This is an author's version of a chapter published in *Philosophical issues in Psychiatry II:*

***Nosology*. Eds K. Kendler and J Parnas. OUP (2012) pp 61-70**

Chapter 4

Title: Is psychiatric classification a good thing?

Rachel Cooper

In this chapter I seek both to ask whether psychiatric classification is a good thing, and also to make it clear how this is an issue on which reasonable people can disagree. When I'm talking of "classification" I have in mind the type of classification facilitated by diagnostic systems such as the DSM.

We can question the wisdom of classifying mentally ill people either in general or in particular cases. At the general level, we might ask whether research programmes such as that associated with the DSM are a force for good or evil. At the particular level, we may query the role of classification in some particular sub-domain, currently, for example, personality disorders stand out as a particularly contested area. Questions at the two levels are of course linked. Those who are generally skeptical about psychiatric classification will worry about classification in many particular cases. Writers such as Thomas Szsaz (1974), for example, would accept only classifications of organic brain disorders as legitimate. At the other end of the scale, even those who are generally pro-classification will agree that some areas of human behavior should not be included in the DSM, for example, they might worry about the potential medicalisation of normal grief (Horwitz and Wakefield 2007),

As tends to be the case with philosophical discussion, much of my argument will be at a fairly abstract level and will be couched in general terms – I shall consider why, in general, classification can be helpful, and what, in general, are the risks attached to classification. However, applying the discussion to particular cases is straightforward, and I shall also mention

particular cases as the chapter proceeds. My overall claim will be that classification in psychiatry is frequently, but not always, a good thing. The chapter is split into three main sections. The first considers the benefits of classification, the second considers the harms that classification can produce, the third, and most tentative, section starts to consider how classificatory projects might best proceed in order to maximize the benefits and minimize the harms.

1.   The benefits of classification

Classification can enable us to gain power over a domain. Where entities fall into groups that are genuinely similar to each other in theoretically important ways then classifying them into groups of like entities is a valuable part of scientific practice. As entities that fall into such natural groups are similar they can be expected to behave similarly. What is all this talk of "genuinely similar", "natural", "theoretically important" doing? Basically I just want to emphasise that in order for classification to yield power over a domain the similarities have to be significant and out there in the world. The classification that I'm interested in is of the sort that has been used to such great advantage in chemistry and the biological sciences. All samples of an element have the same atomic number, and this ensures that their properties are alike. All members of a species are genetically similar and have similar developmental histories; as such they can be expected to thrive in the same sort of habitat, to eat the same food, to have similar life spans, and so on.

Those who favour classification frequently talk of "natural kinds". Members of a natural kind are alike, and natural laws mean that members of a kind will behave similarly. Depending on the author, other conditions have also been added. Natural kinds have been claimed to be universal (in the sense of occurring everywhere), discrete, to have essential properties, and so on and so forth (see, for example, the conditions imposed by Haslam 2003, Zachar 2001). In earlier work I talked in terms of natural kinds, and claimed that at least some mental disorders can be considered natural kinds (Cooper 2005). One of the problems I encountered is that the term "natural kind" has become encrusted with metaphysical baggage. When I talked about "natural kinds", intending to use the term with minimal commitments, I was heard as being committed to all sorts of things.

Now instead of talking of natural kinds I will talk of "repeatables". This makes clear the basic important idea; some entities in the world are alike, and will behave in similar ways. As applied to mental disorders, the idea that there may be repeatables is this: if we consider individual cases of mental disorder some can be seen to be similar to each other. Furthermore some of these similarities will be theoretically important, and in some cases patients who are grouped together will be alike in fundamental ways (maybe they all have the same genetic abnormality, or all have similar levels of some neurotransmitter, or all have similar relationships with their childhood caregivers). If we take cases of mental disorder as our domain and plot them onto a multidimensional quality space (as in cluster analysis) then we will find clusters of similar cases.

If we focus on the right properties, then the clusters that such a process generates will be inductively powerful. External validation on the basis of treatment response, family history, demographic correlates and so on can give additional reason to believe that patients who are being classified together are similar in genuinely important respects. If all goes well, a case that falls in a particular cluster can be expected to behave in ways that are similar to others of its class. The importance of such similarities is obvious if one thinks of treatments. The hope would be that a treatment that is found to work for one member of a class will work for others in that class too.

Note that the key question in looking for repeatables is whether cases of a mental disorder are importantly similar to each other. The question of whether varieties of mental disorder will turn out to be discrete or continuous turns out to be a side-issue. It is the fact that there are similarities between entities that does all the work when it comes to making inductive inferences and grounding explanations. It is because "repeatables" all have similar properties that one will behave like the others of its type. Thus, classifications that vary along dimensions can be as powerful as those that rely on discrete categories. Think of alloys as an example. Knowing that a sample is a particular alloy is as useful, and useful in the same kinds of ways, as knowing that it is a 100% pure metal (if a sample is known to be 55% zinc and 45% copper, one can predict how the sample will behave just as well as if one knew it to be pure copper). There is reason to think that at least some mental disorders will be better mapped by a continuous rather than a categorical classification system. Draft versions of the DSM-V suggest that classifications of personality disorders will go this way.

The idea that mental disorders are repeatables is a very weak claim. It says simply that cases of disorder can be grouped together on the basis of important similarities, and that cases that are grouped together can then be expected to behave in similar ways. One potential worry is that we might be able to pick out many, but inconsistent, potentially useful classifications. In his 1990 paper "Toward a scientific psychiatric nosology" Ken Kendler suggests that this might turn out to be the case with some mental disorders. Kendler takes the case of schizophrenia, and discusses evidence that the criteria that pick out subtypes that best predict treatment response may be different from those that best fit with familial aggregation.

Similar situations occur elsewhere in science, and occur because the world is messy and complex. Within biology, species can be delimited on multiple different criteria. Evolutionary theorists are chiefly interested in groups of organisms based on common descent. Ecologists are more interested in classifying on the basis of current behavior. Thinking about such cases, the philosopher John Dupré urges us to be realists, but "promiscuous realists" (Dupré 1981, 1990, 2001). His thought is that multiple different classification systems can be picked out, with different classification systems being most suitable for different purposes. Thinking in terms of a multidimensional quality space, on Dupré's picture we can discern different groupings if we focus in at different levels of resolution or restrict our attention to particular dimensions of the space. Dupré's suggestion is that we should let a thousand flowers bloom, and that each scientific sub-discipline should be permitted to classify as it finds most useful

Returning to psychiatry, a promiscuous realist would suggest that if, for example, it is the case that one set of criteria best predicts treatment outcome while another set best predicts how disorders run in families then researchers interested in different questions should use different criteria. Though Dupré tends to emphasise the advantages of embracing multiplicity, it should be noted that there is also a downside. If different researchers use different sets of criteria for different sorts of research, then seeking to combine their findings to make an overall judgment becomes problematic. Suppose that those working on schizophrenia did start to use multiple different sets of criteria. In effect, those interested in treatment would be talking about slightly different entities when they talked of "subtypes of schizophrenia" than would those looking at patterns of inheritance. As a consequence, "translation" would be required in those cases where data from both sorts of studies was required together. Recognising multiple sets of repeatables will thus not be cost free and may not always be worthwhile. Still the key point can be upheld – the existence of competing classifications is compatible with a domain consisting of repeatables.

To summarise this section: classification will be useful in so far as mental disorders turn out to be repeatables, in the sense that all cases of a type of disorder are fundamentally similar. We should hope that mental disorders do turn out to be repeatables, as only with repeatability will it be possible to develop therapies that can be hoped to work for all cases of a kind of mental disorder.

2. The Dangers of Classification

And yet, a major tradition sees something problematic about classifying human beings. A sign outside the counsellors' office at my university depicts tins of food and reads "Labels are for tins. Not people" in the same sort of tone that other posters warn of the evils of racism or domestic violence. One infers that labelling people is not a nice thing to do. The idea that classifying people is at least a little bit evil can also be found in the work of many of those who are opposed to 'the medical model" (as found in antipsychiatry, critical psychiatry, postpsychiatry).

Diagnostic labels may cause harm in various ways. For example, in some cases, labels harm individuals by facilitating prejudice against them. Thus racial labels played a role in enabling the system of apartheid practiced in South Africa. Clearly labels can also harm through inaccurately reflecting the structure of reality. For example, when a classification system mistakenly lumps together disorders that are really distinct this may result in patients receiving sub-optimal treatments. Though important, there is little of philosophical interest to say about such harms. In this section I focus on a different and philosophically under-investigated variety of harm. These harms arise because diagnostic labels can enter the narratives by which people make sense of their lives and thereby limit the meaningful futures that a person can imagine. In discussing these harms I will draw on Alastair MacIntyre's (1981) ideas about the importance of narrative for human flourishing, Ian Hacking's (1995) work on the looping effects that affect human kinds (1995), and Carl Elliott's (1999, ch.7) work on diagnosis and identity.

Narratives are important for human flourishing. In recent years this claim has become a commonplace in both philosophy and medicine (MacIntyre 1981, Elliott 1999). Some go so far as to link the narratives that a person tells about themselves with a person's identity (MacIntyre 1981). I don't want to align myself with such radical views here, but I do think that the narratives that structure people's lives are important. At the very least such narratives help to shape what an individual thinks they might do and how they come to understand how they have acted in the past.

Both illness itself and the act of diagnosis can threaten our ability to narrate our lives. The fact that illness can compromise our narrative abilities, for example through distracting us with pain, or by destroying our memory, should be underlined. I don't want to give the impression that I think it is only the talk of doctors rather than also the problems of bodies and minds that cause difficulties for narrative agency. Here though I shall focus on the problems that the act of diagnosis can itself produce.

The nub of the problem is this: we structure our lives with the help of narratives, but we are not the sole authors of our life stories. Others too play a role in shaping what we can sensibly say about ourselves. Some co-authoring occurs by negotiation, but some situations place us in a position similar to that of someone playing the "continue-the-story" game played by children. In the children's game someone writes the first few sentences of a story, they pass it on to the next

player who adds a passage, who passes it on to a third, and so on. The challenge is to continue the story in a way that makes sense given what one's co-authors have said.

The interactions between a patient and a mental health professional as a diagnosis is made can be thought of in a similar vein. When a patient goes to a professional they tell them part of the story, the professional, in making a diagnosis, adds to it, and the patient is left to continue. Sometimes the effect of being diagnosed may be minimal, and arguably diagnosis may enable someone to practice reasonable planning and thus gain control over their life. Suppose I come to think of myself as having depressive tendencies. This may structure my actions in certain ways. Maybe I avoid drugs that have been found to trigger depression in those who are susceptible. Such actions may be reasonable and helpful for me.

However certain diagnoses are more problematic. Certain diagnoses will imply that one's assessment of reality is not reliable (schizophrenia), or that one is essentially manipulative (Borderline Personality Disorder), or that one can never be trusted around children (paedophilia). Once one accepts such a diagnosis as accurate, telling a good story about one's life will become difficult. By telling a "good story" I mean both telling a story with narrative coherence, plot etc. and also telling a story whereby one appears as a decent human being. Plausibly, people need to be able to tell stories about themselves that are good in both senses if they are to think well of themselves. To illustrate the problems that diagnoses can pose let's consider Antisocial Personality Disorder. This is an extreme example, in so far as a diagnosis of ASPD will be one of

the toughest to incorporate into a good story about one's life, however it will clearly illustrate the problems that diagnostic labels can pose.

According to the DSM, ASPD is characterized by a number of undesirable character traits – aggressiveness, irresponsibility, deceptiveness, and so on. In short, someone with ASPD is a bad person. In addition, a powerful tradition has it that that personality disorders are lifelong states that can be highly resistant to treatment. That is someone with ASPD is an irrecoverably bad person. Suppose one receives a diagnosis of ASPD. What does one do then? As I found it hard to imagine how one might respond to receiving such a diagnosis I looked at posts on an online support group for people with ASPD (http://www.psychforums.com/antisocial-personality). There seemed to be three basic ways to respond to diagnosis.

i. Challenge the diagnosis.

Some refuse to believe the diagnosis. Either they give reasons for distrusting the individual clinician who diagnosed them, or they give reasons for thinking that all psychiatric diagnoses are unreliable. Given the esteem with which medicine is held in our culture, challenging a diagnosis will not always be a viable possibility.

ii. "Embrace the dark-side".

Some embrace the idea that they are evil. They have online names like "Lannibal Hector" and "Rage" and swap stories about torturing small animals and homeless people. Amongst such discussions the more sophisticated present themselves as being moral relativists, or think of themselves as Nietzschean supermen. The problem with this option is that it is morally unacceptable. Someone with ASPD who takes this option becomes worse than they were before.

iii. Uncertainty

Some people don't know how to respond to their diagnosis. They have found themselves diagnosed and then, in some cases, abandoned by mental health professionals, and don't have any idea what they should do now they have come to think of themselves as people with ASPD.

Admittedly, the example of ASPD is an extreme one, and one might feel little sympathy for people who manifest the types of behaviours that tend to lead to them receiving such a diagnosis. Still the example of ASPD illustrates how the act of diagnosis can itself harm someone. Coming to believe that one is an untreatably bad person is difficult to live with.

Of course physical diagnoses also limit the narratives that patients can sensibly tell about their lives. Most clearly this is the case with diagnoses of terminal illness. However diagnoses of mental disorder are perhaps particularly hard to incorporate into a good narrative about oneself, both because many mental disorders are chronic conditions, and also because in so far as mental

disorders affect personalities, emotions and beliefs they affect a person considered as an agent more directly than do many physical disorders.

2.   What to do?

Classifying mental disorders can be hoped to bring great benefits. If mental disorders are repeatables, then once a correct classification scheme is achieved, diagnosis can be expected to predict how a case will behave. In particular, a treatment that works for some members of a class can be hoped to work for others. However, classifying people can also harm them. Harms may come about in various ways, but here I have focused on the ways in which diagnostic labels enter into the narratives by which people make sense of their lives and can limit the range of imaginable future courses of action. In this final section I begin to consider how classification systems might be developed so as to maximise their potential benefits and limit the associated risks.

When considering the benefits and harms associated with classification, we can start by noting an unfortunate asymmetry. A classification system can only be expected to be useful if it is at least approximately correct, but it can harm people even if it is wrong. It's my *belief* that there is a disorder such as ASPD, and that I have it, that harms me.

How should we act in such a situation? When hypotheses will harm whether they are right or wrong, but only do good if they are roughly correct, it would be wise to proceed with modesty and caution. As we have seen by considering the case of ASPD, a classification system that makes the general claim "There are people of type X", and a diagnostician who makes the particular claim "You are a person of type X" can do harm. Such labels can enter the narratives that people tell about themselves and limit the possible futures that they can imagine. In such situations it is better that the doubts that surround the validity of classes in a classification system, or an individual's diagnosis, are made explicit. Where there are doubts, a classification system that makes it explicit that there are competing classifications, and that the validity of a category is disputed, will do less damage. How modest is the current DSM? Not very. The foreword to the DSM presents it as being a work in progress, but the language used in the main text suggests that the claims made in the DSM are definite truths rather than contested hypotheses. The harms produced by psychiatric classification could be reduced by making classifications such as the DSM more explicitly tentative.

The harms that result from diagnosis can also be minimised by making hopeful illness narratives – where such exist – more accessible. In general, individuals model the stories that they tell about themselves on those that are readily available in their culture. Publicising examples of people who manage to tell good stories about themselves while living with condition X will thus make it more likely that those who are newly diagnosed with condition X will also come to be able to tell good stories about themselves. For example, in the past, telling a good life story that incorporated a diagnosis of schizophrenia was very difficult. In recent years it has become easier

because hopeful narrative templates have become available via the Hearing Voices Network and similar groups.[1]

In so far as it is knowledge of psychiatric labels that causes harm, one might wonder if things would be better if classification systems were developed and diagnoses made in secret. Maybe the harms caused by psychiatric diagnosis occur because lay people and patients currently know too much, and a more secretive psychiatry would do less damage? Such a suggestion should be rejected however. Within the human sciences, rigid distinctions between those who classify and those who are classified have been linked with a sorry history whereby classification has come to be biased against the less powerful (for studies of gender and psychiatry see Lunbeck 1994, on race see Fernando 2002, Cooper 2007 ch.8). The DSM itself has a history whereby categories have come to be included on dubious grounds – in the past, lobbying by special interest groups, and pressures from the insurance and pharmaceutical industries have played a role (Cooper 2005; Kutchins and Kirk 2003). It is plausible that more openness, rather than less, is needed to combat such tendencies (Longino 1990).

A secret psychiatry is undesirable, however this doesn't mean that patients necessarily have to be told of their diagnosis. Consider the "don't ask, don't tell" policy adopted by many of those who deal with Huntington's Disease. Huntington's Disease is a horrible, untreatable, genetically-caused disorder that develops during middle-age. Genetic tests mean that those who will develop it can be identified. However, many of those who know they are at risk decline the tests. If they

---

[1]   http://www.hearing-voices.org/

are going to die horribly they would rather not know. I think that it would be rational for those

who suspect that they might be diagnosed with certain psychiatric disorders to similarly avoid

finding out. Most notably, there is often going to be little value in being diagnosed with a

personality disorder – as we have seen with the case of ASPD such diagnoses can leave patients

in a position where they are unable to tell a coherent good story about themselves. Knowledge

isn't always a good thing.

Modesty, publicising hopeful disorder narratives, and don't ask, don't tell policies will help limit

the harms of diagnosis, but some damage will still be done. At the end of the day I suggest that

whether one thinks that classifying some subdomain of mentally ill people is on balance a good

or bad idea must depend on a weighing up of the costs and benefits. This approach is consistent

with that which will be adopted by those drafting revisions to the DSM-V where in the case of

revisions "potential benefits…should outweigh potential harms" (Kendler et al. 2009). It should

be noted that as diagnoses affect not only patients, but also their relatives and friends, and

broader society, the costs and benefits to be considered need not be limited to those affecting the

patient.

The idea that only revisions that are likely to do good should be included in the DSM has come

in for criticism. In a recent piece in the Bulletin of the AAPP, Nassir Ghaemi argues that the

gerrymandering of categories will set back research. He thinks that a classification should seek to

mirror the natural structure of the domain of mental disorders, and that this will best enable

research that will one day lead to pragmatic benefits. My suggestion is not that the DSM should lie about the nature of mental disorders, but rather that there may be conditions about which it should keep silent. The fundamental basis for my suggestion is that I think there may be truths that we are better off not knowing. This point is most easily made when one thinks of military research – there are facts about poisons that we do not need to find out. In so far as research in psychiatry can harm people there may similarly be areas where it is best not to conduct research.

Unfortunately, in practice determining whether a new category will be helpful can be extremely difficult. Consider current debates over the possible addition of Psychosis Risk Syndrome (Moran 2009). Advocates argue that inclusion of the new diagnostic category will facilitate early treatment. Critics argue that the stigma that will be associated with the diagnosis, and the side-effects that will result from drug treatment, outweigh any benefits. This case nicely brings out how very hard it is to determine whether a new category will overall do good or be harmful. Some of the consequences of introducing the new category are fairly predictable. Pharmaceutical companies would seek to market drugs for the treatment of such patients. Some employers, particularly those seeking recruits for high-stress jobs, would seek to avoid employing people with Psychosis Risk Syndrome. Other consequences of introducing a new category are hard to predict. What would it be like to be diagnosed with Psychosis Risk Syndrome? Would one spend one's time worrying about becoming psychotic? Or would one be pleased that one's condition was being monitored?

Trying to predict the effects of a new diagnostic category will always be difficult. We can, however, make some generalizations. Where a condition is mild and currently untreatable, patients will have little to gain from it being classified. If that condition affects traits close to an individual's core identity the potential for the classification to do harm may also be great. Prodromal Personality Disorders thus stand out as candidate disorders where classification could be expected to bring no net benefits for patients.

Although we can get some way thinking about whether classification is a good idea on a case by case basis, often the relevant issues will be very difficult, and so we will find ourselves falling back on considerations about the benefits and harms of classification in general. For example, whether one will think that adding some particular new, but currently untreatable, condition is worthwhile will depend in large part on whether one thinks that research is likely to result in a successful treatment being developed. And, one's judgment of the prospects of research will depend on whether one thinks that psychiatric research has generally managed to lead to useful treatments.

Although I am not as optimistic as some, I tend to optimism. I think that mental disorders may well turn out to be repeatables, and that as a consequence the development of successful class-based treatments (that is treatments that will be effective for all those with a particular type of disorder) can be hoped for. On balance, and despite the harms it causes, I think that much, though not all, psychiatric classification is justifiable. However I have shown how this is an issue

on which reasonable people can disagree. Determining when classification can be hoped to bring more benefits than harms is far from clear-cut.

Chapter References:

Cooper, R. (2005). *Classifying Madness.* Dordrecht: Springer.

Cooper, R. (2007). *Psychiatry and Philosophy of Science.* Stocksfield: Acumen.

Dupré, J. (1981). Natural kinds and biological taxa. *The Philosophical Review*, XC, 66-90.

Dupré, J. (1993). *The Disorder of Things*. Cambridge, Massachusetts: Harvard University Press.

Dupré, J. (2001). In defence of classification. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 32, 203-219.

Elliott, R. (1999). *A Philosophical Disease: Bioethics, Culture and Identity.* New York: Routledge.

Fernando, S. (2002). *Mental Health, Race and Culture*. 2nd ed. Houndsmills, Basingstoke: Palgrove.

Ghaemi. N. (2010). DSM-IV, Hippocrates, and pragmatism: What might have been. *A.A.P.P. Bulletin*, 17, 32-35.

Hacking, I. (1995). The looping effect of human kinds. In D. Sperber, D. Premack, and A. James Premack (eds.) (1995) *Causal Cognition: A Multidisciplinary Debate*. Oxford: Clarendon Press. pp. 351-383.

Haslam, N. (2003). Kinds of kinds: A conceptual taxonomy of psychiatric categories. *Philosophy, Psychiatry and Psychology,* 10, 203-217.

Horwitz, A. and Wakefield, J. (2007). *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder.* Oxford: Oxford University Press.

Kendler, K. (1990). Toward a scientific psychiatric nosology. Strengths and limitations. *Archives of General Psychiatry*, 47, 969-973.

Kendler, K. , Kupfer, D., Narrow W., Phillips, K., and Fawcett, J. (revised 21.10.09). Guidelines for making changes to DSM-V. Available online at http://www.dsm5.org/ProgressReports/Documents/Guidelines-for-Making-Changes-to-DSM_1.pdf

Kutchins, H. and Kirk, S. (2003). *Making Us Crazy.* New York: Free Press

Longino, H. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry.* Princeton: Princeton University Press.

Lunbeck, E. (1994). *The Psychiatric Persuasion.* Princeton: Princeton University Press.

MacIntyre, A. (1981). *After Virtue.* London: Duckworth.

Moran, M. (2009). DSM-V developer weigh adding psychosis risk. *Psychiatric News.* August 21. Vol 44: No 16 p.5.

Szasz, T. (1974). *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. Revised edition. New York: Harper & Row.

Zachar, P. (2001). Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry and Psychology,* 7, 167-181.