. This might be true, but the problem is that the distality principle seems to push content past SDMO and all the way out towards FLY again. After all, S was adapted for responding to SDMO by activating the prey-detector as a means of responding to flies, but S was not adapted for responding to flies by activating the prey-detector as a means of responding to SDMO. Neander's proposed solutions to two potential indeterminacy problems may be at odds with one another.

4. The question of representation status

As Neander rightly notes, most work in the naturalisation project has relied on pre-theoretic assumptions about what structures in our minds are representations, and what they represent. Perhaps that is a reasonable starting point when the focus is on the folk psychological states of belief or preference. But when the focus is on subpersonal cognitive mechanism it is less obvious when states are genuinely representational. To her credit, Neander does explicitly discuss this question of representational status, and what she says deserves some discussion.

Why, for example, are some neural cells located in the optic tectum of a normal toad a mental *representation*, as opposed to just another causal relay in a complex biological mechanism? One proposal that Neander criticises is the sender-receiver model. On this proposal, a representation is an intermediary between two systems, or two components in a larger system: a 'sender' and a 'receiver'. In the paradigm model, the sender and receiver are co-adapted. The sender can 'see' the external world, but the receiver cannot. The receiver can 'act' on the external world, but the sender cannot. Moreover, the actions of each player have consequences for the pay-off for the other. Hence, there is a coordination problem. Functionally speaking, it is the job of the sender to coordinate the receiver's response with the state of the world via the production of a 'signal'. A representation just is this intermediary signal produced by the sender, which elicits a response in a receiver.

The sender has the proper function to help the receiver perform its proper function at the right time, via the sending of a signal which correlates with or maps onto the state of the world in the right way.

This potential answer to the representational status question is problematic because it would render countless causal intermediaries in countless biological mechanisms representational. Neander's apt example is the protease sent from the pancreas to the gut. Here, the pancreas (sender) and the gut (receiver) are co-adapted components in the digestive system. Protease is a sign that causes the gut to begin protein digestion. The pancreas helps the gut digest protein at the right time by producing the sign in the presence of protein. Moreover, the amount of protease will normally also vary with the amount of protein entering the digestive tract. In other words, in addition to meeting the basic requirements of a sign/representation in a sender-receiver system, protease also possesses significant causal specificity. The system is more fine-grained than just 'on' or 'off'.

One possible response is to introduce an additional element into the sender/receiver framework; perhaps, as Skyrms has suggested, the signal must be in some sense arbitrary (Skyrms 2010). Neadner tentatively suggests an alternative proposal. She suggests that a representation is an entity in a representing system. What is a representing system? It is a system that has the proper function of 'mirroring' the world, enabling adaptive responses to the environment on the part of the individual possessing the system. What is 'mirroring', in this context? It is the obtaining of a second-order similarity relation between a set of entities in a representational system and a set of entities in the world, that is, where certain relations of similarity hold over the entities in the representing system that mirror certain relations of similarity in the world. When this is the case, any particular representation that might be tokened will have 'an analogous place, in the similarity space of the system's representations, as its content has, in the similarity space of what the system's representations can represent' (pp. 181-2).

An example of such a system is a ruler. Here, certain relations hold over the entities in the representing system (cardinal numbers printed on a ruler representing centimetres) that 'mirror' relations holding over the entities in the world (object height in centimetres). The number 2 on a ruler is an analogue of a relatively small object height, while the number 8 on a ruler is an analogue of a larger object height. This is because the relation between, say, the numerals '2' and '8' on a ruler mirrors the relation between smaller and larger object heights in the world. One worry about this proposal is that protease would count just as much as representational as it does on the sender-receiver account. After all, the range of possible amounts of protease produced by the pancreas can easily be said to 'mirror' the range of possible amounts of protein in the digestive system. Moreover, this 'mirroring' enables adaptive responses to food by the individual possessing a digestive system: Neander writes that the protease sent to the gut from the pancreas

can be said to have the function to model the amount of protein entering the digestive system. She holds, however, that the representational status of protease is 'minimal, and so we need not take this modelling system very seriously as a case of a representational system as such, on my proposal. It is, as it were, barely on the spectrum' (p. 277). However, it is not clear why on her analysis this toy case is, indeed, minimal. After all, were we to compare the protease release system, which possesses significant causal specificity, with a more paradigmatically 'representational' neural state in the cognitive sciences that happens to be just 'on/off', it would appear that on Neander's suggestion the neural state would count as less of a 'modelling system' than the protease release system. We should emphasise that Neander's proposal is explicitly tentative. Perhaps the right conclusion is that the status issue needs *much* offered as more serious attention in this literature.

5. From perception to belief

It is clear that Neander intends her account of the representational character of perceptual states to contribute to the project of explaining mental representation in general, including fully conceptual content. It is less clear to us exactly how she conceives of this contribution. We doubt that she thinks that the content of beliefs and desires can be derived by some definition-like specification from that of perceptual states. That would commit her to a foundationalist-empiricist picture of content, and no project seeking to construct a fully worked out empiricist picture of content of this kind has met with much success since the time of Locke. Nor does she seem to have in mind some direct reduction, in the same general class as those reductions offered by Millikan or Fodor. Unlike Millikan and Fodor, Neander suggests (in Chapter 1) that most concepts probably don't have original intentionality (as opposed to derived intentionality). This, we take it, partly explains the focus of the book on non-conceptual content. If we can explain non-conceptual content, then this (Neander hopes) might allow us to explain conceptual content, which might then allow us to explain the content of person-level, propositional attitude, thought. So the picture is fundamentally bottom-up: start with content at its lowest level and build up from there to belief and desire.

We are not optimistic about a bottom-up project, for it seems to us that human person-level cognition differs from (say) that of the rat in the rat's use of cognitive maps not just in the degree of our adaptive plasticity and complexity. When the rat acts successfully in the world, reliable success depends on adaptive mechanisms built in the mind of the rat by individual natural selection: it depends on the rat's cognitive system working as designed. Human success depends on those mechanisms. But it also depends on another class of mechanisms: languages; mathematical and other notation systems; information stored using public representational systems; artefacts; and community lore, in the case of pre-literate cultures. No alphabet or

notation system is built by any individual, yet these cranes (to borrow a term from Dennett) enhance and transform our cognitive powers. Our success depends on the proper functioning of our cognitive system working as designed too, but it depends as well on these collectively produced and inherited cognitive resources, and these do not have a proper function in Neander's sense. So we do not think the human case can be naturalised just by scaling up from a toad or a rat. We are inclined towards Dennett's indirect strategy. Develop a natural and social history of the origins and expansion of our practices of interpreting one another using folk psychology, while relying on the cognitive sciences to explain why those practices succeeded often enough to be stable. Successfully carrying through the indirect strategy would take the mystery out of intentionality, but without guaranteeing anything resembling a type-type reduction of intentional properties to those of the cognitive sciences.

Adopting the indirect strategy might take some pressure off the project of naturalising representation at the subpersonal level. If we aren't seeking a type-type reduction of the intentional properties of folk psychological states to properties of the cognitive sciences, then we needn't worry about not finding an image of certain notoriously demanding properties of folk psychological states at the subpersonal level. For instance, non-conceptual sensory-perceptual representations needn't be free of content indeterminacy (assuming our beliefs and desires are). Moreover, we needn't worry if, say, non-conceptual representations in early visual processing don't give rise to intensional contexts. (Neander does worry, and seeks to show that they do.) Finally, it wouldn't be a disaster if the representational normativity of non-conceptual visual representations is of a 'thinner' kind than the normativity of folk psychological representations and other paradigm intentional states like sentences: the latter depending on socio-cultural norms as opposed solely to the norms of proper biological functioning.

This approach may undercut those who dismiss representational talk in the subpersonal cognitive sciences (for example, Hutto & Myin (2013)). On their view, a cognitive state is genuinely representational only if it gives rise to intensional contexts and 'thick' normativity. But if we do not expect to find in the cognitive sciences a direct image of the intentional properties of folk psychological states, then we would not expect subpersonal representational states to have all the properties of folk psychological states. Neander's book shows that they probably have some of those properties: aboutness and some degree of normativity. However, they might not have fully determinate content. For example, specifying the selective history of a mechanism always involves idealisation, because some deaths are due to unlucky accidents rather than design flaws, or events in highly atypical environments. Are events in vision labs to be included in the selective history of macaque cognition? There is no unique, privileged idealisation. Moreover, this perspective allows that the extent to which subpersonal and personal states share these

relevant properties is a matter of degree. States of the T5-2 cells in toads, or representations of lines in the human visual system, may well be directed at something in the external world. But these may well be normatively assessable only by a rough notion of accuracy or appropriateness rather than (for example) truth.

We are not suggesting a fictional view of representation in the cognitive sciences. Neander has developed a persuasive case that some subpersonal cognitive states really are representational: they have aboutness (backed by causal-informational relations) and normativity (backed by biological proper function). These two properties support the view that certain subpersonal states are representational, not because they help ground the intentionality of folk psychological states, but rather because both aboutness and normativity (along with other properties) are familiar features of paradigm intentional states like beliefs and desires, and also public signs. The suggestion is that we take the intentionality of folk psychological states, and possibly also public signs, as a *model* for understanding cognition at lower levels in the explanatory hierarchy. By doing so, we suggest that an interesting path opens up between fictionalism regarding representation in the cognitive sciences, on the one hand, and reductionism, on the other. Whether or not we are right about that, this is a fine and thoughtful contribution to the overall project of naturalising the mind.

References

- Dennett, D. C. (1991). 'Real Patterns'. *Journal of Philosophy* 87: 27–51.
- Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT/Bradford Press.
- Hutto, D. and E. Myin, (2013). *Radicalizing Enactivism: Basic Minds Without Content*. Cambridge, MA: MIT Press.
- Levins, R. (1966). 'The Strategy of Model Building in Population Ecology'. *American Scientist* 54(4): 421–431.
- Lewis, D. K. (1973). 'Causation'. *Journal of Philosophy* 70(17): 556–567.
- Skyrms, B. (2010). *Signals*. Oxford: Oxford University Press.
- Stampe, D. (1977). 'Toward a Causal Theory of Linguistic Representation', in P. French, H. K. Wettstein, and T. E. Uehling (eds.), *Midwest Studies in Philosophy*, Vol. 2, Minneapolis: University of Minnesota Press, pp. 42–63.
- Sterelny, K. (2006). 'Local Ecological Communities'. *Philosophy of Science* 73(2): 215–231.

Wimsatt, W. C. (2007). *Re-Engineering Philosophy For Limited Beings: Piecewise Approximations to Reality*. Cambridge, Mass, Harvard University Press.

Woodward, J. (2003). *Making Things Happen: a Theory of Causal Explanation*. Oxford, Oxford University Press.

School of Philosophy,
Australian National University
david.kalkman@anu.edu.au;
kim.sterelny@anu.edu.au

DAVID KALKMAN AND KIM STERELNY

doi:10.1093/mind/fzy036

Advance Access publication 5 September 2018

Virtue and Law in Plato and Beyond, by Julia Annas. Oxford: Oxford University Press, 2017. pp. 234.

Julia Annas's *Virtue and Law in Plato and Beyond* will change the way Plato's political philosophy is understood. Annas sets straight both the historical naïf who supposes that virtue and law are alternatives, and the sophisticate who has learned that whereas Plato advocates rule by philosophers in the *Republic*, he more realistically settles for rule by laws in the *Laws*. Annas briskly demonstrates that the *Republic* says both that a city has to have laws, and that the point of law is to cultivate citizen virtue (rather than merely to keep citizens from harming and being harmed by each other), because the law of a city should aim at the happiness of its citizens and happiness depends on virtue. Even though 'the only society that is any good, the only one worth thinking about, is the society where people live happy lives because their lives are lived virtuously' (p. 31), the *Republic* is replete with talk of law: Socrates and his interlocutors are lawgivers to the ideal city (p. 13) (a role which includes legislating the noble lie, for *pace* Annas's claim that it's the rulers who establish the noble lie (p. 19), Socrates says 'I'll first persuade the rulers' (414d)); citizens are educated to be law-abiding (pp. 17-18); individual and civic degeneration is due to lawlessness (pp. 15-16). (One might note that the centrality of legislation to Plato's political thought goes back further than the *Republic*: in the *Gorgias*, Socrates says that the two arts that care for the soul and comprise the political art are legislation and (corrective) justice (464b-c), and in the *Crito* the personified laws of Athens remark on Socrates' approval of Sparta and Crete for being lawful (53a).) Given the presence of law in both the *Republic* and *Laws*, Annas argues that the real difference between the works is that the *Laws*: (1) describes the best form of government by drawing on the best of existing constitutions rather than by starting from scratch, thereby putting itself into dialogue with citizens of cities governed by those constitutions; and (2) discusses in detail how law effects citizens' virtue.