Disambiguating the question whether mental disorders are natural kinds

Olivier Lemeire – This a draft

Abstract

This paper argues that the difficulty in determining whether mental disorders are natural kinds is partly due to the ambiguity of the phrase 'natural kind' itself. This phrase is used to talk about several different phenomena. On the account defended in this paper, 'natural kind' is ambiguous between several combinations of three more basic concepts, namely that of a *kind*, that of a *representation corresponding to reality*, and that of a *phenomenon belonging to the domain of the natural sciences*. By using the phrase 'natural kind' for various conjunctions of these basic concepts, the phrase is not only ambiguous but also less precise than desirable. This paper argues that it would be less ambiguous and more precise for philosophers of psychiatry to ask whether mental disorders are kinds, whether their representation corresponds to reality, and whether they belong to the domain of the natural sciences.

Introduction

Currently, mental health problems are mostly diagnosed using the DSM-5; the fifth version of the Diagnostics and Statistical Manual of Mental Disorders (APA 2013). This manual lists the different types and sub-types of psychiatric disorders and the associated criteria that must be met for someone to be diagnosed with the disorder. Based on the criteria outlined in this manual, mental health professionals classify different individuals as suffering from the same type of mental disorder. Many questions can be asked about the different types of disorders that are recognized and the criteria based on which they are diagnosed. In philosophy, these questions are often formulated by asking whether types of mental disorders are *natural kinds*. Interesting papers and books chapters addressing this question appear regularly.[1]

Though interesting, these contributions also invite confusion. That is because – as will be argued – the phrase 'natural kind' is ambiguous and imprecise. The phrase is ambiguous because different authors use it with a different meaning. Just like the word 'bank' is ambiguous between meaning *edge of a river* or meaning *financial institution*, so the phrase 'natural kind' is ambiguous between different meanings. It is quite literally the case that many different questions have been asked about types of mental disorders by asking whether they are natural kinds.

The phrase 'natural kind' is also imprecise. Each of the different meanings that this phrase can express is a combination of three more basic concepts, namely that of a *kind*, that of a *representation corresponding to reality* and that of a *phenomenon belonging to the domain of the natural sciences* (henceforth simply called *natural phenomenon*). Some authors, for example, intend the phrase 'natural kind' to denote those kind categories that correspond to reality, a conjunction of the concept *kind* and the concept *representation corresponding to reality*. Others, however, use the same phrase to denote kinds that can be explained in purely biological or neurological terms; a conjunction of the concept *kind* and the concept *natural phenomenon.* In this way, the phrase is ambiguous between several combinations of three more basic concepts, creating confusion.

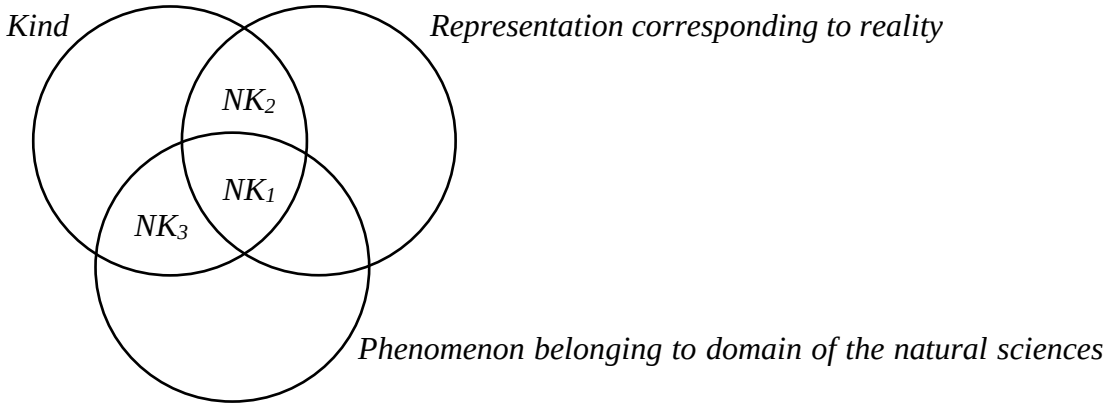
The claim that the phrase 'natural kind' creates confusion is not a new one. Several philosophers have already argued that there is disagreement about the defining criterion of natural kinds (Cooper 2005; Haslam 2014; Zachar 2003). In the first two sections of this paper, a new ambiguity view is proposed and distinguished from a nearby 'disagreement view.' Distinctive of this ambiguity view is that it also provides perspective on how to move beyond the current confusion; one should replace all talk of 'natural kinds' by the more precise notions 'kind,' 'representation corresponding to reality' and 'natural phenomenon.' The third, fourth, and fifth section illustrate how one can do so. The goal of this paper is to argue that whatever

philosophical issues can be ambiguously and imprecisely addressed by asking whether mental disorders are natural kinds, can be addressed unambiguously and more precisely by asking whether there are kinds of mental disorders, whether their representation corresponds to reality, and whether mental disorders are natural phenomena.

1. Ambiguity

The phrase 'natural kind' causes confusion in philosophical debates about mental disorders. On the view defended in this paper, this confusion is largely due to the phrase being ambiguous between several different conjunctions of three more basic concepts. Visually, this ambiguity can be represented as follows:

Figure 1:



Each of the circles represents the extension of one basic concept; the concept of a *kind*, the concept of a *representation corresponding to reality*, and the concept of a *natural phenomenon*. These basic concepts will be discussed in more detail below (sections 3, 4 & 5). The point here is that the phrase 'natural kind' has been used in the literature with (at least) three different meanings – namely NK_1, NK_2 and NK_3 – based on different conjunctions of these three more basic concepts.[2] On its first meaning, 'natural kind' expresses the complex concept NK_1; a conjunction of the basic concepts *kind*, *representation corresponding to reality* and *natural*

phenomenon. On that reading, the phrase denotes those phenomena that satisfy all three of these concepts. Zachar has characterized natural kinds in this way, stating that:

> Typical examples of natural kinds are chemical elements such as gold, biological species such as tiger, and infectious diseases such as tuberculosis. These all: a) are naturally occurring as opposed to artificial; b) have clearly demarcated boundaries separating members of the natural kind from non-members; c) possess observable features that are causally produced by internal properties; and d) these causal properties can be used to objectively validate category membership. (Zachar 2015, 288)

The five different criteria mentioned here by Zachar are a combination of criteria that determine whether something is a kind, whether its category is a faithful representation of reality, and whether something is a natural phenomenon.[3] Something being naturally occurring rather than artificial, for example, is not a criterion that determines whether a class of objects constitutes a kind, nor a criterion that determines whether a category corresponds to reality. Instead, it is more specifically a criterion to determine whether a phenomenon belongs to the domain of the natural sciences. Criteria (b) and (c) are possible criteria of kindhood. Criterion (d), furthermore, is a criterion of realism. By using these four criteria to define 'natural kinds,' Zachar intends the phrase to denote those phenomena that supposedly satisfy all three concepts, like gold, tigers, and tuberculosis.

Other authors, however, use the phrase 'natural kind' to express the complex concept NK_2. On that understanding, the concept is a conjunction only of the more basic concepts *kind* and *representation corresponding to reality*. One example of this interpretation can be found in the following fragment by Kincaid:

> On this view, we have defensible natural kinds in the study of psychopathology if we have evidence that our classifications pick out distinguishable groups of individuals and that those classifications allow us to identify well-grounded patterns between

> membership in that group and other variables. The conception of natural kinds is a liberal one, but that does not mean it makes natural kinds a matter of convention or subjective (Kincaid 2014, 151).

There is no mention in this account of the view that the phrase 'natural kind' refers to phenomena that belong to the domain of the natural sciences, nor is this criterion ever mentioned in Kincaid's text.[4] Instead, what is characteristic of his view is that categories of mental disorders refer to natural kinds if they represent kinds about which we should be realists, rather than think of these categories as conventional or subjective representations of kinds.

A third interpretation of 'natural kind' takes it to express the complex concept NK_3. In that case, it denotes kinds that consist of phenomena belonging to the natural sciences. For example, when considering whether psychopathy constitutes a natural kind, Verplaetse and Focquaert (2020) use the following two conditions:

> A first condition concerns the notion of a kind. One can only speak of a kind when the members that belong to the kind show enough similarities and have sufficiently many points of difference with non-members. […] A second condition elaborates on the notion of naturalness. After all, the kind is not 'cultural' or merely 'conceptual' but 'natural,' which means that manifestations of this category can be reduced to natural processes that are not or not completely influenced by human actions and interests. […] We call this second condition the reduction condition. (Verplaetse & Focquaert 2020, 51, *translated from Dutch original*)

In the view of these authors, 'naturalness' does not concern the extent to which the category of 'psychopathy' corresponds to reality but rather whether it refers to a phenomenon that can be reduced to a natural process. This too is a common way of understanding the phrase 'natural kind,' especially in philosophy of psychiatry and philosophy of psychology. Together with the two other examples, this text fragment provides evidence for the claim that 'natural kind' is

ambiguous between expressing the complex concepts NK_1, NK_2, and NK_3, which are various conjunctions of the more basic concepts *kind*, *representation corresponding to reality*, and *natural phenomenon*.

2. Disagreement

That the phrase 'natural kind' causes confusion in the debate about mental disorders has been noted before (Cooper 2005; Haslam 2014; Zachar 2000). Kincaid and Sullivan, for instance, have previously argued that the "notion of natural kinds can be read in multiple ways" suggesting that they too think of the phrase as being ambiguous (Kincaid & Sullivan 2014, 2). Typically, however, this ambiguity is considered a result of disagreement. In this section, the current ambiguity view is distinguished from this alternative 'disagreement view.'

When the phrase 'natural kind' is introduced in a text, it is often joined by a disclaimer stating that there is a continuum of positions about what constitutes a natural kind. On the one hand of the continuum there would be essentialist theories that propose very demanding criteria for a class to constitute a natural kind (see e.g. Ellis 2001). On the other hand of the continuum there would be very relaxed theories, like the promiscuous realism of Dupré (1993). The popular Homeostatic Property Cluster theory of Boyd is typically placed somewhere in the middle (Boyd 1999). Examples of this presentation of the debate can be found in Cooper (2005), Haslam (2014), and in Kincaid & Sullivan (2014) when they state the following:

> All senses of natural kinds assume that there is something out there in the world that grounds classifications. […] The strongest sense of natural kinds, often called "essentialism," thinks that there are sets of individually necessary and jointly sufficient properties for a given kind that entail a strict in or out classification of all individuals. The weakest sense asserts only that some properties are statistically associated; it allows

> for no sharp, non-arbitrary categorizing of individuals who lie on a continuum. (Kincaid & Sullivan 2014, 3)

On this continuum view, the different senses that exist for the phrase 'natural kind' are the result of proper *disagreement* between authors. That is, on this view everyone who uses the phrase intends it to apply to roughly the same phenomenon – there being 'something out there in the world that grounds classifications' – though there is disagreement about how best to conceptualize this phenomenon and where to draw its boundaries.

In fact, the ambiguity of the phrase 'natural kind' is not just due to this type of disagreement. The quotations provided in the previous section show that not all authors who use this phrase are talking about roughly the same phenomenon. At least some of them are simply talking about a different phenomenon, using the same words. Those who use the phrase 'natural kind' to denote kinds that belong to the natural sciences – rather than the human sciences like psychology or sociology – are not talking about the same phenomenon as those who use the phrase to defend a realist position – rather than a conventionalist or instrumentalist one – about kind categories. The fact that these authors defend different accounts of natural kinds is not because they disagree about a common subject. Their disagreement is only a semantic one about the meaning of the phrase 'natural kind.'

Of course, there is also real disagreement between philosophers who work on natural kinds. There is semantic ambiguity *and* there is disagreement. The disagreement about the notion(s) of 'natural kinds' that does exist, however, can be completely explained as a result of different view about the phenomena denoted by the three more basic concepts. There is real disagreement about how to conceptualize kinds, about how to conceptualize the correspondence of categories to reality, and about how to conceptualize (reduction to) the natural domain. For each of these three phenomena there is a continuum of conceptualizations, ranging from those based on very strict criteria of inclusion to those based on more relaxed criteria. And because the phrase

'natural kind' is used to express various combinations of these more basic concepts, there are also more and less strict versions of the complex concepts NK_1, NK_2 and NK_3. Still, the fact that 'natural kind' can be used to express each of these complex concepts is a case of ambiguity, not of substantive disagreement.

The difference between an ambiguity view and a disagreement view is methodologically important. In philosophy, disagreement about an interesting phenomenon cannot be avoided. If the multiple senses of 'natural kind' were to be a result of philosophers disagreeing about how to conceptualize a single phenomenon, then the most one could do in order to avoid confusion is to clearly define how one intends to use that phrase. It is commendable that so many philosophers already aim to do so when they discuss the question whether mental disorders are natural kinds. Yet if an ambiguity view is correct, one can and should do more. One *should* do more because even when specifying one's own use of the phrase 'natural kind,' it would still be confusing that other authors use the same phrase to talk about a different phenomenon altogether. One *can* do more because whereas substantive disagreement about a complex phenomenon cannot be avoided, semantic disagreement about how to use an ambiguous term is only a matter of language regimentation.

Since the phrase 'natural kind' are ambiguous between expressing the complex concepts NK_1, NK_2, and NK_3, one might suggest that to disambiguate it would be better to use different labels for each of the complex concepts the phrase can express, like 'natural kinds_1,' 'natural kinds_2' and 'natural kinds_3.' A better and more feasible proposal, however, is to simply stop talking about 'natural kinds' altogether. After all, given that the different concepts expressed by the phrase 'natural kind' are conjunctions of three more basic concepts, it would be less ambiguous to simply express one's philosophical views mental disorders by talking of 'kinds,' 'representations corresponding to reality,' and 'natural phenomena' (or alternative specific terms).

Not only would this change in language avoid the ambiguity caused by the phrase 'natural kind,' it would also have the additional benefit of being more precise. Given that 'natural kind' is typically used to express a conjunction of several more basic concepts, *denying* that something is a natural kind is an imprecise claim that can be true in several different ways. If the phrase 'natural kind' is used as meaning NK_2 for instance – the conjunction of the basic concepts *kind* and *representation corresponding to reality* – types of mental disorders could fail to be natural kinds because their representation does not correspond to reality or because they do not refer to kinds, or because they do neither. A person who claims that types of mental disorders are *not* natural kinds (understood as NK_2) can therefore intend to communicate that they fail to offer a realistic representation, fail to refer to kinds, or fail to do both. Hence even if one is clear about one's intended meaning of the phrase 'natural kind,' denying that this term applies to mental disorders would still be an imprecise claim that can be true in several different ways.

The confusion caused by the use of 'natural kind' is therefore not only due to the ambiguity of the phrase but also due to the fact that its different meanings are conjunctions of several more basic concepts. Both these problems can be solved by expressing one's views by simply talking of 'kinds,' 'representations corresponding to reality,' and 'natural phenomena.' The next three sections illustrate that whatever philosophical issues can be addressed about types of mental disorders by asking whether they are natural kinds, can be addressed more precisely and less ambiguously by asking three more precise questions; whether they refer to kinds, whether they provide a representation that corresponds to reality, and whether they are (reducible to) natural phenomena.

3. The kindhood question

One question that can be asked about mental disorders is whether they constitute kinds. Is it correct to say, as many lay people and mental health professionals intuitively do, that individuals who satisfy the same diagnostic criteria listed in DSM-5 suffer from the same *kind* of disorder (Adriaens & De Block 2013)? Or considered independently of current diagnostic manuals; are there even *kinds* of mental disorders? Perhaps psychopathology is a domain of reality that, unlike many other domains, does not have a kind-like structure. To answer these questions requires a criterion of kindhood, that is, an account of what it is for something to constitute a kind. Based on such a criterion one can then consider which, if any, mental disorders constitute kinds.

A first possible criterion of kindhood is that members of the same kind are similar to each other in many respects and differ in many respects from non-members. This is a vague criterion but requires that at the very least members of the same kind have more in common than the properties by which they are classified. This criterion is based on the view that kinds are like clusters of properties that tend to co-occur (Slater 2015). Mental illness would then be 'kind-like' if the domain of psychopathology were to contains clusters of properties that tend to co-occur, possibly including genetic, neurological, behavioral and other types of properties.

To the extent that there are such clusters in the domain of psychopathology, it would be epistemically fruitful to have a reliable and valid way of determining whether an individual instantiates one of these clusters. The resulting groupings of individuals would allow for useful generalizations and predictions. However, for the same reason that it is epistemically fruitful to recognize clusters where they exist, it would be epistemically harmful if research practices and clinical practices were to depend on an expected clustering of properties where it does not exist. Hence criticisms of DSM-5 often focus on the heterogeneity of those individuals who are similarly classified or on the non-specificity of their properties (Insel et al. 2010). By design

DSM-5's *polythetic* classification system allows for substantial heterogeneity between the symptom profiles of similarly diagnosed individuals. Focusing on eighteen DSM-5 categories, Olbert, Gala and Tupler (2014) calculated that "in most categories, 2 individuals with the same diagnosis may share no symptoms in common, and that any 2 theoretically possible symptom combinations will share on average less than half their symptoms." This polythetic approach to classification is based on an expected similarity in other respects, beyond the heterogeneous symptom profiles. Critics argue, however, that this promise has not been borne out, saying for example that:

> *DSM* categories have not been shown to be strongly associated with empirically validated syndromes, biological or psychological correlates, or other variables of interest (e.g., response to treatment, genetics). To the extent that (weak) associations have been demonstrated, it is likely that the categories are essentially 'free riders' contributing little or nothing to the association: i.e., specific symptoms and aggregations of small effects from independent sources are probably doing most of the work. (Poland 2014, 34).

If this is indeed the case and individuals who are similarly diagnosed share only few properties that are distinctive of their disorder, then a kind-diagnosis should carry little weight in the epistemic practices of psychiatry. This is a first aspect of the kindhood debate about mental disorders.

A second possible criterion of kindhood that is often considered is that members of a kind are causally homogenous. On this criterion, the mental health issues of various individuals would be instantiations of a single *kind* of disorder if they have the same cause. When a diagnostic category refers to a class that is causally homogenous, this category is also epistemically useful in a distinct way. Causal homogeneity entails that diagnosing someone as suffering from a type of disorder also provides explanatory information about the causes of

their disorder (Maung 2016). This causal-explanatory power of a diagnostic category is particularly important for a medical discipline like psychiatry, since the goal is often to effectively *intervene* in causal processes (Cooper 2005). It is no wonder then that Kincaid and Sullivan consider the true mark of 'natural kindhood' that a category groups together phenomena that are "subject to the same type of causal explanation […] and respond similarly to the same kinds of causal interventions" (2014, 2). Again, however, many critics have argued that current psychiatric categories do not refer to classes that are causally homogenous and should therefore not be relied on for explanatory practices in psychiatry (Kendler 2012; Maung 2016; Poland et al. 1994).

A third and final criterion of kindhood requires kinds to have categorical boundaries (Haslam 2014). For there to be a difference in kind between mental disorders, it is argued, requires a categorical boundary such that someone either does or does not instantiate the kind. For the most part, current psychiatric categories are defined as if mental disorders have categorical boundaries. Especially in the decade leading up to the publication of DSM-5, however, researchers have argued that the domain of psychopathology is to a large extent dimensionally structured rather than category-like. (Kraemer 2007; Widiger & Trull 2007). To what extent mental disorders are best conceptualized as discrete taxa or as quantitative variations along a continuum, is a question most directly addressed by taxometric analyses (Meehl 1995). The most recent meta-analysis of such psychometric analyses concludes that "dimensional findings overwhelmingly outnumbered taxonic findings," including in the domain of psychopathology (Haslam et al. 2020, 7). If mental disorders are dimensional rather than taxonic, at least some epistemic tasks in psychiatry will be best served by dimensional diagnostics and measurements rather than categorical ones.

The goal here is not to settle these kindhood questions about mental disorders. Instead, the goal is to illustrate that asking whether different criteria of kindhood apply to specific mental

disorders and to the domain of psychopathology more generally is a less ambiguous question than asking whether mental disorders are 'natural kinds.' Nevertheless, these kindhood questions still result in a complex debate. One reason for this complexity is that there are different criteria of kindhood and that these criteria need not always be co-instantiated. In some scientific domains, many kinds do satisfy all three criteria. This is the case when several properties cluster by virtue of being the effects of the same common cause. Kind that are defined based on such causally important common causes – like chemical elements for example – satisfy all three criteria. There is, however, no *a priori* reason to expect to find such kinds in all scientific domains. As Boyd and others have already argued, kinds in the special sciences should rather be thought of as Homeostatic Property Clusters, where the clustering can be supported by more complex causal mechanisms (Boyd 1999; Kendler et al. 2011). One step further still, is to recognize the possibility that there are varying degrees in which groups can be associated with a cluster of properties, be causally homogenous, and have categorical boundaries; three distinct dimensions of kindhood. Answering the kindhood question about mental disorders requires an evaluation of the applicability of each possible criterion to the domain of psychopathology and to draw the consequences for various research and clinical practices in psychiatry.

4. The realism question

A second question that can be asked about mental disorders is whether the categories offered by DSM-5 correspond to reality or instead provide a conventional representation, determined by particular epistemic interests or practical concerns (Beebee & Sabbarton-Leary 2010). The same realism question can also be asked independently of current diagnostic manuals; is it possible to provide an objective representation of mental health issues that corresponds to

reality or is any representation necessarily also subjective; imbued with the particularities of the minds who are responsible for the representation?

This question about the reality corresponding to representations of mental disorders is distinct from the kindhood question introduced in the previous section. For one, some philosophers propose such a strict criterion of realism that categories of mental disorders could refers to groups that satisfy each criterion of kindhood outlined above yet would still only be useful instruments or conventions rather than realistic representations. Secondly, the realism question does not exclusively apply to kind categories but also to other possible ways of conceptualizing mental health issues, for example using network models, dimensional scales, or terms referring to causal properties. Hence it allows for a more precise debate when both the kindhood question and the realism question are distinguished.

This is not to say that both questions are not related as well. Given that the DSM-5 conceives of categories for mental disorders as referring to kinds, a first way of approaching the realism question is by asking whether mental disorders really are as kind-like as they are made out to be. In that case, the realism question reduces to the kindhood question. On Boyd's view, for example, we should be realists about categories – be it categories of kinds or of properties – to the extent that they allow for successful inductive inferences and explanations. This is his accommodation thesis: "[k]inds useful for induction or explanation must always 'cut the world at its joints' in this sense: successful induction and explanation always require that we accommodate our categories to the causal structure of the world" (Boyd 1991, 139). Note that on this view, it is irrelevant to one's realism about categories whether the epistemic aims for which these categories are useful are conventionally selected within a scientific discipline. After all, whatever the epistemic aims within a discipline happen to be, achieving those goals requires categories that accommodate one's epistemic practices to the real causal structure of the world.

Others, however, have a more stringent conception of realism, according to which realism requires total independence of any conventionality (Franklin-Hall 2015). On this view, one should not be a realist about scientific categories when the epistemic aims of a scientific discipline are conventionally selected and influence how a domain is categorized, useful as the categories may be. After all, the category would still not provide a representation of reality as it is independently of our categorizing minds. The categories would also reflect our categorizing minds and hence not provide us with a picture of a mind-independent reality. On this more stringent criterion of realism, the realism question no longer reduces to the kindhood question.

A second aspect of the realism question – beyond the conventionality of one's epistemic aims – is the extent to which DSM-5 categories and other possible representations of mental disorders are reflections of various *practical* concerns. In several papers, Zachar has argued for the instrumentalist position that categories of mental disorders are *practical* kinds (2003, 2014, 2015). According to him, "discovery alone cannot tell us how to classify" (2014, 90) and "concepts for psychiatric disorders are constituted by discoveries *and* decisions" (Zachar 2015, 289). Psychiatric categories are tools that are subject to many goals and interests, like "the clinical goals of practitioners and patients, the various scientific goals of researchers […], the priorities of health service administrators and social policy analysists, and commercial interests" (Zachar 2015, 289). Since the categorization of mental health issues requires one to balance these competing goals, Zachar argues, the resulting categories can at most be 'practical' kinds. On this view too, one should not be a realist about psychiatric categories even if they were to refer to groups that are kind-like in every way outlined above.

Furthermore, the realism question can also be a way of addressing not just the influence of epistemic and practical aims on categorization, but also the conventionally inherent in one's scientific methods. In the case of psychometric analyses, for example, the method often requires

substantive conventional assumptions so that it is not clear how realist one can still be about the resulting representations. As Borsboom explains:

> Thus, the introduction of pragmatic concerns and mathematically convenient fictions is essential to the representational goal of psychometrics, but it compromises the explanatory goal because it endows the model with properties that are very unlikely to hold in reality. As a result, as the model becomes more parsimonious, identifiable, estimable, and mathematically elegant, it often loses its substantive plausibility (Borsboom 2017, 46).

Hence one aspect of the realism question is whether it is possible to determine which aspects of one's methodology are conventional and in what way these conventional aspects affect the resulting representations. It is these and other considerations that are to be taken into account when deciding whether one should be a realist about representations of mental disorders. These considerations can be and should be separated from questioning the kindhood and reducibility of mental disorders.

3. The reducibility question

A third and final question that can be asked about mental disorders is whether they belong to the domain of the natural sciences. For some, this is part of what they mean to ask when they ask whether mental disorders are 'natural kinds.' To what extent can disordered behaviors, emotions and thoughts that appear to be subjective and imbued with meaning, ultimately be reduced to and explained in terms of natural processes? This third question is distinct from both previous questions; whether mental disorders are reducible to the natural domain says little about the extent to which they are kind-like. Similarly, whether mental disorders are reducible says little about whether one should be a realist about representations of mental disorders.

Yet like the other two questions, the reducibility question about mental disorders is a complex one. To simplify then, there are those who hold that psychiatric disorders are ultimately brain disorders, such that "behavioral disorders that characterize psychiatric illness are disturbances of brain function, even in those cases where the causes of the disturbances are clearly environmental in origin" (Kandel 1998). Similarly, Hyman stated that psychiatric disorders "are real illnesses of a real organ, the brain, just like coronary artery disease is a disease of a real organ, the heart" (Hyman 1999). According to this biomedical model of psychiatric disorders, the domain of psychiatry falls within the natural domain and psychiatry is ultimately a form of clinical neuroscience (Insel & Quirion 2005). For some, to say that mental disorders are natural kinds is to say that they are kinds from the natural domain, just like chemical elements or biological species.

On the other hand, there are those who argue that mental disorders are, in some way to be specified, irreducibly social, cultural, or mental (Albee & Joffe 2004; Hacking 1998). For them, mental disorders cannot be understood or explained without also invoking insights from the human sciences, like sociology, anthropology, or psychology. Some, for example, highlight the *social* aspects of mental disorders (Poland 2014). It is not only that mental disorders have social causes but also that in many cases, to be disordered is to instantiate a social property; it is to violate some social norm or to stand in some relation to others (van Riel 2016). According to others, *cultural* factors play a crucial role in the formation of psychiatric disorders. In *Creating Mental Illnes*s, Horwitz argues that "most nonpsychotic symptoms stem from general underlying vulnerabilities that may assume many different overt forms, depending on the cultural context in which they arise. […] Cultural processes, not the unfolding of natural disease, structure the overt manifestation of symptoms into recognizable entities" (2002, 108). Perhaps the hardest problem, however, is the relationship between the mind and the brain, and the extent to which the first can be reduced to the latter (Bolton & Hill 1996). Murphy

summarizes why the reducibility of mental, social, and cultural factors to the natural domain is all but evident as follows:

> [T]he causes of many mental illnesses include a mix of genes and environmental factors, as in Meehl's example […] of genetics and bereavement conspiring to cause depression. The genes and environmental effects are different kinds of processes, not different levels of one process. If we were dealing with one process, describable in different ways, then we could anticipate a reductive account, in which higher-level variables are mapped on to lower-level ones. But even though it is hard enough to imagine a molecular or neurological reduction of a psychological construct like humiliation (which […] is depressogenic), it is even harder to imagine a reductive analysis of socio-cultural factors like unemployment or childhood sexual abuse. They have brain effects, but the brain effects vary across classes of individuals in ways that depend on other environmental and genetic contexts. (Murphy 2017)

There is no need to say more here about the contentious issue of the reducibility of mental disorders. The goal of this final section was merely to introduce a third question that can be asked about mental disorders. Specifying this question allows one to avoid the imprecision and ambiguity of asking whether mental disorders are 'natural kinds.'

Conclusion

This paper has made a case for clearly distinguishing the kindhood question, the realism question, and the reducibility question about mental disorders. In the past, each of these issues has typically been addressed by asking whether mental disorders are natural kinds. Yet as I have argued, this phrase 'natural kind' is ambiguous and imprecise; it is used to express various combinations of three more basic concepts. The result is confusion in a debate that sorely needs precision. Much of this confusion can be avoided by eliminating the phrase 'natural kind.'

Instead, one should ask one of three questions. Firstly, whether mental disorders are kinds, rather than for example variations along a continuum. Secondly, whether representations of mental disorders correspond to reality, rather than that they provide conventional though perhaps useful tools. Thirdly, whether mental disorders can ultimately be reduced to the domain of the natural sciences, rather than being irreducibly social, cultural, or mental. Thus, this paper has disambiguated the question whether mental disorders are natural kinds.

References

Adriaens, P. & De Block, A. (2013). Why we essentialize mental disorders. *Journal of Medicine and Philosophy, 38,* 107-127.

Albee, G. W., & Joffe, J. M. (2004). Mental illness is NOT "an illness like any other". *Journal of Primary Prevention*, *24*(4), 419-436.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th edition). Arlington, VA: American Psychiatric Publishing.

Beebee, H., & Sabbarton-Leary, N. (2010). Are psychiatric kinds real? *European Journal of Analytic Philosophy*, *6*(1), 11-27.

Bolton, D., & Hill, J. (1996). *Mind, meaning, and mental disorder: The nature of causal explanation in psychology and psychiatry*. Oxford University Press.

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies, 61,* 127-148.

Boyd, R. (1999). Homeostasis, species and higher taxa. In: R.A. Wilson (Ed.), *Species: new interdisciplinary essays*. Cambridge, MIT Press, 141-186.

Borsboom, D. (2017). Representation and explanation in psychometric modeling. In: K. Kendler & J. Parnas (Eds.) *Philosophical issues in psychiatry IV. Classification of psychiatric illness*. Oxford UP, 45-50.

Cooper, R. (2005). *Classifying madness* (Vol. 86). Dordrecht, The Netherlands: Springer.

Dupré, J., 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Cambridge MA: Harvard University Press

Ellis, B. (2001). *Scientific Essentialism*. Cambridge UP.

Franklin-Hall, L. R. (2015). Natural kinds as categorical bottlenecks. *Philosophical Studies, 172*(4), 925-948.

Hacking, I. (1998). *Rewriting the soul: Multiple personality and the sciences of memory*. Princeton University Press.

Haslam, N. (2014). Natural kinds in psychiatry: Conceptually implausible, empirically questionable, and stigmatizing. *Classifying psychopathology: Mental kinds and natural kinds*, 11-28.

Haslam, N., McGrath, M. J., Viechtbauer, W., & Kuppens, P. (2020). Dimensions over categories: a meta-analysis of taxometric research. *Psychological Medicine*, 1-15.

Horwitz, A. (2014). The social functions of natural kinds: The case of major depression. In: Kincaid, H., & Sullivan, J. A. (Eds.). *Classifying psychopathology*. *Mental kinds and natural kinds.* Cambridge: MIT Press, 209-226.

Hyman, S.E. (1999). White House Conference on Mental Health, June 7, 1999.

Insel, T. R., & Quirion, R. (2005). Psychiatry as a clinical neuroscience discipline. *JAMA: Journal of the American Medical Association, 294*, 2221–2224.

Kandel, E. R., (1998). A new intellectual framework for psychiatry. AmJ Psychiatry, 155, 457-69.

Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*, 1143–1150.

Kendler, K. S. (2012). The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. *Molecular psychiatry*, *17*(4), 377-388.

Kincaid, H., & Sullivan, J. A. (Eds.). (2014). *Classifying psychopathology*. *Mental kinds and natural kinds.* Cambridge: MIT Press.

Kraemer, H.C. (2007). DSM categories and dimensions in clinical and research contexts. *Int. J. Methods Psychiatr. Res.*, 16(S1), 8 – 15.

Maung, H. H. (2016). Diagnosis and causal explanation in psychiatry. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *60*, 15-24.

Meehl, P. (1995). Bootstrap taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50*, 266-275.

Murphy, D. (2017) Philosophy of Psychiatry. *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.).

Olbert, C. M., Gala, G. J., & Tupler, L. A. (2014). Quantifying heterogeneity attributable to polythetic diagnostic criteria: theoretical framework and empirical application. *Journal of Abnormal Psychology*, *123*(2), 452.

Pober, J. M. (2013). Addiction is not a natural kind. *Frontiers in psychiatry*, *4*, 123.

Poland, J. (2014). Deeply rooted sources of error and bias in psychiatric classification. In: Kincaid, H., & Sullivan, J. A. (Eds.). *Classifying psychopathology*. *Mental kinds and natural kinds.* Cambridge: MIT Press, 29-64.

Poland J., von Eckhardt B., Spaulding W. (1994). Problems with the DSM approach to classifying psychopathology. In: Graham G., Stephens G.L., editors. *Philosophical psychopathology*. MIT Press; Cambridge, MA, pp. 235–260.

Samuels, R. (2009). Delusion as a natural kind. *Psychiatry as cognitive neuroscience: Philosophical perspectives*, *49*, 79.

Slater (2015). Natural kindness. *The British Journal for the Philosophy of Science*, *66*(2), 375-411.

Tabb, K. (2019). Philosophy of psychiatry after diagnostic kinds. *Synthese*, *196*(6), 2177-2195.

Tekin, Ş. (2016). Are Mental Disorders Natural Kinds?: A Plea for a New Approach to Intervention in Psychiatry. *Philosophy, Psychiatry, & Psychology*, *23*(2), 147-163.

Tsou, J. Y. (2016). Natural kinds, psychiatric classification and the history of the DSM. *History of psychiatry*, *27*(4), 406-424.

Turkheimer, E. (2017). The hard question in psychiatric nosology. In: K. Kendler & J. Parnas (Eds.) *Philosophical issues in psychiatry VI. Classification of psychiatric illness*. Oxford UP, 27-44.

van Riel R. (2016). What Is Constructionism in Psychiatry? From Social Causes to Psychiatric Classification. *Frontiers in psychiatry*, *7*, 57.

Verplaetse, J., & Focquaert, F. (2020). Psychopathie. In: *Het tweede handboek psychiatrie en filosofie*. Boom Psychologie en Psychiatrie.

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist*, *62*(2), 71.

Zachar, P. (2003). The practical kinds model as a pragmatist theory of classification. *Philosophy, Psychology and Psychiatry*, 9(9), 219-227.

Zachar, P. (2014). Beyond natural kinds: Toward a" relevant"" scientific" taxonomy in psychiatry. In: H. Kincaid & J.A Sullivan (Eds.). *Classifying psychopathology*. *Mental kinds and natural kinds.* Cambridge: MIT Press, 75-104.

Zachar, P. (2015). Psychiatric disorders: natural kinds made by the world or practical kinds made by us?. *World Psychiatry*, *14*(3), 288.

[1] See for example Beebee & Sabbarton-Leary (2010), Cooper (2005), Haslam (2002, 2014), Kendler, Zachar & Craver (2011), Kincaid & Sullivan (2014), Pober (2013), Tabb (2019), Tekin (2016), Tsou (2016), Zachar (2003, 2014, 2015).

[2] I do not intend to claim that these are the only ways in which the phrase 'natural kind' has been used. Some philosophers also seem to use it to just express on one of the three basic concepts.

[3] For a similar use of 'natural kind' consider the following quotation from Haslam: "The schema understands the concept of 'natural kind' in a particular way: as a naturally existing (i.e., objectively occurring, discovered rather than fabricated) discrete category whose causal basis is a shared, biological essence" (Haslam 2014, 17). Also see Tsou: "In this framework, natural kinds are *naturally-occurring classes of objects, properties or processes that are discovered and exist independent of classifiers* (e.g. electrons, H_2O, fish). By contrast, artificial kinds are *conventional classes that are imposed on nature by and invented by classifiers* (e.g. triangles, money, vehicles)" (Tsou 2016, 407).

[4] For a similar use of 'natural kind,' see Cooper (2005)