

**Forthcoming in Broome, M. & Bortolotti, L. (eds.) Psychiatry as Cognitive Neuroscience: Philosophical Perspectives, OUP**

**Delusion as a Natural Kind\***

*Richard Samuels*

**Ohio State University**

**Abstract:**

This chapter clarifies and defends what I call the NK Thesis: The claim that delusions constitute a natural kind. In doing so, I spell out the relevant notion of a natural kind and show why the most prominent objections to the NK Thesis are unsatisfactory. In addition, I present some *prima facie* reasons for adopting the NK Thesis as a working hypothesis, and argue that careful reflection on the standard objections to the thesis provides some insight into the sort of natural kind that delusions constitute. Roughly put: If the NK Thesis is true, then we have reason to suppose that delusions constitute a generic, multiply realized, cognitive kind.

**Introduction**

Though delusions are widely regarded as a central psychiatric phenomenon, their nature and ontological status has, since the inception of modern psychopathology, been the subject of concerted debate (Jaspers, 1914). One important aspect of this debate concerns whether or not delusions constitute a *natural kind*. Crudely put: Do delusions constitute a scientifically respectable kind in nature in the way that, say, molecules or quarks appear to; or are they more like days of the week, or domestic pets in being, in one respect or another, inappropriate for the purposes of scientific enquiry? It is this issue that I propose to discuss here. In particular, I defend what I will call the *NK Thesis*: The thesis that delusions constitute a natural kind.

---

\* I would like to thank Ben Caplan, Dominic Murphy and Tim Schroeder for stimulating discussions of the issues covered in this chapter. I would also like to thank the editors for their helpful comments on an earlier draft.

Two clarifications are in order. First, I won't be concerned merely with the issue of whether or not some *sub-types* of delusions are natural kinds. For instance, the issue is not merely whether Capgras delusion, or persecutory delusions, or any of the myriad sub-kinds of delusion are natural kinds. Rather, I am interested in the more general issue of whether delusions *as such* –the entire category— constitute a natural kind.

Second, though the best argument for the existence of a natural kind is to provide a detailed, well-articulated and highly confirmed account of the kind in question, no such account will be provided here. I would be delighted if such a proposal were available; but contemporary scientific psychopathology affords no such account. Still worse, the natural kinds conception of delusions has in recent years been roundly rejected by many theorists, and for many different reasons. The following discussion, then, is a largely defensive one that seeks to clear the ground for the more ambitious project of saying what natural kind delusions are. With this in mind, I argue that the main extant objections to the NK Thesis are unsatisfactory. But in addition to this, I develop two more positive lines of thought. First, I argue that there are some *prima facie* reasons for taking the NK Thesis seriously. Second, I maintain that a careful consideration of the main arguments for and against the NK Thesis, suggest a range of conditions that delusions need to satisfy, if there are to constitute a natural kind. In spelling out these conditions I seek to indicate what are likely to be the most fruitful avenues for future scientific enquiry.

Here's the game plan. In section 1, I consider the general issue of what a natural kind is and defend what is sometimes called a *homeostatic cluster* account of natural kinds (Boyd, 1991). The remainder of the chapter interleaves critical assessment of the main objections to the NK Thesis with positive morals about how best to develop the thesis. In section 2, I consider a first criticism of the NK Thesis: The *anti-essentialist objection*. Though widely endorsed, I argue that the objection is misdirected since it attributes to the NK Thesis assumptions that it need not –and should not—endorse. Nevertheless, a consideration of this objection yields a range of positive morals, and provides some *prima facie* reason to take the NK Thesis seriously. In section 3, I outline these morals and reasons. Then in section 4, I consider a second kind of objection to the NK Thesis, which I call *continuity objections*. Such arguments purport to show that the NK Thesis is false because delusions fail to constitute a kind of any sort, natural or

otherwise. I argue that whilst these objections show something interesting –roughly, that delusions exhibit various kinds of symptomatic continuity with other phenomena— they fail to show delusions are not a natural kind because these kinds of symptomatic continuity are wholly consistent with the truth of the NK Thesis. Next, in section 5, I consider third family of objections, which purport to show that delusions are a *mind-dependent* kind and, hence, fail to constitute a natural kind. The apparent plausibility of such arguments, I maintain, result from collapsing some crucial and well-motivated distinctions. Given the appropriate distinctions, the arguments themselves collapse. In section 6, I consider a final class of objections –*heterogeneity arguments*— which purport to show that delusions are not a natural kind because they are in some sense too heterogeneous. In response, I point out that providing a satisfactory theory of a natural kind requires that we identify the relevant granularity of description. In view of this, the heterogeneity arguments do not establish that the NK Thesis is false. Rather they merely impose constraints on what the relevant grain of description must be. I conclude with a brief discussion of what I take to be the central empirical challenge for a natural kind conception of delusions: what I call the *unity problem*.

## **1. What is a Natural Kind?**

In order to evaluate the NK Thesis we need to clarify the notion of a natural kind; and this is not an easy task, since the notion has a long –and checkered—history in which it has been characterized many times over. Indeed for much of the twentieth century, the notion of a natural kind was considered little more than an artifact of an ancient and outmoded metaphysics.<sup>1</sup> But in recent years, the notion of a natural kind has regained philosophical respectability in large measure because it has proven useful to understanding some central aspects of contemporary scientific practice (Boyd, 1991; Griffiths, 1997).

### ***1.1. Why Natural Kinds Matter***

---

<sup>1</sup> Of which more in section 1.4.

Which scientific practices does the notion of a natural kind help clarify? Here are three central candidates.

1. *Inductive Generalization* The notion of a natural kind is intimately connected with the problem of induction (Quine 1969, Boyd 1990, 1991, Hacking 1991, Machery 2005). In brief, inductive generalizations are formulated on the basis of observed instances of a category and yet purport to license inferences about unobserved instances as well. But for such purposes not all categories are created equally. On observing 100 birds that have hearts, for example, it is reasonable to suppose that most—even all—unobserved birds also have hearts. In contrast, observing 100 things over 2 meters tall that are made of wood does not license the analogous inference—that most or all things over 2 meters tall are made of wood. For the co-variation between height and being made of wood would be an *accidental* one. The notion of natural kind is relevant to this issue because it is supposed to effect a distinction between these two kinds of classes: those about which non-accidental, scientifically relevant, inductive generalizations can be formulated—atoms, molecules and species, for example—and those about which few, if any such generalizations can be formulated—e.g. things over 2 meters tall (Machery, 2005). Natural kinds permit this distinction because a central aspect of any natural kind is that its members share many non-accidentally related—though logically unconnected—scientifically important properties (or relations).

2. *Objects of Scientific Discovery* In part because the formulation of non-accidental, inductive generalizations is central to scientific practice, many sciences aim to determine which classes of entities are natural kinds relative to their domain of enquiry. Biologists, for example, have identified cells, species and strings of mRNA as kinds relevant to the formulation of inductive generalization, whilst rejecting *élan vital* and domestic animals as kinds of this sort. Notice that the acceptance of such kinds is typically determined *a posteriori*. That is: where a category is accepted as a kind over which scientific generalization are made, it is very typically on the basis of empirical considerations; and when they are rejected for such purposes, the reasons are similarly empirical.

3. *Targets for mechanistic explanation* As noted earlier, natural kinds are supposed to underwrite a rich inventory of non-accidental, inductive generalizations because they are the kinds of categories whose members possess many non-accidentally

related properties (or relations). But such properties are not *logically* related in the way that, for example, being red and being colored are. Rather they are contingently but non-accidentally associated. As Richard Boyd has put it, the instances of natural kinds have *contingently clustering families of properties*: properties that reliably, though need not invariably co-vary (Boyd, 1991). Consider a paradigmatic example of a natural kind: water. Samples of water tend to possess a wide array of characteristics –transparency, potability, specifiable boiling and freezing points, and so on. Moreover, these characteristics are not logically (or conceptually) necessary properties of water samples in the way that, say, being unmarried is a necessary property of bachelors. Clearly, the existence of such property clusters call out for explanation. And, in fact, the provision of such explanations is a widespread scientific practice. Specifically, much science seeks to explain the existence of reliably co-varying property clusters by identifying and specifying the structures, processes or mechanisms that –under appropriate circumstances— causally explain the contingent clusters associated with (natural) kinds. Again, the task of identifying and characterizing such structures, processes and mechanisms is a largely *a posteriori* matter. In the case of water, for example, the relevant explanatory factor (give or take a bit) turned out to be the chemical structure of water molecules, and was not discovered until substantial developments in chemistry had been made.<sup>2</sup>

### ***1.2 Three Further Conditions on Natural Kindhood***

So, the notion of a natural kind is important, in large measure, because it helps explain a range of central scientific practices. But given the role that natural kinds play in science, what characteristics must they possess? Let me start with three fairly obvious characteristics.

*1. Discreteness.* Given that natural kinds are *kinds*, they must to be reasonably *discrete* classes of entities that can be demarcated from other phenomena. But it is important to note that insisting on discreteness is *not* the same thing as maintaining that kinds cannot be vague: that they must be wholly determinate categories without

---

<sup>2</sup> See Weisberg, 2003 for a more detailed discussion of the chemistry of water.

borderline cases. Children constitute a kind, as do red things, flat things, and hexagonal objects. But all these kinds are vague (Sorensen, 2006). So, for example, to our knowledge there is no precise age –specified, for example, in picoseconds—that marks a precise divide between childhood and other life stages. But this alone would be a poor reason for supposing that children are not a kind of any sort.

Even so, one might think that natural kinds are different in this regard: that they must have strictly determinate boundaries. But this fits poorly with the facts about the kinds invoked in science. For many of the kinds that figure in scientific generalizations also exhibit vagueness. The point is well illustrated by biological kinds –tigers and apes, for example, but also eyes and hearts. These kinds are widely regarded as plausible candidates for natural kind status (Kripke, 1972; Putnam, 1975). Yet they are also vague.<sup>3</sup> Indeed the assumption of absolutely strict boundaries is incompatible with the fact that biological entities evolve gradually over time (Dennett, 1995).

2. *Homogeneity.* A second characteristic of natural kinds is that they need to be fairly *homogeneous* kinds. Some classes of objects are just too heterogeneous to be natural kinds. Consider, for example, the class consisting of left knees, brown hairs and chicken curries. Though it is a genuine class –one might even think it constitutes a kind—no one would suggest that it constitutes a *natural* kind. It is just too heterogeneous –too disjunctive— to figure in the sorts of scientific practices mentioned in section 1.1. For example, we would not expect to find many non-accidental, scientifically relevant, inductive generalizations that ranged over the members of this class. Nor would we expect to find some mechanism operating in all –or even most— of these cases that explained the co-variation of whatever properties members of the class happened to share. This contrasts sharply with the paradigmatic exemplars of natural kinds –water, for example— whose instances typically share many common characteristics. In short, and in contrast to many classes, natural kinds appear to possess real *unity*.

---

<sup>3</sup> Or to vary the diet of examples, consider some categories from materials science –e.g. glass and silica. These are plausibly natural kinds in the domain of materials science, and yet they are vague in a variety of respects.

3. *Mind-independence*. At the very heart of the natural kind concept is the idea that some kinds are *real* kinds in nature, where ‘real’ does not merely mean that they exist, but that they are not *ideal* –i.e. that their existence is in some appropriate sense mind independent. So, for example, electrons are plausibly mind-independent in the relevant sense whilst Tuesdays are not. But *what* is this sense of mind (in)dependence? Given that we are concerned with the status of a mental phenomenon –delusion— some comments are in order, if we are to avoid confusion later on. First, some comments on what the relevant notion of mind (in)dependence is *not*:

1. One might think that the distinction is simply the distinction between *psychological* kinds and *non-psychological* kinds. But this is clearly unsatisfactory. There are long-standing disagreements about the mind-dependence of theoretical entities –e.g. quarks, electrical fields and chemical compounds— and whilst I am prepared to believe that some of the people engaged in such debates are confused, it is hard to accept that everyone is *so* confused that they are arguing over whether quarks (or chemical compounds etc.) are psychological kinds! This strains credulity. Similarly, the status of psychological kinds –beliefs, experiences and, of course, delusions — is a longstanding issue in philosophy and psychology. But by the present standard all such things are *trivially* mind dependent and, hence, non-natural kinds. So, if the present conception of mind (in)dependence were the relevant one, all such debates would be trivial and *very* easily resolvable. Again, this strains credulity.
2. Nor is the relevant distinction merely between those *entities whose existence metaphysically necessitates the existence of minds* and those that do not. Again, this would render all questions about the naturalness of psychological kinds trivial. For of course the existence of beliefs (or desires, experiences, delusions and so on) necessitates the existence of minds.
3. Nor is the relevant sense of (in)dependence mere *causal (in)dependence*. The existence of toy poodles is causally dependent on mental activity. Had breeders not made whatever decisions and judgments were required to breed this family favorite, no such kind of dog would exist. The same is true, *mutatis mutandis*, of the radioactive chemical element, californium. But despite their causal

dependence of mental activity, this in no way implies that such kinds are mind-dependent in the relevant sense.

So, what notion of mind independence *is* relevant to the characterization of natural kinds? It is what Page (2006) calls *individuated independence*. Roughly put: A kind, K, is individuatedly independent if it is circumscribed by boundaries that are totally independent of where we draw the lines. In other words, individuatedly independent kinds are the sorts of kinds whose existence does not (metaphysically) depend on how we categorize things.<sup>4</sup> Perhaps stars or oxygen atoms are individuatedly independent. At any rate, it is plausible that had we never engaged in the cognitive activity of categorizing some things as stars and others as oxygen atoms, those kinds of things could still have existed. In contrast, constellations appear to constitute an individuatedly *dependent* kind. As Page puts it:

We individuate the night sky into constellations. We, or more specifically our ancestors, determined which stars comprise which constellations. We can come up with new constellations whenever we like simply by pointing out a few stars and giving the cluster a name. Furthermore, the boundary between a constellation and its surroundings is very much a function of where we draw the lines (or more aptly, how we connect the dots). Though it is *prima facie* plausible that reality is individuated intrinsically into stars, reality is not individuated intrinsically into constellations, since it is people who divide the night sky into constellations. (Page, 2006)

The present suggestion, then, is this: Natural kinds are mind-independent in the sense that they are individuatedly independent. The NK Thesis is thus committed to the view that

---

<sup>4</sup> Even this might be too strong a condition. For example, currency is surely a kind of economics; and yet the existence of currency depends on the existence of mind that can think of some things as currency *as such*. For present purposes, however, I ignore this concern since a) the kinds of economics are contentious candidates for natural kind status for exactly this sort of reason; and b) the present condition, in fact, makes my case *harder* to sustain and not easier.

delusions are individually independent: that they are more like stars and oxygen atoms than they are like constellations.

### ***1.3 Homeostatic property clusters***

Given the discussion so far we are in a position to see that natural kinds tend to possess at least the following characteristics:

- They are discrete, though not necessarily determinate categories.
- They are mind-independent –i.e. individually independent— kinds.
- They can figure in scientific, inductive generalizations.
- They are associated with contingent clusterings of properties.
- They are empirically discoverable.
- They are associated with empirically discoverable structures, processes or mechanisms that explain the occurrence of contingent clusterings.
- They are homogeneous.<sup>5</sup>

With these characteristics in mind there is, I maintain, a plausible general characterization of natural kinds that fits well with scientific practice. The view in question –the *homeostatic cluster* account— is perhaps the most popular proposal to have emerged from the philosophy of science in recent years. Roughly put, what it maintains is that a kind, K, is natural if:

*H1.* It is associated with a contingent property cluster –a range of characteristics or symptoms, which tend to be co-instantiated by instances of the kind, but need not be genuine necessary conditions for membership.

---

<sup>5</sup> What is the status of these characteristics of natural kinds? Clearly, they are important aspects of natural kinds and their role in scientific practice. But are all (or some) of them necessary conditions on natural kindhood; or are they merely typical, though unnecessary, features of natural kinds? These are interesting questions, though not ones I address here. All I am assuming for present purposes is the following principle: All else being equal, an account of natural kinds that explains these characteristic features of natural kinds should be preferred to one that does not explain them.

*H2.* There is some set of empirically discoverable causal mechanisms, processes, structures and constraints –a *causal essence*, if you will— that causally explains the co-variation of these various symptoms.

*H3.* To the extent that there is any real definition of what it is for something to be a member of the kind, it is not the symptoms, as such, but the causal essence that defines membership. More precisely, to the extent that natural kinds have definitions, it is the presence of a causal essence producing (some of) the symptoms that comprise the property cluster that defines kind membership.<sup>6</sup>

Consider an illness such as influenza. Influenza is, on the homeostatic cluster view, a plausible candidate for natural kind status. First, it is associated with a range of characteristic symptoms –coughing, elevated body temperature, and so on—even though these symptoms do not *define* what it is to have flu. Second, there is a causal mechanism –roughly, the presence of the flu virus— whose operation explains the occurrence of the symptoms. Finally, to the extent that influenza has a definition, it is the presence of the virus –or better, the presence of the virus producing some of the symptoms— but not the symptoms as such, that make it the case that one has flu.

Notice that the homeostatic cluster view does a good job of accommodating the features of natural kinds mentioned earlier. First, on the present view, instances of a natural kind will (more-or-less by definition) tend to exhibit a contingently co-varying cluster of properties. Second, because natural kinds exhibit such property clusters, they will, as required, be able to figure in non-accidental, scientifically relevant, inductive generalizations. Third, on the present view, contingent property clusters are causally explained by empirically discoverable causal essences. Thus it is trivially the case that natural kinds are associated with mechanisms or processes that explain the occurrence of contingent clusterings. Fourth, since empirically discoverable causal mechanisms and processes are mind-independent, natural kinds will also be mind-independent and empirically discoverable. Fifth, on the present view, kinds will tend to be homogeneous in two different respects: a) The fact that the members of a natural kind tend to exhibit a

---

<sup>6</sup> See Boyd 1990, and 1991 for more extensive characterizations of the homeostatic cluster view.

contingent property cluster, means that most members of the kind will be similar in this regard. b) Because members of the kind tend to share a causal essence, there will also be significant homogeneity at the level of causal mechanism as well. Finally, on the present view, natural kinds will be discrete, though not necessarily wholly determinate. In particular, discreteness derives from the causal essence producing some aspect(s) of the contingent property cluster. What this means is that other phenomena—even phenomena that are very similar or indeed wholly overlapping—in manifest properties, will not be wholly continuous with members of the kind since they will lack the appropriate underlying causal essence.

#### **1.4. *Essentialism and Austerity about Natural Kinds***

The homeostatic cluster view does a good job of accommodating the central characteristics that categories need to possess in order to play their requisite role in scientific practice. This, I maintain, provides good reason to endorse it as an account of natural kinds—at least if by ‘natural kinds’ one is concerned with the sorts of categories that figure prominently in science. Nevertheless, some philosophers have suggested that the homeostatic cluster view is too permissive, and should be rejected in favor of some more austere view of natural kinds. Of these more austere views, the most commonly endorsed is *essentialism* about natural kinds, which imposes a range of requirements wholly lacking in the homeostatic cluster view (Ellis, 2001).

Traditional essentialism about natural kinds can be traced back to Aristotle and Locke and is suggested by the work of such 20<sup>th</sup> Century philosophers as Kripke and Putnam (Kripke 1972; Putnam 1975). It has four main tenets:

- E1.* All and only the members of a kind share a common essence.
- E2.* The essence is a property, or a set of properties, that all the members of a kind must have.
- E3.* The properties that comprise a kind’s essence are intrinsic—i.e. non-relational—properties.
- E4.* A kind’s essence causes the other properties associated with that kind.

For example, the essence of the natural kind gold is gold’s atomic structure. That atomic structure is an intrinsic property possessed by all and only pieces of gold. That structure

is a property that all gold must have as opposed to such accidental properties as being valuable to humans. Finally, the atomic structure of gold causes pieces of gold to have the properties associated with that kind, such as dissolving in certain acids and conducting electricity (Ereshefsky, forthcoming).

Traditional essentialism and the homeostatic cluster views have much in common. Both maintain that natural kinds possess *some* sort of essence; that essences causally explain the other properties associated with the kind; and that they are relevant to determining kind membership. Where they differ, however, is in their conception of essences. Traditional essentialism is committed to what have been called *sortal essences*, whereas the homeostatic cluster view is committed only to *causal essences* (Gelman and Hirschfeld, 1999). As I use the terms all sortal essences are causal essences but not *vice versa*. For in addition to figuring in causal explanations, a sortal essence is, as a matter of metaphysical necessity, possessed by all and only the member of the kind and consists of intrinsic properties. In contrast, causal essences imply no such commitments. Causal essences need not be intrinsic. They need not be possessed by all members of the kind. (There may, for example, be deviant, abnormal or borderline kind members fail to instantiate the relevant process, mechanism or structure.) And where present, it need not be metaphysically necessary that a member of the kind instantiate the causal essence. (It may, for example, only be nomologically necessary.)

Which conception of essences is most relevant to providing an account of natural kinds? Traditional essentialism –and the sortal essences it assumes— applies well to some of the categories that figure in scientific explanations. For instance, gold and other elements in the periodic table appear to conform to traditional essentialism. But if we aim to characterize the kinds that figure more generally in scientific practice, then an insistence of sortal essences is overly restrictive. One problem is that many kinds are not characterizable in terms of intrinsic properties. Biological kinds appear to be of this sort, as do some of the kinds of psychology, materials science and arguably of physics.

Even if we reject the assumption that sortal essences consist only of intrinsic properties, the requirement that each natural kind has a sortal essence is still overly restrictive. What such essentialism implies is that there are necessary and sufficient conditions for kind membership, and moreover, that these conditions are also the causally

relevant properties of kind members. But this is overly restrictive because there appear to be many kinds that figure in scientific generalization, which do not satisfy these conditions. Biological kinds are the most widely discussed example. The category of dogs, for example, appears not to be definable in terms of some set of necessary and sufficient conditions that explain the other properties associated with dogs. There seems, for example, to be no phenotypic trait possessed by all and only dogs; and nor does there seem to be a genetic ‘essence’ possessed by all and only dogs. Indeed the process of natural selection appears to work against the production of such stable sortal essences. As a consequence, many philosophers and biologists have rejected the thesis that species have essences (Ereshefsky, forthcoming). Nevertheless, biological kinds are still appropriate for the purposes of formulating robust empirical generalizations, and to that extent are natural kinds of the sort that concern us.

In addition to being overly restrictive, the sortal essentialist conception of natural kinds is unmotivated. This is because a kind does not need to possess a sortal essence in order to play their characteristic scientific roles. Consider the functions mentioned in section 1.1. First, in order to effect the distinction between kinds about which non-accidental, scientifically relevant, inductive generalizations can be formulated, and those about which few, if any such generalizations can be formulated, natural kinds need not have sortal essences. All that’s required is, as the homeostatic cluster view maintains, that kind members share many non-accidentally related –though logically unconnected— properties (or relations). Second, in order for kinds to be objects of empirical discovery, they need not have sortal essences. All that’s required is that it is an empirical matter whether such kinds can figure in robust generalizations and/or that have empirically discoverable causal mechanisms that explain the contingent property clusters associated with (prototypical) kind members. Again, the homeostatic property cluster view accommodates this without the need for sortal essences. Finally, natural kinds need not possess sortal essences in order to underwrite to project of mechanistic explanation –i.e. the search for causal essences. Now it may be that some causal essences are also sortal essences, as in the case of chemical elements. But there is no reason to suppose that this must generalize to all putative natural kinds. So, for example, to the extent that the kind *neuron* is definable by a set of necessary and sufficient conditions, it is that all and only

members of the kind are conducting cells within the nervous system. But this definition does not provide a characterization of the properties in virtue of which neurons possess their characteristic properties. In this case, then, it would seem that defining features and causal essence come apart.

So, essentialism about natural kinds is unattractive for a variety of reasons. Nevertheless, one might still be inclined to argue, on the basis of other considerations, that the homeostatic cluster view is unduly permissive. Most obviously one might claim that the account is overly permissive because it permits paradigmatic examples of non-natural kinds –most obviously, social kinds or human kinds— to be natural kinds as well. But this objection strikes me as unconvincing. First, though the account, taken in isolation, is *consistent* with the claim that social kinds are natural kinds, the same is also true of every other remotely plausible account of natural kinds as well –including the more austere ones. So, for example, taken in isolation, traditional essentialism also fails to preclude the possibility of social-cum-natural kinds. So, if the worry is a serious one, it's one that applies to accounts of natural kinds quite generally and not only to the homeostatic cluster view.

Second, whilst the homeostatic cluster view is consistent with the possibility of social-cum-natural kinds, it's also wholly consistent with their *impossibility* as well. For it may be that there are facts about each social kind –or about social kinds as such— which make it impossible for them to be natural kinds. To put the point another way: The homeostatic cluster view doesn't imply the possibility of social-cum-natural kinds. Rather, it merely fails to rule out this possibility by point of definition. Finally, the fact that an account of natural kinds fails to preclude social-cum-natural kinds as a point of definition is not grounds to reject the account. On the contrary, this is as things should be. For if natural kinds in general are, as I have argued –and virtually everyone accepts— determined *a posteriori*, then –absent some good argument to the contrary— it should also be an *a posteriori* matter whether or not *social* kinds are natural kinds.<sup>7</sup>

---

<sup>7</sup> I suppose that one might revise the complaint: The problem isn't so much that the homeostatic cluster view fails to preclude the possibility of social-cum-natural kinds but that the account –being relatively permissive—makes it too easy for social kinds to be

### ***1.5 Semantics and Natural Kinds***

We have seen that the notion of a natural kind is important to the understanding of scientific practice. But the class of natural kind *terms*, such as ‘water’, ‘gold’, and ‘tiger’ –terms which purport to refer to natural kinds— have also been the object of sustained attention in linguistics and the philosophy of language. This is a rich and complex area of enquiry that we do not have the time to discuss in detail here. Nevertheless, there are two widely accepted characteristics of natural kind terms that I need to flag since they will be relevant to the discussion later on.

First, a very widespread view about natural kind terms is that they are not synonymous with the description of (prototypical) characteristics that speakers associate with the kind (Kripke, 1972; Soames, forthcoming). For example, the term ‘water’ is associated by speakers with a description. It is the kind of stuff that boils and freezes at certain temperatures, that is clear, potable, and necessary to life, etc (Soames, 2008). Even so, the term ‘water’ is not synonymous with such a description; and we can see this because speakers are prepared to apply the predicate ‘is water’ to quantities that lack characteristics specified by the associated description.

A second and related feature of natural kind terms is that they are amenable to substantial conceptual revision. In particular, we routinely accept substantial

---

natural kinds. But we need not take this seriously. First, the complaint is *ad hoc*. The homeostatic cluster view is defended on independent grounds –viz. its power to explain the role of natural kinds in scientific practice. If it makes it easier than one expects for social kinds to be natural ones, so much the worse for the prejudice that it should be harder. Second, it’s far from obvious that it is “too easy” –whatever exactly that means— for social kinds to satisfy the conditions imposed by the homeostatic cluster view. Indeed, it is far from clear that any of the prototypical examples of social kinds –e.g. sociological groups— satisfy the relevant conditions. So, for example, if the above arguments are correct, then homeostatic clusters are both individually independent and have a causal essence: two features that prototypical social kinds are widely believed not to possess.

modifications in the extension of natural kind terms on the basis of empirical enquiry. To take a well-known example: The extension of ‘fish’ was until relatively recently taken to include whales and other marine mammals. But developments in biological systematics undermined this claim. Similarly, at one time Kant could confidently claim that gold is a yellow metal is an analytic, hence, a priori, truth. Yet in the light empirical enquiry, we now suppose that many instances of gold are not yellow –e.g. white gold—and that many samples that superficially resemble gold to a high degree are not gold at all. Precisely how best to explain such phenomena is a point of ongoing debate, which for present purposes we need not address. Nevertheless, as we will see, the existence of such phenomena will be relevant to addressing some common objections to the NK Thesis.

## 2. The Anti-Essentialist Objection

With a clearer conception of natural kinds in hand, we are now in position to consider objections to the NK Thesis. The first objection I discuss purports to show that delusions are not a natural kind because they fail to share an essence. Here is the argument in skeletal form:

1. There is no essential criterion for being a delusion
2. But natural kinds must have essences.
3. So: Delusions are not a natural kind.

According to the argument, then, the NK Thesis is false,

*Response:* Though commonplace in recent discussion (see, for example, Ghaemi, 2004; Zachar, 2000), the present argument is misguided. The main problem is with the second premise. The notion of essence in play here is clearly the notion of a *sortal* essence –one that provides necessary and sufficient for kind membership. Here’s what Ghaemi (2004) has to say on the matter:

There simply is no essential feature of delusions. We need to be clear about this fact, accept it deep in our souls, live with it, and go on from there in our work. This is, in my adaptation of Daniel Dennett’s (1995) phrase, ‘Darwin’s dangerous method.’ Darwin’s key innovation was in realizing that species are not essentialistic; there is no single aspect of the nature of species that necessarily and sufficiently characterizes them... I suggest that many notions in biology and

medicine (including psychiatry) are not essentialistic. Peter Zachar (2001) has shown, forcefully in my judgment, that psychiatric diagnoses should not be conceived as natural kinds with essential features, but rather as pragmatic kinds. ...If psychiatric diagnoses are not essentialistic, it is plausible that neither are delusions. Delusions are not characterized by any single essential feature; so much the worse for essentialism. Now let's move on. (Ghaemi, 2004)

Ghaemi's claim that many kinds don't have sortal essences—characteristics that are necessary and sufficient for kind membership—is plausible. But contrary to what he suggests, this in no way undermines the claim that delusions constitute a natural kind. As we saw in section 1, an adequate account of natural kinds need not insist on sortal essences. Rather, one need only insist on the existence of causal essences. Species and geological formations are plausibly natural kinds. But there are no known sets of necessary and sufficient conditions for many such things. *Mutatis mutandis* for delusion. Thus the mere fact that they fail to share a sortal essence is no reason to deny that they constitute a natural kind.

So, the argument fails; and not because of what it says about delusions but because it recruits a flawed conception of natural kinds. Still, the first premise of the argument deserves some comment. According to this premise, there is no essential criterion or set of criteria for being a delusion. But why accept this? To my knowledge, the most plausible reason is the dismal track record of past efforts to define 'delusion'. (For a useful review, see Garety & Hemsley, 1994). By way of illustration, consider what is arguably the most influential approach to defining 'delusion'—what we might the *standard account of delusion*. Though there are many slightly different versions of the standard account, what they share is a commitment to the idea that delusions are a species of *belief*, which possess the following characteristics:

1. *Falsity*: A delusion is a *false* belief.
2. *Entrenchment*: Delusions are *firmly held* by the patient.
3. *Doxastic Isolation*: A patient's delusion is not accepted by other members of the person's culture or subculture.
4. *Resistance to Rational Persuasion*: Delusions cannot be dispelled by argument— including *good* argument—to the contrary.

5. *Resistance to Incompatible Information*: Delusions are maintained in the face of incompatible information that is available to the patient.

Definitions of ‘delusion’ that cite these sorts of characteristics are widespread in psychiatry and clinical psychology.<sup>8</sup> Nevertheless, as an attempt to specify necessary and sufficient conditions for being a delusion, the standard account is unsatisfactory. First, many –perhaps all—the above are not necessary conditions on being a delusion. So, for example:

- Though delusions are very typically false, they are not invariably so. For example, there are reported cases of hypochondriacal delusion with the content *I am mentally ill*. The belief is true, but for all that it’s still a delusion. (See Fulford, 1994 for further details).
- Delusions are not always doxastically isolated. Sometimes –as in the case of folie à deux —the very same delusion can be held by multiple individuals.

Similar points can be made about the other criteria.

It’s also far from clear that the conditions specified by the standard account are jointly *sufficient* for delusion. By way of illustration, consider the case of a rather stubborn professor who is deeply invested –personally and professionally— in a false, pet theory that they alone defend. Philosophy is, I suspect, replete with such figures! Suppose further that the pet theory was arrived at by faulty reasoning, and maintained in the face of strong argument and evidence to the contrary. Perhaps we have a Lewis-like figure who maintains that there exist an uncountable infinity of concrete universes (Lewis, 1986); or an Unger-ish professor who maintains that they, themselves, do not exist

---

<sup>8</sup> The Standard account is nicely illustrated by the definition of ‘delusion’ in DSM IV: A delusion is a “false belief based on incorrect inference about external reality that is firmly sustained despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary. The belief is not ordinarily one accepted by other members of the person’s culture or subculture (e.g. it is not an article of religious faith)” (DSM IV, p. 821).

(Unger, 1979).<sup>9</sup> Even so, no psychiatrist would return a diagnosis of psychosis. Pet theories –even false, implausible and singularly held theories—are part and parcel of contemporary academia. The standard conditions on delusion are met, but the attribution of delusion withheld. It would seem, then, that the standard account provides neither necessary nor sufficient conditions for delusion.<sup>10</sup>

Let's return to the anti-essentialist objection. Earlier we saw that the argument fails because it incorrectly assumes that the instances of a natural kind must share a sortal essence. We have now seen that the most widespread approach to defining delusion – the standard account— is unsatisfactory: that it fails to specify necessary and sufficient conditions for delusion. But on reflection this should be unsurprising, if the NK Thesis is true. What the NK Thesis is committed to is that an underlying mechanism explains the typical co-variation of a property cluster. Moreover, to the extent that anything is determinative of kind membership it is –as in the case of water— the instantiation of a causal essence. But the standard account clearly does not specify any such causal essence. Rather it characterizes delusion in terms of a cluster of superficial properties that delusions often possess but, as we have seen, need not invariably possess. On the

---

<sup>9</sup> Of course, I do not mean to suggest that the real Lewis or Unger resemble deluded subjects. I would not dream of impugning the quality of reasoning found in the work of these excellent philosophers!

<sup>10</sup> The point is not, of course, that there are no differences between stubborn professors and delusional patients. Rather, the point is that the standard account fails to mark such a distinction. Of course, one might try to shore up the standard account by adding further conditions. But this has been tried without much success. So, for example, one might suggest that the deluded subject, in contrast to the Ungerish academic, *lives* their beliefs – that they are rationally and coherently interfaced with their actions. But the problem with adding this condition to the standard account is that it's not a necessary one. Whilst some deluded individuals live their beliefs, not all do. (For further discussion see Buchanan & Wellesley, 2004.)

assumption that delusion is a *bona fide* natural kind, then, it should be unsurprising that the standard account fails to provide an accurate account of the kind.<sup>11</sup>

### 3. Some Provisional, Positive Morals

Does this mean that if the NK Thesis is correct, then the standard account has no bearing whatsoever on a theory of delusion? I think not. Indeed our discussion of the standard account yields a number of positive morals.

*Positive Moral # 1: The standard account specifies (part of) the property-cluster associated with delusions.* As already noted, natural kinds have an associated syndrome: a cluster of properties that tend to co-vary with each other despite being neither logically related nor definitive of the kind. But if this is so, then delusions must have an associated syndrome if they are to constitute a natural kind. What might this syndrome be? An obvious and plausible suggestion is that the characteristics identified by the standard account at least partially specify the relevant syndrome.<sup>12</sup> That is: delusions tend to possess –though need not invariably possess— the properties of falsity, entrenchment, doxastic isolation, resistance to rational persuasion and resistance to available incompatible information. On this view, then, the error of the standard account is not that the features it specifies are irrelevant to an account of delusions but that the logical character of these features has been misconstrued. They are not conceptually or metaphysically necessary conditions for being a delusion, but aspects of the syndrome associated with delusions.

*Positive Moral # 2: We have prima facie reason to adopt the default assumption that delusion constitutes a natural kind.* If delusions are a natural kind for which we currently possess no good theory, then we should expect that scientific and medical

---

<sup>11</sup> A related point: The standard account is perhaps best viewed as an attempt to provide a nominal definition –an account of what our word ‘delusion’ means—as opposed to a real definition. But due to space limitations, I propose to put this issue to one side.

<sup>12</sup> It’s worth stressing that according to a natural kinds conception of delusions, there may of course be other –empirically discoverable—properties in the cluster. I discuss some candidates later in the chapter.

practices concerning delusion would follow the pattern characteristic of other examples of natural kinds. This expectation appears to be met in at least the following respects:

1. *Covariation*: Natural kinds possess an associated cluster of properties that whilst logically independent and empirically dissociable very often co-vary. As I argued earlier this seems to be true of delusions. Although the criteria identified by the standard account are logically independent and empirically dissociable, they often co-vary with each other. Indeed, they tend to co-vary in precisely those cases that clinicians and psychopathologists construe as paradigmatic examples of delusion. It would seem, then, that delusion, like many other putative candidates for natural kind status –water, dogs, chlorine, and so on—possess an appropriate cluster of associated properties.
2. *Anti-descriptivist Semantics*: As mentioned earlier, a very widespread view about natural kind terms is that they are not synonymous with the description of (prototypical) characteristics that speakers associate with the kind (Kripke, 1972; Soames, 2008). For example, the term ‘water’ is associated by speakers with a description. It is the kind of stuff that boils and freezes at certain temperatures, that is clear, potable, and necessary to life, etc (Soames, 2008). Even so, the term ‘water’ is not synonymous to such a description; and we can see this because speakers are prepared to apply the predicate ‘is water’ to quantities that lack characteristics specified by the associated description. What of ‘delusion’? Given what I have said so far, it is plausible that ‘delusion’ is relevantly similar to ‘water’ in the above respects. For as we have seen, it is not synonymous with the description that clinicians associate with delusion –i.e. the description specified by the standard account. Moreover, as in the case of water, this is clear from the fact that, as indicated in section 2, clinicians are prepared to apply the term to phenomena that fail to satisfy the description. In short: The term ‘delusion’ possesses precisely the semantic character we should expect it to exhibit on the assumption that it is a natural kind term.
3. *Reliable Tracking*: In the history of science many of the most plausible candidates for natural kind status –water, gold, lead, dogs, monkeys, etc.— have paradigmatic instances that are reliably –though not, of course, invariably— identifiable, even in

the absence of any good theory or definition for the kind. There is an obvious reason for this. Good theories and definitions are not the starting point for science but the *end product*. In which case, if, as we suppose, natural kinds, are the sorts of things whose causal essences are empirically discoverable by the scientific community, they must be kinds whose instances can be reliably tracked by the community *prior* to possession of a good theory. Notice that this implies not merely that some members of the relevant community can on some occasions identify instances of the kind, but also that there is a high degree of inter-rater reliability regarding which things are instances of the kind. Absent such reliability, the standard methods of science simply won't work. *Prima facie*, this point applies to delusion. As many theorists have lamented, no good theory or definition of what delusions are currently exists. But despite this, there seems to be high levels of inter-rater reliability within the relevant scientific communities –e.g. psychiatry and clinical psychology— when it comes to identifying instances of delusion (Bell, et al. 2006). It is not merely that psychiatrists, when exercising their clinical judgment, tend to agree. Rather, it seems that quite different methods for assessing delusion –e.g. clinical interview and various standardized scales— reliably converge (Bell, et al. 2006). In short: Assessments of delusion appear to exhibit precisely the sort of reliability that we'd expect on the assumption that they constitute a natural kind.

4. *Empirical Regularities*: Let me add one final consideration that is largely independent of the discussion so far. If the NK Thesis is true, then we should expect appropriate methods of enquiry to yield a body of empirical regularities concerning delusions. Though it is too early to tell with any certainty, there are grounds for optimism on this score. For over the past two decades or so, a wide range of results has emerged regarding delusions. So, for example, there is now considerable evidence of various abnormalities in the reasoning, attention, metacognition and attributional tendencies of delusional patients (Bell, et al., 2006). Some of these phenomena will be discussed briefly in section 6; though for reviews of some relevant results see Bell, et al. (2006); Garety & Hemsley (1994); Garety et al. (2001); Freeman and Garety (2006); and Freeman (2007).

What follows from the above? Though such considerations clearly fail to provide strong grounds for accepting the NK Thesis, I maintain that they do give us *prima facie* reason to take the claim seriously. After all, the scientific practices surrounding delusion appear to exhibit precisely those characteristics that we should expect them to exhibit if the NK Thesis were true. To put the point another way: Though such grounds for accepting the NK Thesis are eminently defeasible, they still provide reason to adopt it as a working hypothesis. In view of this, one question that requires our attention is this: Are there any good reasons for rejecting the *prima facie* plausible assumption that delusions are a natural kind? Most of the remainder of the chapter will be concerned with this issue.

#### **4. Continuity Objections**

So far our discussion has proceeded under the assumption that delusions are a kind of some sort, even if they do not constitute a natural kind. But in fact one prominent kind of objection to the NK Thesis purports to undermine even this weaker assumption. According to such objections delusions do not constitute a kind of any sort, natural or otherwise.

##### **4.1. Vagueness and delusions**

Let's start by disposing of a red herring. One might be inclined to argue that delusions are not a genuine kind, natural or otherwise, on the grounds that the category of delusion has vague boundaries –roughly put, there are entities for which it is indeterminate whether not they are delusions. But as we saw in section 1, mere vagueness does not suffice for rejecting the existence of a kind. Children really do constitute a kind even though it is a kind that lacks precise borders. Nor is vagueness a reason to deny that a kind is a genuine natural kind. On the contrary, many plausible candidates for natural kind status have borderline cases. Eyes, for example, plausibly constitute a natural kind; and yet it does not have determinate boundaries. Consider, for example, a structure that biologists call 'pit-eyes': simple eye-spots of approximately 100 cells, set into a pit to reduce the angles of light that enters the eye-spot. Are such structures *eyes*? It is, I take it, utterly unclear how to categorize such a case. And if you think they are clearly

categorizable as eyes, then what of eye-spots not located in pits? So far as we know, there are no clear answers regarding how best to classify such examples.

#### ***4.2. Continuity with Normal Experience***

So, mere vagueness or indeterminacy does not suffice to undermine the NK Thesis. It is simply beside the point. But another more interesting line of argument is that delusions do not constitute a kind because they are literally *continuous* with other non-delusional phenomena (Strauss, 1969; van Os, 2003; Freeman & Garety, 2006). On this *continuity view*, the right way of thinking about delusions is as a range on a continuum for which there exist no discrete boundaries at all, vague or precise. On such a view, the complaint against the NK Thesis is not merely that delusions comprise a kind with vague boundaries but that there are no well-motivated points at which to draw the distinction between delusions and other phenomena. And since kinds are supposed to be discrete this would suffice to undermine the NK Thesis.

Though the present objection comes in a variety of forms, the most plausible and popular version seeks to defend a continuum account of delusions on the grounds that they are continuous with normal, non-pathological experience (van Os, 2003, Freeman & Garety, 2006).<sup>13</sup> But what is the evidence for such continuity supposed to be? Much of the putative evidence comes from epidemiological studies. So, for example, in one well-known study involving over 7000 people in the Netherlands, it was found that 3.3% of the population had a ‘true’ psychiatrist-rated delusion whilst 8.7% had a ‘not clinically relevant’ delusion (van Os et al., 2000; Johns & van Os, 2001). Similarly, in a study on delusional ideation in primary care patients it was found that delusional ideation was frequent in those with no psychiatric history. For example, 10% of participants with no

---

<sup>13</sup> It has also been suggested, though never presented as an objection to the NK Thesis, that delusions are hard to discriminate from other pathological phenomena –e.g. over-valued ideas (David, 1999). Nevertheless, there is good reason to suppose that the patterns of characteristics associated with delusion are readily distinguishable from those associated with over-valued ideas (Jones, 1999).

psychiatric history reported that they had at some point felt as if there was a conspiracy against them (Verdoux et al., 1998).

Such data are interesting, perhaps even surprising. But it in no way adjudicates between kind and continuity conceptions of delusion. In order to see this, we need to look more closely at the character of the inference from the data. What the epidemiological data strongly suggests is that delusional phenomena are prevalent not merely among those with a psychiatric history but also amongst the population at large. But how does this cut in favor of a continuum conception of delusions? The idea, in brief, is that it supports a continuum hypothesis because according to such a view ‘psychotic symptoms should be present not only in subjects identified as “cases of psychosis”, but also in a proportion of subjects from the general population that does not fulfill the clinical criteria of “case of psychosis”’ (Verdoux & van Os, 2002). In contrast, a kinds conception of delusion makes no such prediction. So, the continuum hypothesis wins.

The present issues and evidence are important and intriguing. Nevertheless, as an argument against the NK Thesis it is far from overwhelming. First, the present data are wholly *compatible* with a kinds conception of delusions, including the NK Thesis itself. In order to see this, consider the following –mutually compatible— possibilities. First, it is quite possible that psychotic pathology is more widespread than initially supposed. That is, what the data may in part reflect is the (rather unsurprising) fact that the general population contains people with psychotic symptoms who have not (as yet) acquired a documented psychiatric history.<sup>14</sup>

Second, the claim that delusions constitute a natural kind in no implies that delusions are *only* present in pathological cases. Instead it may be that someone can be both deluded and not mentally disordered. By widespread consensus, having a mental disorder or psychopathology presupposes that some *evaluative* condition is met. So, for example, according to one prominent view, mental illness or mental disorder require that the state is in some way *harmful* (Wakefield, 1992), But notice, if delusions are a natural

---

<sup>14</sup> As Matthew Broome has pointed out to me, the general population clearly contains such people. After all, on first becoming ill a person will lack a psychiatric history even though they still have a clear-cut psychotic illness.

kind, then this creates the possibility of some –indeed many— delusions that are not harmful, either to the deluded subject, or to anyone else for that matter. That is, there may be people who meet the following conditions:

- They exhibit the characteristic pattern of properties associated with delusions,
- The presence of the property cluster is explained by the activity of some mechanism, which in the case of delusions, is responsible for the covariation of the property cluster, but
- There are no harmful effects either for the individual or for those around them.

Under these conditions, someone would have a genuine delusion and yet fail to be mentally disordered.

Finally, it is important to be clear on the sort of continuity that the available evidence supports. At most, what the data suggest is a *symptomatic* continuum: nonpathological subjects –and those with mood disorders— exhibit delusion-like states; states that possess properties that are very similar to, or overlap with, those associated with the psychotic delusions (Verdoux & van Os, 2002). But the existence of such symptomatic continuity is wholly consistent with the NK Thesis. Recall: According to the NK Thesis, symptoms, as such, are not definitional of having a delusion. Rather it is the instantiation of the salient causal essence that determines category inclusion. But what this means is that there can be many phenomena that are superficially very similar to delusion and yet are not delusions because the manifestation of the symptoms fails to depend on the appropriate kind of underlying causal mechanism. The point is in no way specific to delusions. It applies quite generally to natural kinds. So, for example, influenza is symptomatically quite similar to many other illnesses. For example, there are many diseases that characteristically involve an elevated body temperature, congestion, sore throat, and so on. But that doesn't mean that a continuum theory of influenza should be adopted. Rather, such continuity in symptomology between flu-suffers and others is to be explained by the fact many mechanisms can produce flu-like symptoms. Despite this, only those people who exhibit such symptoms as a consequence of the activity of a flu virus have flu.

So, it would seem that the current evidence for continuity is compatible with the NK thesis. Nonetheless, one might think that a continuum view is at an advantage because it

predicts the epidemiological data, whereas a kinds view of delusion does not. But in fact the continuum hypothesis does not do such a good job of predicting the extant data. What the data suggest is not merely that: Delusion-like symptoms are present in a proportion of subjects from the general population. It also suggests that: ‘True’ psychiatrist-rated delusions are present in the general population. The continuum model predicts the former fact. But it does not predict the latter. At best it is silent on the matter. Indeed one might also think that it fits most naturally with the converse prediction. For on perhaps the most natural way of formulating the continuum hypothesis, it should in fact be *surprising* that non-pathological populations manifest ‘true’ delusions. Specifically, what the continuum approach is widely thought to imply is that ‘delusions are not qualitatively different from normal beliefs, but simply represents a more extreme end of the population spectrum or distribution of anomalous mental phenomena’ (Bell et al. 2006). But if this is so –if delusions are just the tail of the curve—then one would expect ‘true’ delusions to be those found amongst those people with the most anomalous mental lives –i.e. those who comprise the clinical population.

Still, doesn’t the continuum approach have the advantage of predicting the fact that delusion-like phenomena are found in the population at large? Yes. But it is only a very slight advantage. The reason is that the NK Thesis, when combined with additional plausible assumptions, also makes the same prediction. Earlier, I commented that many different putative natural kinds bear striking symptomatic similarities to other kinds. The symptoms of influenza are remarkably similar to those of many other illnesses. Likewise, the superficial properties of gold are very similar to those of various compounds. Jadeite and nephrite are superficially very similar. And so on. Now the mere fact that these things are natural kinds in no way *implies* that such syndrome overlaps exist. But the fact that *so many* natural kinds exhibit such overlaps is inductive grounds for supposing that delusions would exhibit symptomatic continuities and overlaps, if the NK Thesis were true. Thus the NK Thesis plus this apparent fact about natural kinds renders symptomatic continuity unsurprising, if not probable.

### ***5. Mind Dependence Objections***

Let's turn to a different sort of objection to the NK Thesis: one that purports to put pressure on the assumption that delusions constitute a mind-independent kind. Though the objection comes in a number of forms, in what follows I consider what I take to be the three most plausible versions.

### ***5.1 Delusion as a folk psychological kind***

According to one version of the objection, delusions are not a natural kind but an artifact of our folk psychology: our commonsense mode of thought about the mental. Murphy (2006) expresses the position as well as anyone:

[W]hether or not something is a delusion is a matter of how it strikes us, and that depends on how well it comports with our understanding of what people are like, both in general terms and within our culture. It does not depend on some psychological mechanism or a formal property of beliefs. Even if we can identify belief formation mechanisms, it is unlikely that any mechanism has *completely* broken down even when subjects are delusional, since even in the maddest psychotics we find some preservation of normal reasoning alongside the delusional reasoning. (Murphy, 2006, p.180)

Murphy is not claiming that delusions are not a natural kind merely because they are a part of our folk conception of the world. This would be an obviously poor argument since there is no incompatibility between the naturalness of a kind and it being a kind identified in our folk conception. Water is plausibly a natural kind, though 'water' and the concept it expresses are part of commonsense. Similarly, species terms and the concepts they express are thoroughly folkish, but that alone doesn't mean that dogs or cats don't constitute natural kinds. Rather, it might be—and plausibly is the case—that some of our folk concepts pick out natural kinds.

Why, then, would one think that the present consideration threatens the NK Thesis? According to the present view, it is because *what it is to be* a delusion is determined by how it strikes us. That is, being a delusion is a *response-dependent* property. In the present case, the claim is that how something strikes beings like us—beings with our commonsense psychology—is metaphysically determinative of what it is to be a delusion. But if this is so, then delusions won't be a natural kind. Natural kinds

are supposed to be discoverable, mind-independent kinds whose existence is not parasitic on our taxonomic practices. But according to the present view, being a delusion is metaphysically determined by our folk psychological classificatory practices. In which case, delusions will be mind-dependent in a way that precludes them from constituting a natural kind.

*Response:* The present line of reasoning conflates the metaphysics of delusion with its epistemology. The relevant metaphysical issue concerns the nature of delusions: roughly, what is it to be a delusion. The relevant epistemic question concerns the evidential basis for our judgments about delusion: roughly, the sorts of evidence we invoke in judging that someone is deluded. What's right about the present objection is that our judgments about delusion are not based on some scientific account of either the mechanisms on which delusions depend or the formal properties of belief. Instead, the judgments of clinicians depend largely on the sorts of evidence that are available on the basis of commonsense psychological considerations alone –e.g. the reasonableness of the belief, its failure to conform to cultural norms, the extent to which it is resistant to countervailing evidence, and so on. But it clearly doesn't follow from this epistemic point alone that the *nature* of delusion is exhausted by how things strike us. On the contrary, it is wholly compatible with a conception of delusions as a natural kind. Compare: Our standard methods of determining whether something is water are made on the basis of considerations drawn from folk wisdom. Prior to developments in 18<sup>th</sup> Century chemistry no one ever judged something to be water on the basis of anything other than such folkish considerations. But it clearly doesn't follow from this alone that water is not a natural kind. For all Murphy's argument shows, the same is true of delusions as well. What's required is some reason to suppose that in the case of delusion things really are different: that issues about the nature of the kind, and issues about how we know things about instances of the kind *ought* to be collapsed. But no such argument is forthcoming; and in the absence of such an argument, the present considerations simply beg the question against the NK Thesis.

## 5.2. *The cultural relativity of delusion*

Still, there are various ways one might try to develop the view that delusions are response-dependent in a way that threatens the NK Thesis. One obvious strategy turns on the widely noted fact that the content of delusions is highly sensitive to social context. Indeed, the diagnosis of delusions often turns heavily on whether the patient's belief is accepted by other members of their culture or subculture. By way of illustration, consider the case of religious delusions. If I believe that there is an all-powerful deity, I do not have a delusion. I merely subscribe to a widely held, socially sanctioned religious viewpoint. If I believe, however, that *I* am God, then I will likely be judged delusional. What makes for the difference? We might suppose that both beliefs are false, fixed, incompatible with the available evidence and highly resistant to counter-argument. But what seems to make the difference is that in one case the belief is socially sanctioned (within a particular culture) whilst in the other case it is not. Delusions would thus appear to be response-dependent at least to the extent that what it is to be a delusion depends on what beliefs are socially prevalent. In which case, it would seem that the NK Thesis is false.

*Response.* What the sensitivity of delusions to social context shows is that the nature of delusion, as such, cannot be characterized in terms of their contents. But the fact that what counts as a delusion varies from culture to culture need not undermine the natural kind status of delusions so long as there is some culturally *invariant* way to characterize the nature of this cultural sensitivity. Fortunately, there seems to be just such a characterization. In brief, what the cultural sensitivity of delusions appears to track is the insensitivity of delusions to an important source of epistemic warrant and epistemic defeat: *testimony*. Many of the beliefs that we adopt are held on the basis of testimony. That is, we adopt beliefs on the basis of reasons, evidence or information provided by (putatively) authoritative sources. So, for example, we accept much on the basis of reading the newspaper, or consulting encyclopedias, or talking with local experts, or on the basis of the prevailing 'commonsense' or popular consensus—itself a species of testimony. Much the same applies to religious belief. When I believe that there is an all-powerful God, I do so in part on the basis of widespread access to testimony: (putative) experts—e.g. priests and rabbis—television shows, popular opinion, and so on. The belief

may well be false, and such testimony may ultimately be subject to defeaters. But there is little doubt that testimony is a genuine source of warrant; and there is little doubt that in societies where theism is widespread, many such lines of testimony exist, and most of us are exposed to it from an early age. This contrasts strongly with the case of religious delusions –such as the belief that I am God. Here there are no such lines of testimony since no one else holds the belief. Moreover, if testimony is to guide belief, then such a delusional belief would likely be defeated many times over by the testimony of others since no-one else accepts the delusion as true. Of course, none of this makes much difference to the patient with a religious delusion. They are, it would seem, resistant to testimony, at least where the delusion itself is concerned. So the suggestion is this: The cultural relativity of delusions is not of the kind that threatens the NK Thesis. What it shows is that delusions are resistant to testimony –a culturally invariant feature of delusion— and not that being a delusion is a response-dependent property.

### ***5.3. The normativity of delusion***

Let's now turn to what I take to be the most challenging version of the mind-dependence objection: one that focuses on the *normative* character of delusions. Delusions are the sorts of states that are normatively assessable. Indeed, one might suppose that they are *essentially*, negatively assessable. When one is deluded there is necessarily something *wrong*. But the objection continues: the relevant norms are *social* norms. They are norms that depend on our cultural modes of thought. In which case, the existence of delusions, as such, is essentially dependent on cultural modes of thought. And if this is so, then delusions are not individually independent kinds and, hence, not natural kinds at all.

*Response:* For the present objection to the NK Thesis to succeed, it needs to be the case both that some normative conditions are essential to being a delusion *and* that these very same normative conditions are social/cultural ones whose existence depends on our modes of thinking. But we have little reason to suppose that this is so. Though delusions are plausibly subject to *some* social norms, these social norms are not essential to being a delusion; and though there are some normative conditions that are plausibly essential to being a delusion, these norms are not plausibly social. Let me explain.

In order to appreciate the situation we need first to consider the question: What norms are delusions subject to? There are two plausible candidates: Medical norms and norms of rationality. Medical norms are those that determine whether or not something is pathological: a disease, illness or disorder. Norms of rationality are those norms that determine whether or not something is rational or reasonable. Let's consider these types of norm in turn.

Medical norms determine whether or not a given state is pathological: whether or not there is something medically *wrong* with the organism. Precisely how best to characterize such norms is, I suspect, a largely open matter; and it may be that they are not a social norms at all. Even so, there is widespread consensus amongst those interested in the notions of disease and disorder that such norms are at least partially social in character. So, for example, even Jerome Wakefield –whose *harmful dysfunction* account of disorder is amongst the least socially-laden of extant proposals— readily accepts that the notion of harm should be understood in socio-cultural terms. Roughly, the relevant notion of harm is to be characterized in terms of a reduction in well being, as defined by social values and meanings.

Suppose that this is true. Medical norms are at least partially social norms. Does this pose a problem for the NK Thesis? Only if delusions are *necessarily* harmful or otherwise problematic relative to the relevant medical norms. But why suppose that this is true? As far as I can tell, there is no good reason to suppose that it is. Indeed, if our earlier discussion of the continuity objection ran along the right lines, then it seems quite possible for delusions not to be pathological. In saying this, I do not deny (of course) that delusions are *very typically* pathological. Rather the point is a modal one: Though delusions are very often pathological –we might even suppose that they are in fact invariably pathological— the connection between pathology and delusion is, for all that's been said so far, a contingent one. And absent some reason to suppose that the connection is a necessary one, the (putative) social character of such norms poses no threat to the NK Thesis.

Let's now consider the relationship between delusion and norms of rationality. First: Are delusions essentially irrational or unreasonable? How best to answer this question far from clear. To be sure, delusions are *very typically* irrational or

unreasonable. We might even suppose that they are *invariably* irrational. But whether to treat the connection as a necessary one –as opposed to merely a contingent, though (highly) reliable one— is not at all obvious. Perhaps the initial description that we associate with delusions –and use to identify them— partially characterizes them as irrational. This would, in fact, appear to be the case –at least to the extent that the standard account of delusion is the description we associate with delusions. For it mentions two properties that impugn the (epistemic) rationality of delusions: their resistance to rational persuasion, and their resistance to incompatible information. But, to repeat: the fact that irrationality is part of the description that we associate with (prototypical) delusions does not imply that irrationality is an *essential* property of delusions.

Still, let us suppose for the sake of argument that delusions are necessarily irrational or unreasonable. Would this pose a problem for the NK Thesis? Not unless the relevant rational norms are *social* norms. The problem for the present version of the mind-dependence objection is that a social constructionist view of rational norms really isn't plausible. This is a familiar topic that has been the subject of longstanding discussion; and we do not have time to discuss it in detail here. But let me mention just two considerations. First, a social norms conception of rationality appears to make it metaphysically impossible for a person to rationally criticize the prevailing epistemic and practical norms of their own society. But this doesn't seem at all impossible (Gibbard, 1990). Second, of the many extant theories of rationality currently entertained by philosophers, the vast proportion articulate rational norms in a way that does not reduce them to social norms. Consider, for example, Bayesian accounts of rationality (Howson & Urbach, 1993) or reliabilist accounts (Nozick, 1993), or pragmatic accounts (Stich, 1990). All render rationality in a manner that is not social in any sense that threatens the NK Thesis. But if this is so, then the proponent of the present objection to the NK Thesis needs to show that a social conception of rationality is preferable to such alternatives. I am pessimistic that such a position could be sustained.

## **6. Heterogeneity Objections**

As we saw earlier, if the NK Thesis is true, there must be some underlying causal essence –e.g. mechanism, process or structure—that explains the symptom cluster associated with delusions. In this section, I consider a range of objections that challenge this requirement on the grounds that there is no sufficiently heterogeneous kind of mechanism or process that can play this role.

### **6.1 Causal Heterogeneity**

A common view about delusions –often associated with the so-called biopsychosocial model—is that ‘many factors are implicated in delusion development, and the contribution of each in individual cases varies’ (Freeman & Garety, 2006). From this it is common to conclude that they are unlikely to admit of unified explanation (ibid.); and from this, it may seem like only a small step to the conclusion that delusions are not a natural kind. For what the NK Thesis demands is that delusions admit of some kind of explanatory unity.

But appearances are, at least on this occasion, misleading. The claim that something is an instance of a natural kind is wholly consistent with its being produced and maintained by *many* different causal factors. To take an obvious example, the development of an organism –a tiger, for example—depends on a huge array of causal factors: gravitational forces, appropriate levels of nutrients and oxygen, thermal conditions, the occurrence of appropriate patterns of gene expression and regulation, the presence of chemical gradients within the growing organism, and so on. The factors relevant to maintaining a (live) tiger are similarly numerous. But it surely does not follow from this alone that tigers are not a natural kind.

Why not? There are at least two reasons. First, in order to understand the role that the identification and citation of causal essences plays in scientific research, we need to respect the distinction between background causal conditions and explanatorily salient causal mechanism. There is nothing here that is peculiar to the case of delusion. All causal explanation presupposes such a distinction. So, for example, when I kick a ball, there are many enabling conditions that must obtain in order that my action succeeds. Oxygen must be present, gravitation must be in force, appropriate levels of ambient light

must be available so that I can see the ball, and so on. Moreover, there are many prior conditions that must hold in order that I can kick the ball. I must have been born, life on earth must have evolved, the planet itself must have developed out of cosmic particles, and so on.<sup>15</sup> Nevertheless, a causal *explanation* of ball kicking need not cite all these factors. Rather they are treated as background conditions that need to obtain, but are not central to explaining the specific phenomenon. The same is true of delusions. Though there may be many factors that contribute to their occurrence, this alone is not incompatible with the idea that they are all produced by some kind of mechanism since most of the causal factors responsible for delusions may be part of the background conditions as opposed to aspects of the mechanism itself.

Second, it is important to distinguish between different *levels* of explanation. The identification of a natural kind is typically made relative to a particular level of explanation. Sodium is a kind of chemistry. Quark is a kind of physics. Neuron is a kind of biology, and so on. The mere fact that many different causes of delusion poses no problem for the NK Thesis if they are causes that obtain at different levels of explanation. This is because the causes need not be *distinct* causal factors so much as factors that bear various dependency relations to each other. So, for example, the fact that my behavior can correctly be said to be caused both by my mental states and by neural states, is not by itself reason to suppose that my behavior is caused by a multiplicity of *distinct* causal factors since in this case presumably mental activity is not independent of neural activity but rather metaphysically depends on —e.g. is identical to, realized by or supervenes on— neural activity. In such cases, higher and lower level causes don't compete: they are not evidence of heterogeneity but of dependency.

## ***6.2 Neural Heterogeneity***

So the mere fact that delusions have many different causes in no way undermines the claim that delusions are a natural kind. Nevertheless, the NK Thesis is committed to the claim that delusions depend on the activity of some kind of mechanism (process or

---

<sup>15</sup> Indeed, on the assumption that causal is transitive every causal relation that link my kicking the ball to the big bang will be *a* cause of my kicking the ball!

structure, etc.). And this leads to a new version of the heterogeneity argument. Since mental phenomena in general depend on neural activity, presumably delusions, qua mental phenomena, depend on neural mechanisms (process or structure, etc.). In which case, it may seem that the NK Thesis mandates that there exist some kind of neural mechanism on which delusions depend. But this appears to pose a problem. For what neuroscientific research suggests is that different delusions depend on different neural states, processes and mechanisms. For example, both the neuroanatomy and neurochemistry of Capgras delusion appear quite different from the neuroanatomy and neurochemistry of persecutory delusions or passivity phenomena. But if this is so, then it seems implausible to claim that the same neurobiological mechanism is responsible for all delusions. In which case, the NK Thesis would appear to be false.

*Response:* So formulated, there is a general, and I think, plausible response to the present objection. Indeed the response falls out of one of the most widespread (and plausible) assumptions of contemporary cognitive science and philosophy of mind: the *multiple realizability of the mental*. In brief, according to this thesis, psychological kinds (such as, pain and belief) can be realized by many distinct physical kinds. Most importantly for our purposes, multiple realizability implies that a psychological kind can be realized by many different neural kinds. In such cases, there need be no one-one relation between a psychological kind and an underlying neural kind. Instead, the psychological kind will bear a *one-many* relation to neural kinds. Or to put the point slightly more precisely, a psychological kind that is multiply, neurally realized, will participate in a *one-to-many* dependency relation with neural kinds, where these kinds are characterized in the terms of of *biological neuroscience*: roughly, the chemistry, physiology, morphology and anatomy of neurons and neuronal assemblies (Stoljar & Gold, 1998).<sup>16</sup>

The above characterization of multiple realizability, suggests a response to the present objection. If delusions are multiply realizable, then different delusions can be

---

<sup>16</sup> Multiple realizability has been the subject of intensive philosophical attention over the past three decades. For more detailed discussions see Block, 1997; Kim, 1992; Pylyshyn, 1984; and Shapiro, 2000.

realized by quite different kinds of neural state and process. In which case, it should be wholly unsurprising that neuroscience fails to identify a neat one-one correspondence between delusions and neurobiology. Instead we should expect the emerging picture to be a complex one marked by a one-many relation between delusion and underlying neurobiology. But this alone doesn't show that delusions are not a natural kind of any sort. All it shows is that they are not a *neurobiological* natural kind –i.e. a kind that is (non-disjunctively) characterizable in the vocabulary of biological neuroscience. In which case, what follows from the NK Thesis is that delusion will be a natural kind relative to some *other* level of explanation in much the same way that quarks and species are kinds relative to levels of explanation other than biological neuroscience.

But what might this level be? In the case of quarks and species, we seem to know the answer –fundamental physics and evolutionary biology, respectively. But what is the right level for the purposes of delusion? There is an obvious suggestion that surely merits further exploration. Historically, the multiple realizability of the mental was widely (and correctly) taken to suggest that a science of the mind requires levels of description and explanation more abstract than those afforded by biological neuroscience. Specifically, the multiple realizability thesis was central to motivating the view that psychological science requires computational/functional levels of description. Such levels of description have been central to contemporary cognitive science and cognitive neuroscience, among other reasons, because they permit stable mechanistic descriptions that abstract away from neurobiological heterogeneity. Such descriptions have proven hugely productive in understanding myriad psychological phenomena. Indeed, there has in recent years been considerable effort to apply this approach to the study of delusions. (See, for example, Bell, et al, 2006; Birchwood et al., 2000; and Freeman, et al, 2002). This leads to a further positive moral.

*Positive Moral # 3: Delusions as a cognitive kind.* If the NK Thesis is true, then delusions are a multiply neurally realized kind. Moreover, given extant efforts to understand psychological phenomena that are multiply realized, it is plausible to adopt the working assumption that the relevant level of description is a computational-cum-functional one.

In short: The most plausible version of the NK Thesis is one that views delusions are a *cognitive kind*.

### 6.3 Cognitive Heterogeneity

But there is a reformulation the previous objection that requires our attention. According to the multiple realizability thesis, mental states of the same type—in the present case individual delusions—are realized by different neural kinds, where those kinds are individuated via the traditional methods of neurobiology—e.g. chemistry, anatomy and physiology. But in the case of delusions it seems not merely that delusions are multiply realized in this sense. Rather it seems that they also depend on different kinds of *psychological* states and processes—where such types are individuated by their functional-cum-computational properties. So, for example, according to one familiar view—one that clearly will not generalize to all delusions—Capgras delusion depends crucially on states of the face recognition system (Ellis & Lewis, 2001). Similarly, according to a prominent account of passivity phenomena—again, an account that clearly won't generalize to all delusions—delusions of control depend crucially on disruption to forward models within our motor systems (Blakemore et al. 2003). For present purposes, the crucial point is that these differences are not merely neurobiological but also involve functional-cum-representational differences between the kinds on which delusions are supposed depend. As a consequence, it would seem that we not only have plurality at the level of biological neuroscience but also at the level(s) of psychology. Even abstracting from the messy neurobiological details, then, delusions appear not to be a unified kind.

*Response:* How should an advocate of the NK Thesis respond to the present point? One option would be to suggest that delusions are multiply realizable not only relative to neurobiological levels of description but also relative to psychological levels of description as well. But this response is unsatisfactory. Multiple realizability—at least as ordinarily construed—is a one-many, *inter*-level relation. When a kind K is multiply realized by a set of kinds K', K'', ..., there exists a one-many relationship between K at level L and the many kinds at L\*. So, for example, corkscrews are plausibly multiply realizable in this sense because there is a one-many relation between the functional kind *corkscrews* and numerous arrangements of atoms that can be instances of the kind. With

this in mind, the problem with claiming that delusions are multiply realizable relative to our psychology is that it's utterly unclear that there is any other level of explanation relative to which they will constitute a single kind. To the extent that delusions comprise a coherent kind, it is surely most plausible to maintain that they comprise a psychological kind of some sort. Indeed, delusions appear to a *paradigmatic* example of a psychological kind. But if delusions are multiply realizable relative to our psychology, then what would the level of unity be? Sociology? Historical? What? The problem is that there seems to be no plausible option other than psychology. In which case, if delusions constitute a unified kind at all, then it would seem that they have to constitute a psychological kind.

A second, more plausible response to the neural heterogeneity objection turns on the distinction between specific and generic natural kinds. By way of illustration, consider a case from chemistry. Magnesium is plausibly a natural kind. It is one of the chemical elements: one that exhibits a complex cluster of properties, and is subject to many robust empirical generalizations. But metals also constitute a natural kind, though it is a more generic kind that has magnesium (and iron and aluminum etc.) as subordinate kinds or species. The suggestion is that advocates of the NK Thesis can respond to the present objection by claiming that what goes for chemical kinds is true of psychiatric kinds as well. Again, this leads to a positive moral.

*Positive Moral # 4: If the NK Thesis is true, then delusions are most plausibly construed as a generic kind—one that may well subsume many different subordinate or species kinds.* Indeed, this seems to be the *right* result for proponents of the NK Thesis to embrace. For there really does seem to be myriad different kinds of delusions; and an adequate version of the NK Thesis is required to capture not merely the unity in the kind but also the theoretically salient dissimilarities as well. Precisely what these theoretically salient groupings are should not, of course, be considered a matter for *a priori* speculation. Rather if the NK Thesis is true, then it is a matter to be determined by further empirical inquiry. This is, among other things, because the matter of how best to subdivide delusion will turn crucially on which sorts of mechanisms turn out to be responsible for different kinds of delusions. And this is clearly an empirical matter.

## 7. Conclusion: The Unity Problem

This chapter has been an extended exercise in philosophical ground clearing intended to clarify and defend the thesis that delusions constitute a natural kind. Specifically, I have argued that the main extant reasons for rejecting the NK Thesis are unsatisfactory. The argument from anti-essentialism fails because natural kinds need not have sortal essences. Continuity arguments fail because, at most, they establish a continuity of symptoms wholly compatible with the NK Thesis. The arguments for mind-dependence fail because they do not show that delusions are individuatively dependent; and heterogeneity objections fail because they only impose constraints on the sort of natural kind that delusions would need to be, as opposed to showing that they are not a natural kind of any sort.

In addition to these negative conclusions, I have also argued that there are *prima facie* reasons for taking the NK Thesis seriously; and labored to identify both the general conception of natural kinds that should be relevant to the present discussion and those more specific properties that delusions would need to possess, if the NK Thesis is true. Specifically, if the NK Thesis is true, then delusions most likely constitute a multiply realizable, generic cognitive kind whose members characteristically, though not invariably, exhibit those properties enumerated by the standard account of delusion: falsity, entrenchment, doxastic isolation, resistance to rational persuasion and resistance to available incompatible information. On this view, then, there may be many species of mechanism responsible for the production of those properties characteristic of delusions. But these various species of mechanism will need to bear important similarities to each other. They will need to be mechanisms of the same generic kind.

This raises perhaps *the* fundamental explanatory challenge for any natural kinds approach to delusion; what I call the *Unity Problem*. If many different subtypes of mechanism are responsible for delusions, why treat delusions *as such* as a natural kind? It must be because these mechanisms are, themselves, of the same kind. For by assumption, natural kinds are individuated by their causal essences. But what is it about these various mechanisms that make them sub-types of some more general mechanism type, as opposed to merely a heterogeneous collection of *different* mechanisms? This is the Unity Problem. An adequate answer must specify a unifying characterization of the relevant

mechanism type; and moreover, must do so without merely deferring to the fact that they produce similar effects –i.e. those properties associated with delusion. With the ground clearing complete, it is this challenge that should form the focus of future naturalistic research on the nature of delusion.

## References

Bell, V., Halligan, P. & Ellis, H.D. (2006). Explaining delusions: a cognitive perspective. *Trends in Cognitive Sciences*, 10, 5, 219-226.

Bell, V., Halligan, P.W. & Ellis, H.D. (2006) Diagnosing delusions: A review of inter-rater reliability. *Schizophrenia Research*, 86 (1-3), 76-9.

Blakemore, S.J., Wolpert, D.M., & Frith, C.D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, 6(6) 237-242.

Blakemore, S.J., Oakley, D.A. & Frith, C.D. (2003). Delusions of alien control in the normal brain. *Neuropsychologia*, 41, 1058-1067

Block, N. (1997). “Anti-Reductionism Slaps Back.” In Tomberlin 1997, 107-132.

Birchwood, M., Iqbal, Z., Chadwick, P. & Trower, P. (2000). Cognitive approach to depression and suicidal thinking in psychosis: 1. Ontogeny of post-psychotic depression. *British Journal of Psychiatry*, 177, 516-521.

Boyd, Richard (1990), “What Realism Implies and What It Does Not”, *Dialectica* 43: 5–29.

Boyd, R. (1991), “Realism, Antifoundationalism, and the Enthusiasm for Natural Kinds”, *Philosophical Studies* 61: 127–148.

Buchanan, A. & Wessely, S. (2004). Delusions, action and insight. In *Insight and Psychosis*, 2nd edn (ed. X. F. Aamdor and A. S. David). Oxford University Press: Oxford.

Dennett, D.C. (1995) *Darwin's Dangerous Idea*. Simon & Schuster.

Ellis, B. (2001) *Scientific Essentialism*, Cambridge University Press: Cambridge.

Ereshefsky, M. (Forthcoming) "Natural Kinds in Biology" In Craig, E. (ed.) *Routledge Encyclopedia of Philosophy*

Langdon, R., Ward, P.B., & Coltheart, M. (2008) "Reasoning Anomalies Associated With Delusions in Schizophrenia" *Schizophr Bull*.

Freeman, D., Garety, P.A., Kuipers, E., Fowler, D., & Bebbington, P.E. (2002). A cognitive model of persecutory delusions. British Journal of Clinical Psychology, 41, 331-347.

Freeman, D., & Garety, P.A. (2006) "Delusions". In *Practitioners' Guide to Evidence-Based Psychotherapy* (eds.) J.E. Fisher & W. O'Donohue. New York: Springer Academic.

Freeman, D. (2007) "Suspicious minds: The psychology of persecutory delusions", *Clinical Psychology Review*, Volume 27, Issue 4, pp. 425-457

Fulford, K.W.M. (1994) 'Insight, Delusion and the Intentionality of Action: Framework for a Philosophical Psychopathology', in G. Graham and G.L. Stephens (eds.), *Philosophical Psychopathology*, MIT Press.

Gelman S. & Hirschfeld, L. (1999) "How Biological is Essentialism?". In *Folkbiology* (eds.) Medin, D. & Atran, S. MIT Press

Garety, P.A. & Hemsley, D.R. (1994). *Delusions: investigations into the psychology of delusional reasoning*. Oxford University Press: Oxford.

Ghaemi, S.N. (2004). The perils of belief: *Delusions reexamined*. *Philosophy, Psychiatry, and Psychology*, 11, 1, 49-54.

Griffiths, Paul E. (1997), *What Emotions Really Are*. Chicago: University of Chicago Press.

Hacking, Ian (1991), "A Tradition of Natural Kinds", *Philosophical Studies* 61: 109–126.

Howson, C. & Urbach, P. (1993) *Scientific Reasoning: The Bayesian Approach, 2nd ed.* Chicago: Open Court.

Johns, L.C. and van Os, J. (2001) The continuity of psychotic experiences in the general population. *Clin. Psychol. Rev.* 21, 1125–1141

Kim, J. (1992). "Multiple Realization and the Metaphysics of Reduction." *Philosophy and Phenomenological Research*, 52: 1-26.

Kripke, S. (1972) "Naming and Necessity", in *Semantics of Natural Language*, D. Davidson and G. Harman (eds.), Reidel

Lewis, D. (1986) *On the Plurality of Worlds*. Blackwell

Machery, E. (2005) "Concepts are not a Natural Kind" *Philosophy of Science*, 72 pp. 444–467

Murphy, D. (2006) *Psychiatry in the Scientific Image* MIT Press

- Nozick, R. (1993) *The Nature of Rationality* Princeton University Press
- Page, S. (2006) "Mind-independence disambiguated: separating the meat from the straw in the realism/anti-realism debate", *Ratio XIX* 3 pp. 321-335
- Putnam, Hilary (1975), "The Meaning of 'Meaning'", in Hilary Putnam, *Mind Language, and Reality*. Cambridge: Cambridge University Press, 215–271.
- Pylyshyn, Z. (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.
- Quine, Willard V. O. (1969), "Natural Kinds", in William V. O. Quine, *Ontological Relativity and Other Essays*. New York: Columbia University Press, 114–138.
- Shapiro, L. (2000). "Multiple Realizations." *Journal of Philosophy*, 97: 635-654.
- Soames, S. (Forthcoming) "What are Natural Kinds?" *Philosophical Topics*
- Sorensen, R. (2006) "Epistemic Paradoxes, "Nothingness", and "Vagueness" *Stanford Encyclopedia of Philosophy* (ed.) E. Zalta at <http://plato.stanford.edu/entries/>
- Stich, S. (1990) *The Fragmentation of Reason* MIT Press
- Stoljar, D. & Gold, I. (1998) 'On Cognitive and Biological Neuroscience' *Mind and Language*, Vol. 13, Number 1, pp. 110-131
- Unger, P. (1979) "I do not Exist", in *Perception and Identity*, G. F. MacDonald (ed.), London: Macmillan
- van Os, J. (2003) Is there a continuum of psychotic experiences in the general population? *Epidemiol. Psychiatr. Soc.* 12, 242–252

Wakefield, J. C. (1992). Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder. *Psychological Review*, 99, 232-247.

Weisberg, 2003 "Water is not H<sub>2</sub>O" in *Philosophy of Chemistry: Synthesis of a New Discipline*, a volume of *Boston Studies in the Philosophy of Science*.

Zachar, P. (2000). Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry, and Psychology*, 7, 167-182.