

Emotions, concepts and the indeterminacy of natural kinds

Taylor, John

DOI:

[10.1007/s11229-018-1783-y](https://doi.org/10.1007/s11229-018-1783-y)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Taylor, J 2018, 'Emotions, concepts and the indeterminacy of natural kinds', *Synthese*.
<https://doi.org/10.1007/s11229-018-1783-y>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Emotions, concepts and the indeterminacy of natural kinds

Henry Taylor¹ 

Received: 19 September 2017 / Accepted: 5 April 2018
© The Author(s) 2018

Abstract A central question for philosophical psychology is which mental faculties form natural kinds. There is hot debate over the kind status of faculties as diverse as consciousness, seeing, concepts, emotions, constancy and the senses. In this paper, I take emotions and concepts as my main focus, and argue that questions over the kind status of these faculties are complicated by the undeservedly overlooked fact that natural kinds are indeterminate in certain ways. I will show that indeterminacy issues have led to an impasse in the debates over emotions and concepts. I first consider and reject one way of resolving this impasse. I then suggest a different method, which places more emphasis on a close analysis of predictive and explanatory practices in psychology. I argue that when we apply this method, a new position emerges: that it is indeterminate whether concepts or emotions are natural kinds. They are neither determinately natural kinds, nor determinately not natural kinds. Along the way, we will see that natural kinds have been put to two completely different theoretical uses, which are often been blurred together, and that they are ill-suited to fulfil one of them.

Keywords Natural kinds · Psychology · Emotions · Concepts · Indeterminacy · Homeostatic property clusters

1 Introduction

One of the most pressing questions in cognitive science is which mental faculties (if any) are natural kinds. There is extensive debate over the kind status of perception (Burge 2010), consciousness (Irvine 2013; Shea 2012), seeing (Block 2012), colour

✉ Henry Taylor
j.h.taylor.1@bham.ac.uk

¹ Department of Philosophy, University of Birmingham, ERI Building, Edgbaston, Birmingham B15 2TT, UK

constancy (Davies 2016), the senses (Nudds 2011), attention (Prinz 2012), emotions (Griffiths 1997; Prinz 2002) and concepts (Machery 2009). This list could be extended.

Why are these questions so important? By investigating which faculties are natural kinds, we aim to discover which categories pick out genuine scientifically interesting divisions in nature, worthy of investigation. This is opposed to the categories that pick out groups of entities that have no scientific interest or import. Identifying natural kinds and disregarding categories that are not natural kinds is a key source of progress in science as a whole.

Many have suggested that certain mental faculties be dropped from scientific study, because they are not natural kinds. Paul Griffiths says:

It is unlikely that all the psychological states and processes that fall under the vernacular category of emotion are sufficiently similar to one another to allow a unified scientific psychology of the emotions... emotions are not a natural kind (2004a, pp. 901–902)

Griffiths concludes that the term ‘emotion’ should be:

eliminated from our psychological vocabulary (1997, p. 15).

Edouard Machery says something similar about concepts:

A growing amount of evidence suggests that concepts do not constitute a natural kind (2005, p. 445).

Machery then claims that:

“concept” ought to be eliminated from psychology (2009, p. 246).¹

I will start by examining the account of natural kinds that these debates rely on, which is the homeostatic property cluster (HPC) view. By examining *indeterminacy* in the HPC view, I will show how the debates over the kind status of emotions and concepts have reached an impasse (Sects. 2–4). I then investigate ways that we might decide whether emotions and concepts are kinds, and thus how we might resolve this impasse. I draw a distinction between two quite different theoretical roles that the HPC view has been put to, which I call the taxonomic and metaphysical roles. I argue that the HPC view is ill-suited to fulfil the taxonomic role, because of issues pertaining to indeterminacy (Sect. 5). I thus dismiss one method for resolving the question of whether emotions and concepts are kinds. Finally, I suggest another method, which is much more closely focussed on an examination of the predictive and explanatory practices of psychology. I argue that when we deploy this method in the case of emotions and concepts, the result we get is that it is *indeterminate* whether they are kinds or non-kinds (Sect. 6). I close by drawing out some consequences of these arguments (Sect. 7). For reasons of space, I will mainly focus on emotions and concepts, but many of the points apply generally. So, the paper should be seen as making general observations about psychological kindhood, with emotions and concepts as concrete case studies.

¹ See also Griffiths (2004b, 2008, p. 216) and Machery (2010a, b).

2 Homeostatic property clusters

Ultimately, one of the main purposes of this paper is to argue for the claim that it is indeterminate whether emotions and concepts are natural kinds. However, before we can do this, we must settle certain issues concerning *how* best to go about assessing whether or not they are kinds. That is the job of the first part of the paper. Once we have a clear idea of the best method for assessing kindhood, we will be able to tackle the issue of the kind status of emotions and concepts directly.

There are many differing accounts of natural kinds. However, when we consult the literature, we find that the overwhelmingly dominant view (at least as far as biology and psychology are concerned) is the homeostatic property cluster (HPC) account, from Richard Boyd (1989, 1991, 1999a, b, 2003, 2010a, b, 2013). The HPC view has a wide following (Kornblith 1993, Wilson et al. 2007), and it has even been described as the ‘consensus’ (Samuels and Ferreira 2010, p. 222) or ‘received’ view (Ereshefsky and Reydon 2015, p. 969).

There are certain core features of the HPC account that will be important for this paper. The first is that members of a kind must instantiate a ‘property cluster’: a set of properties that reliably repeats itself in nature in ‘an important number of cases’ (Boyd 1989, p. 16; 1991, p. 129). It is not the case that all of the entities within the kind must instantiate all of the properties within the cluster. Rather, instantiation of the properties can admit of exceptions (1989, p. 16). In this way, Boyd distances himself from traditional kind essentialism (e.g. Ellis 2002).

This property cluster alone is insufficient for kindhood: the properties must also be underwritten by a homeostatic mechanism: a (typically causal) mechanism that generates the properties in the cluster (Boyd 1989, p. 16). This mechanism *explains* why the properties tend to cluster together. It also explains why it is possible to infer from the fact that some members of the kind instantiate certain properties to the claim that other members of the kind will probably instantiate those properties: because these properties are caused by the same mechanism (Boyd 1991, pp. 130–131; 1999a, p. 68).

This is connected with Boyd’s insistence that the properties in the cluster should be determined by the causal structure of the world, rather than by a priori stipulation (Boyd 1989, p. 16). Boyd says:

the unity of the property-cluster which defines [the kind] is *causal rather than conceptual*... a natural kind is associated *causally* with a large family of methodologically important properties (1991, p. 141).

This quotation raises two important points. First, Boyd insists that there must be a *large* number of properties in the cluster that defines the kind. Second, Boyd insists that the properties in the cluster must be ‘methodologically important’. By this, Boyd means that the properties must be *of interest* to the scientific discipline that is examining it, and that they must be useful for scientifically interesting generalisations, predictions and explanations. Boyd is explicit on this issue (1999a, p. 69; 1999b; 2003, p. 538; 2013, p. 53). We will return to these points below.

Boyd's influence on the debates concerning the kind status of mental faculties is clear. Machery explicitly allies his account with Boyd's (2005, pp. 447–448). Like Boyd, Machery emphasises that members of a kind must share:

a large set of scientifically relevant properties... because of some common causal mechanism (2005, pp. 447–448. Cf. 2009, p. 232).

Griffiths is also explicitly in agreement with Boyd in his discussion of emotions. He spends much time developing the HPC view (1997, chs. 7–9). Griffiths also emphasises that members must be underwritten by a homeostatic mechanism (1997, p. 242), an idea that is inspired by Boyd's work (Griffiths cites Boyd extensively).

3 A space of kinds

In this section, I show that the HPC view can fail to give a determinate answer to whether a certain set of entities counts as a natural kind. I attribute this to three main factors. I suggest that this indeterminacy contributes to making the view ill-suited to helping us decide whether a certain collection of entities counts as a kind. In Sect. 4, we will apply these issues to questions concerning emotions and concepts.

I must be clear that this section is *not* intended in the spirit of hostility toward the HPC view. To say that the HPC is ill-suited for this role is not to criticise it, because as we shall see later, there are good reasons to think that the HPC view should not be deployed in this role at all.

The first factor comes from vagueness in language. As noted above, Boyd and his followers claim that members of a kind must tend to share a 'large' number of properties in order to qualify as a natural kind. Obviously, the word 'large' is vague. This raises an issue: how we interpret Boyd's account will depend upon how strictly we interpret this vague criterion. For example, someone who interprets 'large' as meaning *a very great many* properties would be more likely to *exclude* a certain collection of properties shared by a group of entities from qualifying those entities as forming a kind. Conversely, someone who interpreted 'large' in a more liberal way would be more permissive, and would be willing to count far *fewer* common properties as sufficient for a group of entities to count as a kind. These two interpretations of Boyd would deliver very different constraints on what it takes for something to count as a HPC kind. Two thinkers could then examine the same collection of entities, and one would claim that they do not form a HPC kind, whilst the other would insist that they do, simply as a result of these different constraints. Because of vagueness in the term 'large', the HPC view does not give a determinate answer to the question of how many properties must be present for something to count as a kind.

Vagueness in language also leads to problems when we consider other elements of Boyd's view. We noted that one criterion that a collection of entities must fulfil to be a HPC kind is that it must be possible to make a large number of scientific generalisations about the entities in question. Once again, this introduces vagueness into the view, regarding how many generalizations would have to hold true, or how successful such inductive practices would have to be.

A second factor contributing to indeterminacy over whether the HPC view counts a certain set of entities as a kind comes from *fineness of grain*. One thinker could claim that a set of entities counts as a natural kind because they tend to share some properties specified at an extremely coarse-grained level of detail. Someone else could insist that the class of entities must share properties at a much more fine-grained level of analysis.

A third factor concerns mechanisms. Recall that the HPC view involves the claim that the properties in the cluster are supported by a common causal homeostatic mechanism. There is no consensus on the nature of psychological mechanisms (Bechtel 2008; Piccinini and Craver 2011; Machamer et al. 2000; Illari and Williamson 2012). However, on all of these accounts of what a ‘psychological mechanism’ is, there is room for disagreement similar to those noted above. For example, suppose we take the individuation conditions on mechanisms to be their inputs, outputs and intermediate processes (Machamer et al. 2000). Even if all parties agree on this, it is clearly true that all three of these elements can be specified at different levels of grain. One party could insist on specifying inputs, outputs and processes at a coarse-grain, whilst the other party could demand that more fine-grained distinctions be made between different varieties of input, output and processes. Different opinions on these matters will lead to differences in whether something counts as one mechanism or not, which can obviously result in a difference of opinion over whether some properties are supported by the same mechanism or not, and a resultant disagreement about kindhood. In a sense, we can see this third factor as an application of the second factor (fineness of grain) to the individuation of mechanisms.

I have outlined three factors that can result in the HPC account’s inability to give a determinate answer to whether a certain group of entities is a kind: vagueness in language, fineness of grain in the individuation of properties and the correct individuation of mechanisms. We can imagine a multi-dimensional space of different interpretations of the HPC account. Because of the three factors, it is possible to interpret Boyd’s account in a more liberal, or a more strict way along several different dimensions. These differences will deliver different verdicts on the question of whether a certain collection of entities is a natural kind. In what follows, we will see precisely how these considerations interact with the debates over emotions and concepts.

Two important clarifications: first, I am not saying that the HPC view will *never* help us to decide whether a particular group of entities is a kind or not. Even with such a vast space of possible interpretations, there might be clear-cut cases. Second, this section is only meant to demonstrate the vagueness in the *HPC view itself* when it comes to deciding whether a group of entities forms a kind, not claim that there is *no way* to decide whether a group of entities forms a kind.

4 Concepts and emotions

We now turn to examining some concrete cases where the considerations outlined in Sect. 3 have impact on questions of psychological kindhood.

4.1 Concepts

Machery argues that concepts themselves do not form a kind, but claims that there exist three individual sorts of concepts (prototypes, exemplars and theories), each of which is itself a kind (2009). Machery explicitly relies upon the HPC view in his argument. He bases his argument on the claim that concepts do not share a large set of scientifically interesting properties, which permit many generalisations (2005, p. 450; 2009, pp. 241–242). In this respect, Machery is directly using the HPC view to give criteria by which to judge kindhood. Furthermore, (as we shall see) Machery invokes the criteria for natural kindhood given by Boyd at many points in order to respond to his critics.²

One of Machery's most prominent opponents is Dan Weiskopf.³ He thinks that concepts are a natural kind (2009, 2010).⁴ Weiskopf agrees with Machery that prototypes, exemplars and theories are each natural kinds. The point of difference is that Weiskopf believes that the overall class of concepts *taken as a whole* is a natural kind, which has the various other kinds as sub-kinds, whilst Machery rejects this claim.

Weiskopf's argument is that there are several important properties that are shared by all concepts, and several interesting generalisations that can be made about them (notice that these arguments are explicitly based upon criteria from the HPC view). Weiskopf argues that one such property is conceptual combination: all concepts can be combined with at least some other concepts (according to certain syntactic rules) to produce thoughts that have those concepts as constituents. For example, we combine WOODEN and SPOON to form the concept WOODEN SPOON (Weiskopf 2009, pp. 164–165).

This is Machery's response:

[the claim that all concepts are involved in conceptual combination] is correct. It is, however, unclear why this is taken to justify conserving the notion of concept. Evidence suggests that in conceptual combination, prototypes, exemplars, and theories fulfil different functions... Thus, prototypes, exemplars and theories are likely to be used by different subprocesses of the process underwriting concept combination (2009, p. 245).

Weiskopf points out that all concepts are involved in conceptual combination, and he is happy to accept this abstract and coarse-grained property as sufficient to count as a scientifically relevant property, which aids the case that concepts are a natural kind. Machery responds by insisting that at a finer level of analysis, the processes involved

² Terminological point: Machery reserves the term 'natural kind' for homeostatic cluster kinds that support many interesting generalisations (2010b, p. 238). So on this terminology it is possible for something to be a homeostatic cluster kind and *not* a natural kind. I will stick to the nomenclature on which these terms are interchangeable, but nothing will turn on this.

³ For others, see Piccinini and Scott (2006), Margolis and Laurence (2010), and Strohminger and Moore (2010).

⁴ Weiskopf's understanding of kinds is: 'Kinds are understood here as groupings of entities that participate in a body of empirically discovered reliable generalizations, and which participate in those generalizations due to some set of properties they have in common' (2009, p. 147).

in ‘conceptual combination’ are different (they fulfil different functions, and work in different ways).

Also in support of his claim that concepts form a natural kind, Weiskopf points out that concepts are all stored in long-term memory (2009, pp. 166–167). Machery responds:

[concepts] are stored in long-term memory all right, but the rules that govern storage, permanence, and retrieval are likely to be different (2009, p. 245).

To Weiskopf, ‘being stored in long term memory’ is sufficient to count as a scientifically interesting property, and thus the fact that concepts share this property can aid his case that concepts are a natural kind. To Machery, this is not enough. Machery demands that they must all have similarity at a finer grain of analysis: they must all be stored in long term memory *in the same way*, following the same set of rules.

The links of this to Sect. 3 should be clear: Weiskopf and Machery are placing different constraints upon how much various entities must have in common in order to count them as a natural kind. Weiskopf is liberal, he only requires concepts to share a few properties, specified at a coarse-grain of analysis in order to count concepts as a natural kind. Machery is stricter, for him concepts must share *more* properties, and they must exist at a much *finer grain*.

We find similar issues elsewhere. Samuels and Ferreira (2010) point out various properties that concepts share, and various generalisations that can be made about them, arguing that this qualifies them for kindhood. Machery’s response is to defer to the HPC account of natural kinds and insist that the properties Samuels and Ferreira mention only support ‘few’ as opposed to ‘many’ scientific generalizations, and that this disqualifies them from counting as a natural kind (2010b, pp. 237–238). Again, a strict interpretation of the claim that natural kinds must support ‘many’ interesting generalisations will deliver a verdict in line with Machery; a more liberal interpretation will vindicate the converse verdict.⁵ This occurs yet elsewhere. Edwards opposes Machery’s view by insisting that concepts share various properties in common, though (by Edwards’ own admission) these are only very coarse-grained properties (2011, pp. 125–129).

The third factor highlighted in Sect. 3 (the level of grain at which to individuate mechanisms) is also relevant. Weiskopf argues that all concepts are involved in *conceptual acquisition*, a property that is underwritten by the mechanism of ‘structural alignment’ (2009, pp. 165–166).⁶ Machery’s response is:

it is unlikely that except at a very coarse grain, the processes involved in the acquisition of prototypes and of exemplars are really similar (2009, p. 245).

Weiskopf is happy with an abstract characterisation of the mechanism that underpins conceptual acquisition, Machery demands a more fine-grained specification of the mechanism.

⁵ Machery and Samuels and Ferreira both mention that natural kinds are vague (Machery 2010b, p. 238; 2005, p. 448; Samuels and Ferreira 2010, p. 223) but both of them claim that their interpretation is the correct one.

⁶ Weiskopf’s point is intended to apply to exemplars and prototypes, not theories.

At this point, the following suggestion may be made: the fact that concepts share *at least some* properties or mechanisms (albeit at an abstract level) indicates that they are natural kinds, but that it is very often useful to make divisions amongst sub-kinds, and invoke the more specific properties that Machery mentions. However, this suggestion simply begs the question against Machery. After all, Machery's claim is that these properties and mechanisms do not qualify concepts for kind status *at all*, whether or not these kinds contain sub-kinds. So we cannot assume that these properties and mechanisms are sufficient for kindhood without question begging.

4.2 Emotions

Griffiths thinks that emotions aren't a natural kind. Louis Charland (1995, 2002) thinks they are.⁷ He explicitly commits to the HPC view of natural kinds (2002, p. 512). Charland's argument for the claim that emotions are natural kinds is based around two claims. First, he argues that all emotions 'involve their own distinct mode of representation' (p. 522). The kind of representation that Charland has in mind is that 'emotions are normative or evaluative judgements' (p. 522), though (unlike cognitivists) he is not committed to the claim that these judgements are cognitive or propositional (they could be more akin to perception).⁸ He draws very heavily on Panksepp (1998) in this view. The second part of Charland's argument is that 'there are reliable generalizations and principles of inference that govern emotional behaviour' (p. 522), though Charland admits that there are only a small number of laws that govern these (p. 524). In short, according to Charland, all emotions are 'felt affective' representational states, and it is these unifying features are enough to qualify emotions as a kind.⁹

Griffiths replies to Charland by saying that:

the category of felt affective states is so broad, it is natural within this framework to seek distinctive *kinds* of processes involving affective feelings (2004b, §2.2).

Charland claims that emotions all share the property of being 'felt affective states', and claims that this is sufficient to count members of the category as natural kinds.

⁷ At least, he claims that the hypothesis that they are is plausible (2002, p. 526). Terminological point: Charland expresses his view by saying that *emotion* (singular) is a natural kind, by which he means that emotions (plural) as a whole form a natural kind, as opposed to the claim that some *individual* emotions (such as fear or hate) are natural kinds (2002). In this paper, when I say that *emotion(s)* (in the singular or plural) form a natural kind, I mean that the category *as a whole* forms a kind. I won't discuss the kind status of individual emotions.

⁸ Charland also claims that emotion is a *neurobiological* kind, based on the claim that different emotions share clusters of neurobiological properties across mammalian species (2002, pp. 517–520). Discussion of this would take us too far afield. Even if emotions are a neurobiological kind, it doesn't follow that they are psychological kinds. Something can be a kind for one discipline, but not another (cf. Boyd 1999b, pp. 159–160, Magnus 2012, pp. 39–45). Thanks to an anonymous referee for drawing my attention to this.

⁹ Charland's method deserves comment. Unlike Griffiths, Charland does not examine those faculties referred to by the vernacular category 'emotion'. Rather, he primarily focusses on 'basic' emotions (2002, p. 531) and argues from the claim that *organisms that have* emotions ('emoters') form a natural kind to the claim that emotion is a natural kind (2002, p. 512). This method is interesting, but for reasons of space I shall discuss it no more here.

Griffiths' view is more divisive: he claims that the differences between different sorts of felt affective states shows that the category of felt affective states is not a natural kind. Charland uses the commonalities shared by some entities, specified at a coarse level of grain, and concludes that the entities are a natural kind. Griffiths insists that this is not enough, he demands *more* commonalities at a *finer* level of detail and concludes that they are *not* a natural kind. Similarly, Martha Nussbaum claims that all emotions have the unifying feature of being 'evaluative judgements' (2001). Griffiths' reply follows the general theme that I have been uncovering, of claiming that there are important differences between kinds of evaluative judgements (2004b, §3).

5 Two functions for the HPC view

We have uncovered an impasse in the debates over concepts and emotions. The impasse has been traced to various common factors: differences of opinion over the number of properties, the level of grain of properties and the individuation of mechanisms that qualify entities for natural kindhood. How can we resolve this impasse? It should be clear that we cannot defer to the HPC view to help us here. This is because the HPC view is indeterminate in the ways outlined in Sect. 3: it does not itself tell us the level of grain at which to individuate properties and mechanisms, the number of properties that must be present, and so on. Deferring to the HPC view will not help us here, because the HPC view itself is just too indeterminate to deliver an answer one way or the other. My main focus has been on emotions and concepts, but of course these three factors are features of the HPC view itself, so they are general problems that we should expect to arise with any project that relies on the HPC view to decide whether or not something is a kind.

Does this mean that the HPC view should itself be rejected? No. The HPC view has been asked to fulfil at least two different theoretical roles. One is as a tool for carving up entities into kinds and non-kinds. On this role, the HPC view is treated like a cookie cutter: it is used to help us chunk reality up into kinds and non-kinds. Call this the *taxonomic* role. Boyd himself deploys the view in a taxonomic role, when he uses it to decide which biological categories form natural kinds. He primarily argues that species are natural kinds (1991, 1999b) but he also argues that higher taxa are too (2010b).¹⁰ The arguments of this paper provide some reason to be sceptical of the HPC view *when it is deployed to fulfil the taxonomic role*, because of the three sources of indeterminacy isolated above.

However, this is not the only role that the view has been put to. One of Boyd's main stated motivations for the HPC view is to explain in very general terms what the structure of natural kinds is, in a way that steers between the extremes of traditional essentialism and conventionalism. To Boyd and his followers, such a view must also be compatible with a sensible scientific realism (Boyd 1989, pp. 6–9; 1991, pp. 127–133). In order to fulfil this role, an account of natural kinds must tell us very generally what

¹⁰ Boyd's method is complex, as he claims that the HPC view should itself be modified when it is applied to higher biological taxa in order to fit the conclusion that they are natural kinds (2010b). Analysis of these issues is beyond the scope of this paper. Thanks to an anonymous referee for drawing my attention to this.

sorts of things natural kinds *are*, how they are structured and what the core features of kinds are. Call this the *metaphysical* role. To fulfil this role, the account needs a certain amount of specificity: it needs to be specific enough to show how the features of HPC kinds differ from the features ascribed to kinds by the essentialist and the conventionalist. However, it can perfectly well contain this level of specificity (and thus fulfil the metaphysical role) whilst not delivering a definitive answer on how many properties must be in the cluster, what level of grain they must be individuated at, how mechanisms should be individuated, etc. Without giving clear answers to these questions, the HPC view is of limited use in fulfilling the taxonomic role but it can still perfectly well fulfil the metaphysical one. It's a bad cookie cutter, but it can still be a good explanation of what kinds are.¹¹

Similar things go for the accounts of mechanisms that I have mentioned. The various accounts specify what the core features of psychological mechanisms are, what their metaphysical structure is. They make use of notions such as 'changes', 'functions', 'dispositions', 'component parts', 'start and termination conditions', 'inputs' and 'outputs' (Bechtel 2008; Illari and Williamson 2012; Machamer et al. 2000). These accounts do not give precise ways to individuate all of these features, so they are of limited use in helping us sort out how many mechanisms are present in a given brain area, or in telling us precisely where one mechanism ends and another begins. But they needn't fail at giving us a very general account of what mechanisms are. They may fulfil the metaphysical role without fulfilling the taxonomic one.

This point has far-reaching implications beyond emotions and concepts. For example, Carl Craver criticises the HPC view for providing 'little guidance in arbitrating among taxonomies of scientific kinds' (2009, p. 582) and for being 'unable to settle disputes among those who disagree about the taxonomy of natural kinds' (2009, p. 584). With the distinction of this section, we can put Craver's argument in a wider context: he may be correct that the HPC view is ill-equipped to do this, but it is important to temper this conclusion with the claim that the HPC view may still be worth preserving for other theoretical purposes.

6 The indeterminacy view

Let us summarise. We have examined certain features of the HPC view pertaining to its indeterminacy. We have found that because of this indeterminacy, the view is ill-suited to be used as a tool to help us divide up kinds and non-kinds (i.e. it is ill-suited to fulfil the taxonomic role). So, it cannot help us adjudicate the disputes over whether emotions

¹¹ P.D. Magnus distinguishes between what he calls the 'taxonomy' question and the 'ontology' question. For Magnus, the taxonomy question concerns what the general features of a class of entities are that marks it off as a natural kind, rather than an arbitrary class, and the ontology question concerns what kind of metaphysical structure the kind has, or what 'manner of stuff' it is made of (2015, p. 2. See also Magnus 2014 and Hawley and Bird 2011). The distinction suggested here is similar (the metaphysical role is similar to Magnus' ontological question), but the taxonomic role put forward in this paper is different from Magnus' taxonomy question. The taxonomic role considered here does not concern giving the general features that distinguish kinds from arbitrary classes, but using the HPC view as a cookie cutter to actually perform the taxonomising of kinds from non-kinds. It is *this* role that I claim the HPC view is ill-suited to. Thanks to an anonymous referee for drawing my attention to Magnus' work.

and concepts are natural kinds. In this section, I will outline what I take to be a more profitable approach to questions of kindhood, which places less emphasis on using the criteria in the HPC view to adjudicate questions of kindhood, and has an increased emphasis on examining the explanatory and predictive practices of psychology. I argue that when we do this, the view that emerges is that it is indeterminate whether they are kinds or not.

We have already seen that the HPC view cannot help us decide one way or the other when it comes to the kind status of emotions and concepts. Can we not already infer from this to the conclusion that it is indeterminate whether they are kinds or not? Such an argument would be problematic because it would involve inferring from the fact that the criteria in the HPC view fail to give a determinate answer to the question of whether emotions and concepts are kinds to the conclusion that it *really is* indeterminate whether they are kinds or not. This would be to assume that the criteria in the HPC view should be used as the last word on the kind status of a group of entities. It would be to use the HPC view as a cookie cutter. We have already seen that this is problematic.

6.1 Practice and kindhood

We have seen the inadequacy of using the HPC view directly to help us decide on the kind status of emotions and concepts. To determine a better way, we should return to the point that natural kinds are, most fundamentally, those things that underpin the predictive and explanatory successes of a science. This suggests a way of resolving the problems identified. Specifically, whether a set of coarse-grained properties and mechanisms shared by a particular set of entities qualify those entities for kindhood should not be determined by interpretation of the HPC view. Rather, it should be determined by whether *those properties and mechanisms underpin successful explanations and predictions within psychology*. Notice that the claim is not just epistemic. It is not just that featuring in predictive/explanatory successes is a good guide to kindhood (though it is). It is also metaphysical: it is that featuring in such practices is part of what makes something a kind at all.

This suggests a different method: take the set of coarse-grained properties and mechanisms shared by concepts and emotions, and see how they play out in prediction and explanation. If they feature in successful prediction and explanation, they are indicative of kindhood. If not, they are not. In this way, the question of whether they are definitive of kindhood does not come down to an interpretation of the HPC view itself, but rather to what features in successful research programmes in psychology. Boyd occasionally hints at an approach like this:

considerations of explanatory and inductive significance determine the appropriate standards of individuation for the property cluster itself (1999b, p. 144).

It is worth re-iterating why there is a need for such a fresh approach. This need is generated by the problems uncovered above, concerning difference of opinion over whether emotions and concepts are natural kinds, coupled with the claim that the HPC view itself cannot help us here, because indeterminacy in the view makes it a

bad cookie cutter. So, the issues raised in Sects. 1–5 set up the need for a different approach, and the focus on prediction and explanation is just such an approach.¹²

6.2 Emotions and concepts

Suppose we examine the predictive/explanatory practices that the properties shared by emotions and concepts feature in. We might find that they are clearly successful (they feature in most or all of the predictions and explanations that form the theory); or we might find that they are clearly not successful (they feature in none, or that they feature in predictions which are false). If either of these is the case, all well and good: we will have our answer to the one way or the other. However, there is a third possibility, which is that they feature in explanatory/predictive practices that are *neither clearly successful nor clearly unsuccessful*. That is, we find that the properties shared by concepts and emotions *do* feature in some predictive/explanatory practices, but they are ones that fall within a grey area between success and non-success. If this were the case, then our examination of the explanatory/predictive practices of psychology would not have fixed an answer to the question of whether emotions or concepts are natural kinds: it would have left this question unfixed. They would be neither indicative of kindhood, nor of non-kindhood.

How could explanatory/predictive practices fall within such a grey area? There are several ways that this could be the case. Here I outline three. First, if we find that the few properties shared by concepts and emotions *do* feature in predictions and explanations, but they only feature in a few such explanations and predictions, compared to the number that can be generated by making distinctions between different kinds of concepts and emotions. A second way that such explanatory/predictive practices can fall into a grey area between clear success and clear non-success is if they do make such predictions/explanations, but that these only form background assumptions of the theory in question, rather than those predictions that are subject to direct experimental testing. A third way that they could fall into the grey area is if they do make certain predictions, but that these predictions have to be made more precise before they are tested. These are all ways that the predictive/explanatory successes that the properties shared by emotions and concepts could fall into a grey area between clear success and non-success.

Indeed, I think that a combination of these factors is the case for both emotions and concepts. I do not have the space for an exhaustive survey of empirical work that is relevant to this issue, but I will give two case studies to illustrate my view.¹³

Recall that Charland (drawing on Panskepp) claims that emotions share the property of being ‘affective appraisals’, which govern and move the behaviour of organisms

¹² Obviously, I do not wish to imply that the various interlocutors that I have discussed are *unaware* of the importance of the predictive and explanatory success of the discipline. On the contrary, this issue is discussed (e.g. Machery 2009, p. 245; 2010b, pp. 237–238). My point is just that they *over-rely* on the criteria in the HPC view itself, and that a greater emphasis on psychological practice is required.

¹³ I will focus on the *properties* shared by concepts and emotions, rather than the *mechanisms* that underpin them, as there is more agreement about what properties are shared by concepts and emotions, so we can proceed without too many contentious assumptions. This does not damage my argument.

that have emotions (2002, p. 512). Griffiths responds by claiming that this category is too broad for emotions to count as a natural kind. Given that this is one of the foci of the debate, it will be helpful to examine what role the ‘affective appraisal’ property plays in prediction and explanation. I will take the application of this programme to emotional behaviour in lambs as my main case study.

What we find when we examine the role that the ‘affective appraisal’ property (shared across emotions) plays in this inter-species study is that this property *is* used to support a particular prediction. Specifically, it can be used to predict that animals with emotions should be able to appraise situations, and instigate reactive behaviour in response to this appraisal. Based on this shared property of emotions, it has been predicted that reactive behaviour should be seen in lambs, cattle, rats, etc. (Panksepp 2000, p. 137; Désiré et al. 2004; Lanier et al. 2000). When we further examine the way that this overarching prediction features in the work on cross-species emotional behaviour, we see that it is a background assumption, which motivates the more specific study of emotions in particular species. For example, Désiré et al. clearly invoke this prediction in the rationale for their study of emotional behaviour in lambs:

Appraisal theories... are intended to be applicable to different degrees of cognitive complexity across species... (2004, p. 363).

The property of ‘being an affective appraisal’ (shared by emotions) appears in a prediction, which features as a background motivation. That is, it predicts that we should find affective behaviour across mammalian species. This, in turn, motivates the project of investigating emotional behaviour in nonhuman mammals at all.¹⁴ This may be thought to support the claim that the properties shared by emotions are predictively/explanatorily useful. This would, in turn, support the claim that emotions are kinds. However, when we examine the predictions that are actually subject to experimental testing in this research programme (and the predictions and explanations that are most successful), we find that they rely on more fine-grained distinctions between different *kinds* of affective appraisals. These fine-grained distinctions are used to make specific behavioural and physiological predictions, and it is these predictions (based on these fine-grained properties) that are experimentally tested.

For example, if we take the fine-grained appraisal faculty that appraises a situation for *suddenness* then we can predict that animals with this appraisal mechanism should exhibit startle reflexes and an increased heart rate when exposed to sudden stimuli. A range of experiments tested this very fine-grained prediction: an animal is exposed to a sudden stimulus and its behavioural/physiological reactions are then measured for particular kinds of response, and compared to the predicted startle response (e.g. Désiré et al. 2004, pp. 364–346). The prediction is confirmed in a wide variety of nonhuman mammals, making it highly successful (Désiré et al. 2004, p. 371; Yeomans et al. 2002). Then, a fine-grained appraisal mechanism for *suddenness* is postulated

¹⁴ Notice that the claim is not that the property of being an affective appraisal requires extensive background assumptions in order to be explanatorily and predictively successful, but rather that these properties can be used to make predictions that themselves form only a background assumption of the theory. Thanks to an anonymous referee for drawing my attention to this ambiguity.

to explain this behaviour.¹⁵ Similarly, by using the fine-grained property of *being an appraisal for novelty*, then we can make another specific prediction: that on exposure to novel stimuli, animals should exhibit an orienting response (evinced by looking behaviour) and a variable heart rate. Such a prediction can be tested: expose an animal to novel stimuli, observe its behaviour, and compare the behaviour to the predicted orienting response. Again, this prediction is widely confirmed (Turpin 1986; Bertson et al. 1992; Désiré et al. 2004).¹⁶ Such behaviours and physiological reactions would be explained by postulation of a different appraisal mechanism that appraises for novelty.¹⁷

We have examined the predictive/explanatory role of the abstract property of ‘being an affective appraisal’ that is shared by all emotions. We have found that the role it plays in the theory is not straightforwardly clearly successful or clearly unsuccessful: it *does* feature in predictions, which feature in the theory as background assumptions. However, the predictions and explanations that are made using the more fine-grained properties that are not shared across all emotions are those that the experiments primarily test, and it is these predictions and explanations that are most clearly successful. We will come back to what this means for kindhood shortly, but first let us return to concepts. As we saw above, a core question about concepts is whether the property (possessed by all concepts) of conceptual combination qualifies concepts for kindhood. To adjudicate this, we should look at whether this property features in successful predictions/explanations.

We can take Costello and Keane (2000) as a case study, as they give a computational model of conceptual combination that makes use of exemplars, theories and prototypes. On this view, when atomic concepts are combined, there are several possible distinct interpretations of the resultant complex concept. That is to say, there are several possible meanings that could be assigned to the complex concept that is the output of the combination process. According to Costello and Keane, these interpretations are judged along three dimensions, and the most plausible interpretation (based on these criteria) is assigned to the combined concept (2000).

In this model, the general coarse-grained property shared by all concepts (that they are involved in conceptual combination) figures to generate a certain prediction: that concepts of all kinds (exemplars, theories, prototypes) should be combinable to generate new concepts. This features as an assumption in the background of Costello and Keane’s theory. They assume that *all* kinds of concepts (exemplars, theories and prototypes) are accessible to the combination process, and can be used by these processes to generate conceptual combinations (2000, pp. 305–306). That is to say, they take the property of ‘being available to conceptual combination’, and assume that this

¹⁵ Cf. Scherer (2001, pp. 114–115), where particular predictions are made by dividing the category of ‘affective appraisals’ up into sub-kinds of emotion.

¹⁶ Lambs were also tested for their responses to unpredictability, but the results are less clear-cut (Désiré et al. 2004, p. 373).

¹⁷ Charland comments that emotional behaviour cannot be explained without using affective terms, a point he takes to support emotion’s kind status (2002, p. 523). This doesn’t vindicate emotion as a kind. My claim is about the limitations in the predictive and explanatory power of the coarse-grained property ‘affective appraisal’. The theoretical indispensability of particular affective terms (such as ‘novelty’) is compatible with this.

property is a feature of *all* concepts, and they use this as a background assumption in their theory. This assumption allows them to explain features of complex concepts by reference to the fact that the many individual kinds of concepts (such as prototypes, exemplars or theories) are all inputs to the combinatorial process.

To put all of this in simpler terms: Costello and Keane refer to the fact that all concepts work as input to the combinatorial process in their explanation of how complex concepts are formed. This is the assumption that all kinds of concepts are available for combination at work. So here we see the property of ‘being available for conceptual combination’ functioning as part of the explanatory apparatus of the theory.

However, in order to predict which particular meanings will be assigned to certain particular complex concepts (and explain how they acquired those properties), we must refer to particular features of specific *sub-types* of concepts. For example, in explaining how complex concepts can have emergent properties,¹⁸ Costello and Keane make reference specifically to exemplars, and claim that it is exemplars over which the computational processes operate (at least in these cases) (2000, p. 34; Gray and Smith 1995).¹⁹ Here we refer to exemplars in giving our prediction of which properties a certain complex concept has, and explaining how it could have such emergent properties. The abstract and coarse-grained property of ‘being available to conceptual combination’ that is shared by all concepts cannot do this, because it is insensitive to the particular specific properties of certain sub-types of atomic concepts (not shared by all concepts) that are an indispensable part of the explanation.

Other cases of conceptual combination need to make use of prototypes specifically. Take the hypothetical case of the complex concept CACTUS FISH. Costello and Keane take the correct interpretation of this complex concept to be something like ‘a prickly fish’. One factor contributing to this prediction is that prickliness is a typical feature of cacti (thus explaining why a cactus fish would be prickly). The encoding of *typical* properties is a feature of *prototypes* (cf. Johnson and Keil 2000). So prototypical features figure in the prediction and explanation of which particular interpretation is given of the complex concept.²⁰

Recall that we decided to settle the question of kindhood by asking whether the properties shared across emotions and across concepts were explanatorily and predictively successful. We took ‘affective appraisals’ and ‘availability to conceptual combination’

¹⁸ Emergent properties are typical properties of the referents of complex concepts, which aren’t typical properties of the referents of their constituent concepts. For example, people typically associate properties such as ‘small’, ‘pretty’ and ‘caged’ with pet birds, even though these aren’t properties typically possessed by pets or birds.

¹⁹ The claim is that in interpreting PET BIRD, we retrieve memories of instances of pet birds, and then use their properties (being small and in a cage) to interpret the complex concept.

²⁰ It may be replied that (assuming Costello and Keane’s basic model to be correct) conceptual combination is governed by the same three criteria in all instances [plausibility, diagnostic accuracy and informativeness (2000)]. This may be taken to be evidence of shared properties possessed by concepts, which feature prominently in the theory, thus supporting kindhood. But this inference wouldn’t be right for two reasons. First, it is doubtful that the criteria do govern the combination of *all* concepts, as they only account for noun-noun combinations, and cannot account for metaphor (Costello and Keane 2000, pp. 303–304; Keane 1997). Second, the fact that they govern conceptual combination is not really a shared feature of concepts themselves, but rather of the combinatorial process that operates *over* concepts.

as our examples of properties shared by emotions and shared by concepts respectively, because these are two core sources of disagreement in the debates over emotions and concepts. What are we to make of our investigation? Take the predictive/explanatory practices that the properties shared across concepts, and across emotions feature in. Should we say that they are *successful* or that they are *unsuccessful*? Above I said that an explanatory/predictive practice can fall into a grey area between successful and unsuccessful. This would be the case if the predictions feature only as a background assumption of the theory, or if there are only a few predictions and explanations compared to those that are generated by more specific properties, or if further refinement is required before the predictions are experimentally tested. In such cases, it is neither clearly the case that these practices are sufficiently successful, nor that they are clearly unsuccessful.

I have argued through an examination of two case studies that this is the case with the properties shared by concepts and emotions. Both of them do generate predictions that are part of the background of the theory, but they generate only a few predictions compared to those generated by more specific properties, and both of them require refinement before the predictions that are primarily tested in each case. In the emotions case, they generate the prediction that all emotional animals should show certain reactive behaviours (which is a background assumption of the theory), but distinctions are made between different kinds of reactive behaviour in the predictions that are tested. In the concepts case, they generate the assumption that all concepts are available for combination (which is a background assumption of the theory), but this property is refined to explain the meanings assigned to particular complex concepts in particular cases. In a nutshell, it is unclear whether to say that these practices are ‘successful’ or not, because they *do* produce certain predictions/explanations, but it is just not clear whether these are *successful enough*.

The message is this: psychology makes use of a range of predictive/explanatory practices. Some of them are clearly successful and others are clearly unsuccessful. However, with emotions and concepts, it is not like that. Rather, they have certain features that make it very unclear whether to think of them as successful or not. They fall into the grey area. This is the main conclusion of this subsection. We should now turn to the link between this claim and psychological natural kindhood.

6.3 Indeterminacy

We now have two claims on the table. The first is that we should judge the kind status of emotions and concepts by whether they feature in successful explanatory/predictive practices in psychology (rather than using the HPC view as a cookie cutter). This is a methodological claim. The second is that, upon investigation, these practices fall within a grey area between success and unsuccess. This is an empirical claim. We can now marry these claims. The result is that the natural kind status of emotions and concepts remains unfixed by the practices of psychology. Thus, given the current state of psychological research it is simply *indeterminate* whether concepts or emotions are natural kinds. This I call the indeterminacy view.

Note that I am *not* saying that it is indeterminate whether concepts and emotions are kinds because they share only a few coarse-grained mechanisms and properties in common. That in itself would not be a problem, if these coarse-grained properties/mechanisms and properties were highly successful in the theoretical context of psychology. Rather, the indeterminacy arises from the particular role that these properties play in the predictive/explanatory practices of psychology.

An objector may resist this conclusion. They may accept the case studies of Sect. 6.2, and accept that emotions and concepts play a modest role in predictions/explanations in psychology. However, they may take this to support the claim that they *are* natural kinds after all, on the basis that they play *at least some* predictive/explanatory role. It will be instructive to unpick the reasoning of this objection. The objection infers from the claim that emotions and concepts possess properties that support some modest explanatory/predictive practices to the conclusion that they are kinds. In order for this inference to work, we need to make an additional assumption: that these sorts of modest predictive and explanatory practices are *sufficient* to qualify a set of entities for kindhood. But this additional assumption is not itself justified by an examination of the explanatory and predictive practices of psychology. Psychology itself does not fix whether these modest explanatory and predictive practices are sufficient for kind status. To see the problematic nature of this assumption, consider that another thinker may make the converse assumption: they may assume that these modest explanatory/predictive successes are so modest that they indicate that emotions and concepts are *not* kinds. But both of these inferences are unsatisfactory. Both of these inferences partially involve examination of psychological practice (in order to establish the modest predictive/explanatory practices that concepts and emotions figure in). However, they also both essentially rely on assumptions about the link between these modest predictive/explanatory practices and kind status. These assumptions themselves are not supported by examination of psychological practice because whilst psychology makes use of a range of such practices, it does not fix a precise point at which they qualify for kindhood. We should be wary of such assumptions that do not themselves derive from an examination of psychology.

Three clarifications: first, the indeterminacy view is *not* epistemic. It is not that we just don't know whether concepts or emotions are natural kinds. Rather, it really is indeterminate whether they are or not. To repeat a point made above, this is because the kind status of a group of entities is determined (at least in part) by the explanatory/predictive practices of science, and these practices (in this case) do not fix the answer one way or the other. Second, the indeterminacy view could change in the future, if the explanatory/predictive practices of psychology change enough. There is nothing suspicious about this. Part of what makes something a kind at all is how it features in scientific practice, so we would expect kinds to change depending on the state of the science. Third, note that this indeterminacy is different from other kinds of indeterminacies that have been pointed out before. It is known that HPC natural kinds can have extensions that do not have sharp boundaries (Boyd 1999b, p. 144). The indeterminacy argued for here is quite different: it is that it can be indeterminate

whether a particular category is a kind or not at all, regardless of whether its extension is sharp.²¹

7 The consequences of indeterminacy

Suppose we accept that it is indeterminate whether emotions and concepts are natural kinds. What are the consequences of this view? I highlight three. First, one of the overriding assumptions fuelling the debates over the kind status of emotions and concepts is that psychology's practices have a certain *neatness*: that there is a clear cut answer to whether a certain faculty is a natural kind. If the indeterminacy view is taken seriously, this is false. The practices of psychology are just messier than that. Different sorts of messiness in psychology have been noted before. For example, I have previously argued that psychology sometimes uses a variety of different concepts for the same phenomenon, which carve up the mind differently (Taylor 2017). It has also been shown that there can be a fuzzy boundary between two psychological kinds (Buckner 2015). The current paper is in agreement with these views in spirit, and adds to them another kind of messiness: indeterminacy over whether something is a kind at all. Of course, these results have repercussions beyond philosophy of psychology, to our general view of the unity (or disunity) of science itself (e.g. Dupré 1993).

The second consequence of the indeterminacy view regards the kind status of other mental faculties such as consciousness, perception, seeing, the senses and constancy. If the arguments of this paper are correct, then we can add the indeterminacy view as a contender view of these faculties. Of course, I am *not* saying that it really is indeterminate whether these faculties are kinds. Clearly that would require further argumentation. My claim is only that the indeterminacy view should be put on the table as an option in these debates.

To see the third consequence of the indeterminacy view, consider that the claim that concepts and emotions are not a natural kind is often used as sufficient reason for *eliminativism* about the terms 'concept' and 'emotion': that they should be dropped from our scientific terminology. Likewise, the converse claim that concepts and emotions are natural kinds is often given to justify preservation of these terms. If the indeterminacy view is accepted, it will mean that *both* sides of the debate are wrong: natural kinds cannot be used to support eliminativism *or* preservationism, because both options require there to be a determinate answer to the question of kindhood.

²¹ Magnus (2012, pp. 47–55) discusses a similar point: that kinds can be more or less natural depending on the strength of the science's epistemic success. My point here is in the same spirit, but different: it is that it can be entirely indeterminate whether something is a kind *at all*, not just that kinds themselves come on a spectrum. Magnus is also motivated by more overarching concerns in the philosophy of science, rather than specific issues about psychological kindhood. Magnus himself places two constraints on kindhood: that they support inductive and explanatory success in the discipline, and they be necessary for this success [he calls this latter criterion the 'restriction clause' and develops alternative versions of it (2012, pp. 48–67)]. Notice that we can accept these criteria on kindhood and still accept that it is indeterminate whether concepts and emotions are kind. If emotions and concepts are *necessary* for modest predictive and explanatory practices they feature in, then it would not be clear whether the practices qualified them for kind status in the first place.

With the indeterminacy view, natural kinds would be powerless to help us decide what to say here. Yet we must say something, as it is a profoundly important question about whether radical conceptual change is required in psychology or not. At this point, we have two options. The first would be to stick with the assumption that the terminology of science should track natural kinds, and then infer from the claim that it is indeterminate whether emotions and concepts are natural kinds to the conclusion that it is simply indeterminate whether the terms ‘emotion’ and ‘concept’ should be eliminated or preserved. The second would be to reject the assumption that natural kinds should be the overriding factor in dictating scientific terminology, and to settle the eliminativism/preservationism question based on other factors. Myself and Peter Vickers have made suggestions like this before. We have argued that the most important factor in adjudicating disputes over eliminativism should be purely pragmatic factors to do with how much the term is liable to promote verbal disputes and other methodological problems, rather than whether the term refers to a kind (Taylor and Vickers 2017). Whether this pragmatic approach will ultimately work out, we must leave for another time.

Acknowledgements Thanks to Ned Block, Robin Hendry, Maja Spener, Peter Vickers and three anonymous reviewers for comments on previous drafts of this paper. Thanks also to Edouard Machery for helpful conversation. I am grateful to the Leverhulme Trust and the Isaac Newton Trust for an Early Career Fellowship (ECF-2015-088) that supported me while the paper was written. The research for this paper started with my PhD thesis. I am thankful to Durham University for a Durham Doctoral Studentship, and to the Royal Institute of Philosophy for a Jacobsen Studentship that supported me during this work.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bechtel, W. (2008). *Mental mechanisms*. Sussex: Psychology Press.
- Bertson, G., Boysen, S., & Cacioppo, J. (1992). Cardiac orienting and defence responses. In B. Campbell, H. Hayne, & R. Richardson (Eds.), *Attention and information processing in adults and infants*. Hillsdale: Erlbaum.
- Block, N. (2012). The grain of vision and the grain of attention. *Thought*, 1, 170–174.
- Boyd, R. (1989). What realism implies and what it does not. *Dialectica*, 43, 5–29.
- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61, 127–148.
- Boyd, R. (1999a). Kinds, complexity and multiple realization: Comments on Millikan’s “historical kinds and the special sciences”. *Philosophical Studies*, 95, 67–98.
- Boyd, R. (1999b). Homeostasis, species and higher taxa. In R. Wilson (Ed.), *Species: New interdisciplinary perspectives*. Cambridge: MIT Press.
- Boyd, R. (2003). Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism. *Philosophy and Phenomenological Research*, 66, 505–553.
- Boyd, R. (2010a). Realism, natural kinds and philosophical methods. In H. Beebe & N. Sabbarton-Leary (Eds.), *The semantics and metaphysics of natural kinds*. New York: Routledge.
- Boyd, R. (2010b). Homeostasis, higher taxa and monophyly. *Philosophy of Science*, 77, 686–701.
- Boyd, R. (2013). What of pragmatism with the world here? In M. Baghramian (Ed.), *Reading Putnam*. New York: Routledge.
- Buckner, C. (2015). Transitional gradation in the mind. *British Journal for the Philosophy of Science*, 67, 1091–1115.

- Burge, T. (2010). *The origins of objectivity*. New York: OUP.
- Charland, L. (1995). Emotion as a natural kind: Towards a computational foundation for emotion theory. *Philosophical Psychology*, 8, 59–84.
- Charland, L. (2002). The natural kind status of emotion. *British Journal for the Philosophy of Science*, 53, 511–537.
- Costello, F., & Keane, M. (2000). Efficient creativity. *Cognitive Science*, 24, 299–349.
- Craver, C. (2009). Mechanisms and natural kinds. *Philosophical Psychology*, 22, 575–594.
- Davies, W. (2016). Colour constancy, illumination and matching. *Philosophy of Science*, 83, 540–562.
- Désiré, L., Veissier, I., Després, G., & Boissy, A. (2004). On the way we assess emotions in animals: do lambs (*Ovis aries*) evaluate an event through its suddenness, novelty or unpredictability? *Journal of Comparative Psychology*, 118, 363–374.
- Dupré, J. (1993). *The disorder of things*. Cambridge: Harvard.
- Edwards, K. (2011). Higher level concepts and their heterogenous implementations: A polemical review of Edouard Machery's *Doing Without Concepts*. *Philosophical Psychology*, 24, 119–133.
- Ellis, B. (2002). *Scientific essentialism*. Cambridge: Cambridge University Press.
- Ereshefsky, M., & Reydon, T. (2015). Scientific kinds. *Philosophical Studies*, 172, 969–986.
- Gray, K. C., & Smith, E. E. (1995). The role of instance retrieval in understanding complex concepts. *Memory & Cognition*, 23, 665–674.
- Griffiths, P. E. (1997). *What emotions really are: The problem of psychological categories*. Chicago: University of Chicago Press.
- Griffiths, P. E. (2004a). Emotions as natural and normative kinds. *Philosophy of Science*, 71, 901–911.
- Griffiths, P. E. (2004b). Is emotion a natural kind? In R. Solomon (Ed.), *Thinking about feeling: Contemporary philosophers on emotion*. New York: Oxford University Press.
- Griffiths, P. E. (2008). Current emotion research in philosophy. *Emotion Review*, 5, 215–222.
- Hawley, K., & Bird, A. (2011). What are natural kinds? *Philosophical Perspectives*, 25, 205–221.
- Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for the Philosophy of Science*, 2, 119–135.
- Irvine, E. (2013). *Consciousness as a scientific concept: A philosophy of science perspective*. Dordrecht: Springer.
- Johnson, C., & Keil, F. (2000). Explanatory understanding and conceptual combination. In F. C. Keil, & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 327–359). Cambridge, MA: MIT Press.
- Kornblith, H. (1993). *Inductive inference and its natural ground: An essay in naturalistic epistemology*. Cambridge, MA: MIT Press.
- Lanier, J., Grandin, T., Green, R., Avery, D., & McGee, K. (2000). The relation between the reaction to sudden, intermittent movements and sounds and temperament. *Journal of Animal Science*, 78, 1467–1474.
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1–25.
- Machery, E. (2005). Concepts are not a natural kind. *Philosophy of Science*, 72, 444–467.
- Machery, E. (2009). *Doing without concepts*. New York: Oxford University Press.
- Machery, E. (2010a). *Precis of Doing Without Concepts*. *Behavioral and Brain Sciences*, 33, 195–206.
- Machery, E. (2010b). The heterogeneity of knowledge representation and the elimination of *concept*. *Behavioral and Brain Sciences*, 33, 231–244.
- Magnus, P. D. (2012). *Scientific enquiry and natural kinds: From planets to mallards*. New York: Palgrave Macmillan.
- Magnus, P. D. (2014). NK ≠ HPC. *The Philosophical Quarterly*, 64, 471–477.
- Magnus, P. D. (2015). Taxonomy, ontology and natural kinds. *Synthese*. <https://doi.org/10.1007/s11229-015-0785-2>.
- Margolis, E., & Laurence, S. (2010). Concepts and theoretical unification. *Behavioral and Brain Sciences*, 33, 219–220.
- Nudds, M. (2011). The senses as psychological kinds. In F. Macpherson (Ed.), *The senses*. New York: OUP.
- Nussbaum, M. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge: CUP.
- Panksepp, J. (1998). *Affective neuroscience*. Oxford: OUP.
- Panksepp, J. (2000). Emotions as natural kinds within the brain. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions*. New York: Guilford.
- Piccinini, G., & Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, 183, 283–311.

- Piccinini, G., & Scott, S. (2006). Splitting concepts. *Philosophy of Science*, 73, 390–490.
- Prinz, J. (2002). *Gut reactions: A perceptual theory of emotion*. Oxford: Oxford University Press.
- Prinz, J. (2012). *The conscious brain*. New York: Oxford University Press.
- Samuels, R., & Ferreira, M. (2010). Why don't concepts constitute a natural kind? *Behavioral and Brain Sciences*, 33, 222–223.
- Scherer, K. (2001). Appraisal considered as a process of multi-level sequential checking. In K. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion*. Oxford: OUP.
- Shea, N. (2012). Methodological encounters with the phenomenal kind. *Philosophy and Phenomenological Research*, 84, 307–344.
- Strohinger, N., & Moore, B. (2010). Banishing the thought. *Behavioral and Brain Sciences*, 33, 225–226.
- Taylor, H. (2017). Attention, psychology and pluralism. *British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axx030>.
- Taylor, H., & Vickers, P. (2017). Conceptual fragmentation and the rise of eliminativism. *European Journal for the Philosophy of Science*, 7, 17–40. <https://doi.org/10.1007/s13194-016-0136-2>.
- Turpin, G. (1986). Effects of stimulus intensity on autonomic responding: The problem of differentiating orienting and defense reflexes. *Psychophysiology*, 23, 1–14.
- Weiskopf, D. (2009). The plurality of concepts. *Synthese*, 169, 145–173.
- Weiskopf, D. (2010). The theoretical indispensability of concepts. *Behavioral and Brain Sciences*, 33, 228–229.
- Wilson, R. A., Barker, M. J., & Brigandt, I. (2007). When traditional essentialism fails: Biological natural kinds. *Philosophical Topics*, 35, 189–215.
- Yeomans, J., Li, L., Scott, B., & Frankland, P. (2002). Tactile, acoustic and vestibular systems sum to elicit the startle reflex. *Neuroscience and Biobehavioural Reviews*, 26, 1–11.