Why I am not a Literalist

Zoe Drayson, January 2020
Forthcoming in *Mind and Language*: http://doi.org/10.1111/mila.12306

1. Semantics and the mind-sciences

In the 1980s, when Fodor's Granny was being appalled by relativism in Harvard and graffiti in subway stations, philosophers were appealing to cognitive science to settle questions about the semantics of psychological state ascriptions.[1] The truth-aptness of our folk psychological discourse, such as our ascriptions of beliefs and desires, was widely held to rest on the scientific posits of computational psychology.

The semantics and the science of the mind have largely gone their separate ways over the intervening years. Some philosophers have drawn conclusions about mental states from semantic evidence while remaining neutral with respect to cognitive science,[2] and other philosophers have used scientific evidence to justify claims about the nature of our minds while saying nothing about psychological state ascriptions.[3]

Carrie Figdor's (2018) book *Pieces of Mind*: *The Proper Domain of Psychological Predicates* is an attempt to reconcile the two approaches to the mind in a position she calls Literalism, which argues for a semantic claim about psychological predicates on the basis of a scientific claim about the nature of psychological properties. Her semantic claim focuses on recent scientific work (rather than everyday folk-psychological discourse) which ascribes psychological terms to

[1] Granny featured in Fodor's writings as the voice of conservatism and the defender of both folk psychology and the role of semantics in determining metaphysics. The reference to subway graffiti and Harvard relativism comes from Fodor (1984).
[2] See e.g. Stanley and Williamson (2001), who appeal to the semantics and syntax of embedded questions to argue that know-how is a propositional mental state without making scientific claims about the realizers of such states.
[3] See e.g. Clark (2013) and Hohwy (2013), who appeal to neural models of predictive coding to argue that our mental processes are hierarchical Bayesian computations, but offer no account of what makes our ascriptions of any particular psychological states ascription true or false.

non-humans, such as cases where scientists describe plants as intelligent, bacteria as goal-directed, or neurons as having preferences. Figdor argues that we should understand such ascriptions literally, rather than figuratively. Her main argument for this semantic claim is the scientific claim that non-humans have been discovered to possess the psychological properties in question, and that these are the same psychological properties possessed by humans. Figdor summarizes Literalism as follows:

> "Literalism holds that psychological predicates are being used with literal intent to pick out the same scientifically-discovered structures across the relevant human and non-human domains." (61)[4]

Figor argues that we previously understood our psychological concepts in a human-centred way, and thus that the truth of Literalism would demonstrate that our psychological concepts are undergoing radical conceptual change.

I will explore Figdor's scientific claim first, before addressing its role in supporting her semantic claim. In Section 2, I will evaluate Figdor's claim that quantitative scientific modeling provides strong evidence that non-humans possess the same kinds of psychological properties as humans. Figdor draws on two quantitative models to support her case: her first case is a model of decision-making which has been applied both to humans and to fruit-flies; her second case is a model of preferences which has been applied both to humans and to individual neurons. (I'll focus on the decision-making case, but my arguments are equally relevant to the preference case.) I will argue that Figdor neglects several important features of scientific modelling, which considerably weakens her scientific claim that humans and non-humans have the same psychological properties. In Section 3, I'll propose that even if we accept her scientific claim, it does not provide the appropriate support for her semantic claim. As a result, we lack the evidence required to establish Figdor's claim that our psychological concepts are undergoing radical conceptual change. In Section 4, I will conclude by raising some broader questions about attempts to combine semantics and the mind-sciences.

[4] All page references refer to Figdor (2018) unless other noted.

2. F igdor's scientific claim about psychological properties

2.1. The drift-diffusion model of decision-making

Figdor's scientific claim is that psychological properties are not restricted to humans: she argues, for example, that fruit-flies make decisions. She draws on the fact that scientists have successfully applied the Drift-Diffusion Model (DDM) of decision-making to both humans and fruit-flies, and takes this to provide strong evidence that fruit-flies have the same decision-making properties as humans. In this section I will introduce the DDM (2.1), then highlight some important features of scientific models which Figdor neglects to consider (2.2). I will argue that if we take these features seriously, they diminish the evidence that the DDM is supposed to provide for non-human psychological properties (2.3). I conclude by considering and rejecting a counterargument that Figdor could raise (2.4).

In the simplest decision-making tasks studied by psychologists, subjects are given a forced choice between just two options. One such experiment involves a perceptual discrimination task, where the subject is presented with two stimuli (e.g. images) and asked to decided which one contains a certain target (e.g. motion, or a face). The subject's decision-making process involves accumulating evidence until they make their discriminative choice. As one would expect, there is a trade-off between response time and accuracy in such tasks: subjects asked to make the decision quickly perform less accurately, and subjects asked to focus on accuracy will take longer to make their choice. When experimenters introduce noise into the stimuli, by blurring or pixelating the images, the task becomes more difficult: subjects' response time increases and their accuracy decreases.

There are several ways to model these two-alternative forced-choice tasks, one of which is the DDM. The DDM is a quantitative approach which models the process of evidence accumulation as a 'random walk' through mathematical space, acquiring and integrating evidence for the two alternatives until a decision threshold is reached. The more difficult the task, the slower the rate of evidence accumulation, or 'drift'. The DDM models the formal relation between the drift rate and the rates of accurate and inaccurate response.

The DDM was developed as a model of human decision-making, but has since been successfully used to predict the decision-making behavior of monkeys and pigeons. More recently, DasGupta and colleagues (2014) have applied the DDM to model fruit-flies. They first presented the fruit-flies with a two-choice odor-discrimination task, and trained them to choose the more concentrated of the two samples. In subsequent trials, experimenters recorded the flies' behavior when presented with different combinations of odor concentration. The DDM predicts that in more difficult tasks, where the two odor concentrations are more similar, the fruit flies' drift rate and accuracy rate would be lower. This is exactly what the experimenters found. A group of flies with the FoxP mutant gene, which is associated in humans with motor and speech disorder, took more time to give similarly accurate responses than flies in the control group.

Figdor takes this data to show that the DDM successfully models decision-making in fruit-flies, just as it does in humans. She proposes that the successful uses of such models across human and non-human domains provides "quantitative evidence of structural commonality for specific capacities" (61): in this case, evidence that humans and non-humans share the same psychological capacity for decision-making. Figdor's reasoning, however, relies on a number of assumptions about scientific models and their interpretation. In particular, it assumes that a certain representational relationship holds between the model and the world, which licenses us to draw conclusions about properties of the world from properties of the model. While Figdor acknowledges that there are philosophical questions concerning the representational and epistemic role of scientific models, she proposes that these questions can be set aside for the purposes of her argument (32, n.4). In what follows, I will argue that answers to these questions have crucial implications for Figdor's argument.

2.2. Two features of scientific models

2.2.1. The epistemic constraint on scientific models

Scientific models are epistemic tools that allow us to learn about the world. Figdor clearly thinks that we can learn about an aspect of the world (the psychological properties of fruit-flies, for example) from applying a scientific model (such as the DDM) to the world. In order to play this role, a model must be a *representation* of the world – but mere representation is not sufficient to justify us in drawing conclusions about the world from the model. We might establish a representational relation by stipulation, for example, but this would not allow us to learn substantial information about the world from the model. To play this epistemic role, models must represent the world "in a way that allows us to form hypotheses about their target systems" (Frigg and Nguyen, 51). Various representational relations have been proposed to meet this epistemic constraint, involving features such as similarity, isomorphism, or inferential role. If we want to use the DDM to draw conclusions about the psychological properties of humans or non-humans, the appropriate epistemic relation must exist between the model and the world. Given these considerations, it is not clear why Figdor thinks she can set aside questions about how models represent the world. I will come back to this matter shortly.

2.2.2. Abstraction and idealization in scientific models

Where the appropriate epistemic relation holds between a model and the world, the model can act as a useful surrogate for the world [Swoyer]: we can manipulate features of the model where it would be impossible or implausible to manipulate the corresponding features of the world. The model is easier to manipulate that the world because models are abstractions from the world, in the sense that they leave out some of the messy details of real-world phenomena. Different kinds of scientific models operate at different levels of abstraction. *Qualitative* models represent causal features of the world using causal features of the model: we could model the spatial relations between the molecules of a liquid with the spatial relations between plastic balls, for example, while abstracting away from other features of the liquid. *Quantitative* models, on the other hand, abstract away from causal relations completely by using the formal relations between variables in a mathematical equation to represent the formal relations between aspects of the world. A quantitative model of a liquid could model the formal relation between its density and its flow-rate while abstracting away from the spatial relations between the liquid's molecules. Psychological models can be either qualitative or quantitative: the process of accumulating evidence in decision-making, for example, can be modeled as a complex causal mechanism in the brain, or as a single variable such as drift-rate.

While all models ignore certain details of the world, some models represent the world in ways which are at odds with how the world actually is. Some quantitative models of liquids, for example, don't just neglect their discrete molecular structure but rather represent liquids as having a non-discrete structure: they build in the false assumption that density and velocity vary continuously at every point. These idealized models are useful in allowing us to model the large-scale behavior of a system more simply.

More abstract or idealized models can often be applied outside their original domains. Quantitative models of fluid dynamics, for example, can be used to predict the flow of automobile traffic, because the variation between flow-rate and density in traffic exhibits the same formal relation as in liquids. While this highlights an interesting relation between liquids and traffic at a certain level of abstraction, the successful extension of a model to a different domain does not justify us in inferring that the new domain has similar ontological properties to the original domain. The application of models from fluid dynamics to traffic does not allow us to conclude that traffic is actually liquid. So why should the application of the DDM to fruit-flies permit Figdor to conclude that fruit-flies are decision-makers?

2.3. Challenging Figdor's scientific claim

Having elucidated these features of scientific models, I will now return to Figdor's claims about decision-making in humans and non-humans on the basis of the DDM. Figdor proposes that the DDM represents decision-making properties when applied to humans, and that its application to certain non-human species is strong evidence that those species possess the same kind of decision-making properties. Her argument relies on establishing that the DDM can play the appropriate evidential role in the first place, and on demonstrating that the evidence it provides is strong enough to outweigh any counterevidence. I have concerns about both of these steps.

First, we can ask whether the DDM plays the same representational role when applied to humans and to non-humans. In virtue of what does the DDM represent decision-making in humans? Figdor does not discuss this matter, and gives no indication of whether she takes this representational relation to play an epistemic role. Notice that it would seem strange to say that the DDM provides us with evidence that humans are decision-makers: the DDM was developed as a quantitative model of decision-making in humans on the pre-existing assumption that humans are decision-makers. We might still learn about features of human decision-making from modelling it with the DDM, but I suggest that we do not learn from applying the DDM to humans that humans have the psychological property of being decision-makers. By way of comparison, consider the representational relation between models in fluid dynamics and the properties of liquids they represent: we might use the models to learn about certain properties of liquids, but these models do not provide evidence that liquids are fluid. Rather, it was our pre-existing assumption that liquids are fluid which enabled us to develop models of fluid dynamics.

Now consider the relationship between the DDM and fruit-flies. If we accept that the behavior of fruit-flies in forced-choice tasks can be modelled by the DDM, this indicates that the model represents properties of fruit-flies. But this doesn't tell us that the epistemic constraint can be met: it doesn't tell us that we can learn about fruit-flies from the model, and *a fortiori* it doesn't tell us *what* we can learn about fruit-flies from the model. In particular, it doesn't tell us that the DDM provides evidence that flies are decision-makers. Compare the application of fluid-dynamic models to traffic: we can accept that the models represent properties of traffic (its rate of flow, for example) and we might learn that traffic density covaries with traffic flow, without concluding that traffic has other liquid-like properties. We can accept that the application of quantitative model from one domain to another "provides strong evidence that two domains have important similarities" (135) while denying that all of the properties of the original domain should be attributed to the new domain. So if Figdor want to use the DDM as evidence that fruit-flies have the same decision-making properties as humans, she needs to say more about the representational relationship between the DDM and both humans and non-humans.

Even if we were to accept that the application of the DDM to fruit-flies provides some evidence that fruit-flies make decisions, we still face my second concern: why think this evidence is sufficient to outweigh any evidence we have to contrary? After all, the DDM is one model among a number of qualitative and quantitative models that we use to represent decision-making. We have information-processing models, for example, which represent the interactions between the cognitive components of decision making: formulating the decision problem, identifying the decision goal, compiling a set of available options from long term memory, opting for an appropriate strategy or decision rule, and using it to evaluate the options to make a selection (Wang and Rube 2007). We also have neural models of decision-making which represent how these information-processes might be implemented in the brain, including models which represent activity in the lateral intraparietal cortex as correlating with choice-making (Gold and Shadlen 2007). These information-processing and neural models of decision-making have not been successfully applied to creatures such as fruit-flies, because they lack the relevant neural structures or reasoning skills. Don't these considerations count as evidence against the claim that fruit-flies are decision-makers?

By way of comparison, consider the fact that models from fluid dynamics can be applied to traffic. If we were tempted to take this as evidence that traffic is liquid, we would soon reconsider in the fact of counterevidence from other sources, including qualitative models of traffic. Figdor appears to acknowledge the role of counterevidence in the case of decision-making models: she allows that if could apply the DDM to an inanimate object, such as a body of water, we would probably not conclude that the object was a decision-maker. In such a case, she thinks that the model would still provide "relevant scientific evidence that the body of water makes decisions" but that we would have "legitimate scientific reasons" for denying this conclusion (86). I propose that Figdor's reasoning here should apply to her own case. We can accept that DDM provides us with some scientific evidence that flies make decisions, while thinking that we have legitimate scientific reasons not to endorse these interpretations.

To summarize: I've questioned two key ideas that Figdor's scientific claim seems to rely upon. I've argued that Figdor needs to do more to show that the application of quantitative models to non-humans provides evidence of their decision-making properties, and that this evidence is not outweighed by counterevidence.

2.4. Figdor's defense of the scientific claim

The first of my problems concerns the nature of the representational relationship between the DDM and fruit-flies. Figdor seems to think that DasGupta et al.'s (2014) application of the DDM to fruit-flies demonstrates that an appropriate representational relation holds. She thinks that when a quantitative model is extended from one domain to another, the model will represent the same properties in both domains:

> "There are clear reasons to think the construals of the same mathematical models used across domains are univocal, especially when taking into account the explicit intent on the part of modelers." (61)

It is not clear why we should accept this claim, and it is worth noting that the explicit intent of the modeller is not generally considered sufficient to fix the representational relation between a model and the world (Frigg and Nguyen 2017, 51). The cases where this is most likely to work are cases where the representational relation is stipulated, and thus unlikely to fulfil the epistemic constraint discussed above.

I want to focus, however, on my second criticism of Figdor's argument, that she fails to take into account our counterevidence to the claim that fruit-flies make decisions. Figdor acknowledges that talk of fruit-flies as decision-makers may seem counterintuitive, but argues that the application of a quantitative model from one domain to another provides evidence of important similarities between the two domains "whether or not intuition agrees" (135). The counterevidence to which I am appealing, however, does not come from intuition but from other scientific models of decision-making (such as information-processing models and neural models) which do not seem to apply to the case of fruit-flies. Figdor seems to think that such models are based on anthropocentric intuitions, however: they will only apply to creatures that are qualitatively similar to humans. She argues that quantitative models like the DDM provide much stronger evidence for psychological properties in non-humans "because they rely on a well-respected scientific method for discovering similarity that is independent of qualitative similarity to the human case" (55).

In what sense are qualitative models of decision-making anthropocentric? Information-processing models in cognitive psychology, in particular, are not even biocentric: they characterize psychological properties at a level of abstraction from the brain, in terms of computational states that can be implemented in different physical substrates. The foundations of artificial intelligence are grounded in the idea that a nonbiological system could have the same psychological properties as a human. So if Figdor thinks that such models are anthropocentric, she must have something else in mind. She seems to think that qualitative models are anthropocentric in the sense that they are taking particular features of human psychology as the measure of psychological properties more generally. She proposes that when these models cannot be applied to non-humans, it is because they are using “an ideal hyperrational, baroquely complicated human cognizer as the standard for real psychological capacities” (87). But if the claim is just that qualitative models of psychological properties were developed as models of human psychological properties, then the same allegation can be levelled at quantitative models. In the DDM, for example, the mathematical variable of drift in the DDM is an abstraction from the causal process of evidence accrual represented by information-processing models. If information-processing models of decision-making are anthropocentric in taking human psychology as their starting point, then quantitative models like the DDM seem to face the same charge.

Figdor clearly thinks that there is an important difference between qualitative and quantitative models, and that this difference allows quantitative models to avoid the anthropocentric bias of qualitative models. But such a difference could only reside in the formal nature of quantitative models: the fact that they model the mathematical relationship between numerical variables rather than the causal relationship between concrete properties. Quantitative models of psychological properties are arguably unbiased in the sense that they do not require us to think of psychological properties as human, but they also do not require us to think of psychological properties as physically implemented or causally efficacious. Notice that this makes quantitative models compatible with positions like dualism and epiphenomenalism, which sit uneasily with Figdor’s naturalism. At the very least, therefore, Figdor owes us an account of the sense in which qualitative models build in an anthropocentrism that quantitative models lack, and whether there’s a sense in which quantitative models are specifically non-anthropocentric rather than more generally non-concrete, non-causal, or non-physical.

I have argued that there is no easy way to understand how qualitative models are anthropocentric in a way that quantitative models avoid. As a result, it is not clear why we should accept the quantitative evidence that fruit-flies are decision-makers and ignore the qualitative counterevidence. This leaves Figdor’s scientific claim, that humans and non-humans have the same psychological properties, lacking justification.

3. Figdor's semantic claim about psychological predicates

3.1. Quantitative models and Literalism

Figdor's relies on her scientific claim, that psychological properties are possessed far beyond the human, to argue that we are justified in taking scientific attributions of psychological predicates to non-humans literally. Much of the time, her argument seems to be that the application of a quantitative psychological model to non-humans provides us with (defeasible) evidence that non-humans are decision-makers. Sometimes, however, Figdor suggests that quantitative models are playing a reference-fixing role rather than an evidential one. She suggests that quantitative models are providing our psychological predicates with "scientifically determined reference points" (124): in the case of decision-making, the reference class of the predicate 'decides' is determined by all and only those domains that can be modelled by the DDM. But this is a stipulative use of scientific modelling (akin to claiming that all and only those domains which can be modelled by fluid dynamics constitute the reference class of the term 'liquid') which is clearly distinct from an evidential use of scientific modelling. It is questionable whether the same scientific model can play both these roles.

Although I think Figdor could say much more here about the philosophical properties of scientific models, I want to allow her claim to have established that our psychological predicates like 'decides' can apply to non-humans such as fruit-flies. What I want to question is Figdor's further conclusion that our psychological concepts are changing as a result. She claims that we have not merely discovered that the world includes more decision-makers than we previously thought, but that our concept of decision-making has changed its meaning.

3.2. Conceptual change

Figdor's Literalism claims that if humans and nonhumans have the same psychological properties, then our psychological predicates apply literally to both. She also proposes that our psychological predicates are undergoing radical conceptual change, on the grounds that our previous psychological concepts were human-centered while our new psychological concepts are not human-centred:

> "Scientifically driven referential revision can also result in a newly precise standard for a kind. […] Human-centered concepts and capacities are adjusted in the light of new evidence to make them not human-centered." (144)

The truth of Literalism alone does not entail such claims about conceptual change. It is compatible with Literalism to simply claim that our reference-class for the psychological predicate in question has expanded: our predicates pick our properties which are more abundantly instantiated than we previously thought.

Figdor proposes that we are undergoing major conceptual revision because our previous psychological concepts were human-specific, whereas our new psychological concepts apply equally to non-humans. She thinks that our previous psychological concepts were human-specific because she thinks that traditional approaches to the mind are 'Exceptionalist': they take the view that "human capacities are metaphysically exceptional" (166). Figdor thinks that her scientific claim forces us to reject an Exceptionalist metaphysics in favour of an Anti-Exceptionalist metaphysics, according to which "psychological capacities are possessed by a far wider range of kinds of entities than often assumed" (5). This alteration in metaphysical views is responsible for the conceptual change, according to Figdor.

I can find little evidence to support the idea that our traditional metaphysics of mind takes human psychological properties to be exceptional. Functionalism, for example, is notorious for allowing that octopodes, Martians, even entire nations can be in the same psychological states as humans. Behaviorism identifies psychological states with behaviors or dispositions to behave, but with no restriction to human behavior. Even type-identity theory doesn't restrict psychological states to humans: pain-in-humans may be a different physical state type from pain-in-dolphins, but there is no claim that the human case is special or otherwise exceptional.

I would like to suggest that Figdor's allegations of Exceptionalism are best understood as allegations of metaphysical chauvinism, as discussed by Block (1978). On Block's account, whether a metaphysical theory like functionalism is chauvinist is a matter of whether it withholds psychological properties from creatures or systems that do in fact have those psychological properties. This depends on how we specify the particular type of behavior, functional role, or brain state that is identical to each psychological state type. If we made it a criterion for having a belief, for example, that one can articulate the content of one's belief, we might end up inaccurately withholding belief states from prelinguistic infants and nonlinguistic creatures. Similarly, if fruit flies are decision makers, but we deny this because they don't meet our pre-existing metaphysical characterization of decision-making, then we are guilty of chauvinism. This seems to be the same mistake that Figdor refers to as Exceptionalism.

Why think that traditional approaches to the mind adopt an Exceptionalist or chauvinist metaphysical view? This claim of Figdor's seems to go unsupported, and without a reason to think that our psychological concepts were previously human-centred, we lack the conditions for conceptual change. Figdor also fails to acknowledge Block's important point that if we reject metaphysical chauvinism, we run the risk of adopting metaphysical liberalism: ascribing psychological properties to creatures or systems which do not in fact have them.

4. Conclusion

Figdor's attempt to combine the semantics of psychological predicates with scientific claims about psychological properties is an ambitious one. I began this paper by comparing it to the 1980s debate between realists and eliminativists over the status of folk psychology, and it is worth considering why that particular debate is no longer an active one. That debate relied on the assumption that there are direct connections between claims about the nature of psychological properties and the truth or falsehood of our psychological predicate ascriptions. As De Brigard (2015) points out, that assumption is underwritten by a particular version of scientific realism: once we acknowledge the alternative approaches we can take to the issue, there are no easy moves from claims about properties to claims about predicates and vice-versa. The moves that are available to us are largely constrained by the sorts of things we want to say about scientific models and their interpretation – precisely the kinds of issues that Figdor thinks can be set aside for the purposes of her project. I hope to have demonstrated here that the semantic and epistemic features of scientific modelling are essential to understanding the relationship between psychological predicates and psychological properties.

References

Block, Ned (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science* 9:261-325.

Clark, Andy (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36 (3):181-204.

DasGupta S, Ferreira CH, Miesenböck G (2014) FoxP influences the speed and accuracy of a perceptual decision in Drosophila. In Science 23;344(6186):901-4.

De Brigard, Felipe (2015) What Was I Thinking? Dennett's Content and Consciousness and the Reality of Propositional Attitudes. In C. Muñoz-Suárez, F. De Brigard (eds.), Content and Consciousness Revisited, Studies in Brain and Mind 7 (Springer)

Figdor, Carrie (2018). Pieces of Mind: The Proper Domain of Psychological Predicates. Oxford, UK: Oxford University Press.

Frigg R., Nguyen J. (2017) Models and Representation. In: Magnani L., Bertolotti T. (eds) Springer Handbook of Model-Based Science. Springer Handbooks. Springer, Cham

Gold, I and Shadlen, M (2007) The Neural Basis of Decision Making. Annu. Rev. Neurosci. 2007. 30:535–74

Stanley, Jason & Williamson, Timothy (2001). Knowing How. *Journal of Philosophy* 98 (8):411-444.

Hohwy, Jakob (2013). The Predictive Mind. Oxford University Press UK.

Wang, Y and Ruhe, G (2007) The cognitive process of decision making. Int'l Journal of Cognitive Informatics and Natural Intelligence, 1(2), 73-85.