Essays in the Philosophy of Psychiatry

by

Jessica Sin

A.B. Chemistry and French
Dartmouth College, 1997

M.A. Chemistry
Princeton University, 2000

Ph.D. Chemistry
Princeton University, 2002

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Philosophy

at the

Massachusetts Institute of Technology

September, 2007

© 2007 Jessica Sin. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly
paper and electronic copies of this thesis document in whole or in part in any
medium now known or hereafter created.

Signature of Author ..........................................................

Certified by .................................................................
Richard Holton
Professor of Philosophy
Thesis Supervisor

Accepted by .................................................................
Alex Byrne
Professor of Philosophy
Chair, Committee on Graduate Students
Essays in the Philosophy of Psychiatry
by
Jessica Sin

Submitted to the Department of Linguistics and Philosophy
on July 2, 2007 in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Philosophy.

Abstract

This dissertation consists of three chapters in which I address metaphysical and epistemological issues that arise in psychiatry, with particular attention paid to anti-psychiatric concerns.

In Chapter 1, I consider three versions of anti-realism about psychiatric illness. I argue that Szasz’s version of anti-realism should be rejected because it rests on a misunderstanding of illnesses more generally. Although I do not offer any clear refutations of labeling theory or cultural relativism, I point out the serious disadvantages of holding either view. I argue that in the absence of compelling reasons to endorse either labeling theory or cultural relativism, we are within our rights to remain realists about psychiatric illness.

In Chapter 2, I address an epistemological concern that the scientific legitimacy of psychiatric taxonomy is compromised by the role that value judgments play in the study of mental disorders. I claim that this worry presupposes a view of science according to which objective observation and theory construction would not even be possible. I argue that, on a revised understanding of science proposed by Helen Longino, a scientifically legitimate psychiatric taxonomy is within our reach.

Finally, in Chapter 3, I turn to the metaphysical problem of providing an account of disorders. An important part of a realist view of mental disorders includes an account of disorders. I claim that in light of available evidence of the heterogeneity of disorders, it is unlikely that disorders share an essence, and I argue that previous attempts to provide a Lockean account of disorders fail for this reason. I propose instead that disorders are homeostatic property cluster kinds of the sort first described by Richard Boyd.

Thesis Supervisor: Richard Holton
Title: Professor of Philosophy
Acknowledgments

I would like to thank the members of my committee. I thank my thesis advisor, Richard Holton, for his guidance throughout the dissertation writing process. I would like especially to acknowledge Sally Haslanger for generously offering me many hours of her time to discuss this work at all the stages of its development. I would also like to thank Alex Byrne for his helpful comments on various drafts of this work.

I would like also to express my gratitude to Sabina Berretta, M.D. and Dost Ongur, M.D., Ph.D. for welcoming me into their research groups where I gained valuable insights into some of the conceptual and methodological problems in basic and clinical psychiatric research. I would like to acknowledge Allison Lin and Miriam Lundy for lively and illuminating discussions which have inspired some of the work presented here.
## Contents

1 Anti-Realism about Mental Illness ........................................... 9
   1.1 Introduction ......................................................................... 9
   1.2 General Disputes about Realism .......................................... 10
   1.3 Scheff's Labeling Theory ..................................................... 14
   1.4 Szasz's Myth of Mental Illness ............................................. 19
   1.5 Cultural Relativism ................................................................ 22
      1.5.1 The Ambiguity of Available Evidence ............................. 26
      1.5.2 The Disappearance of Genuine Disagreement ................. 28
      1.5.3 The Impotence of Criticism ......................................... 31
   1.6 Conclusion ........................................................................... 34

2 Psychiatric Taxonomy ................................................................. 35
   2.1 Introduction ......................................................................... 35
   2.2 Science, Objectivity and Value-Neutrality .............................. 37
      2.2.1 Objectivity ................................................................. 37
2.2.2 Value-Neutrality ........................................ 39
2.2.3 The Scientific Method .................................... 41
2.2.4 Psychiatric Taxonomy as Pseudoscience .................. 43
2.3 The Traditional View Rejected .............................. 47
  2.3.1 Is inquiry guided by contextual values always unscientific? 48
  2.3.2 Do contextual values always introduce distortions and biases into observation and theory construction? ............ 51
  2.3.3 Summary ............................................. 57
2.4 Anything Goes? ............................................. 57
2.5 Conclusion .................................................. 59

3 What kind of thing is a disorder? 61
  3.1 Introduction .............................................. 61
  3.2 Lockean Kinds ............................................. 63
  3.3 Disorders as Purely Natural Lockean Kinds .................. 65
    3.3.1 The Statistical Deviation Account ..................... 65
    3.3.2 The Biological Disadvantage Account .................. 66
  3.4 Disorders as Partially Natural Lockean Kinds ................. 68
  3.5 Disorders as Property Cluster Kinds ........................ 73
  3.6 Conclusion ............................................... 81
Chapter 1

Anti-Realism about Mental Illness

1.1 Introduction

In the 1960s, an anti-psychiatric movement emerged and challenged the legitimacy of psychiatry as a medical discipline. (Berlim, Fleck & Shorter 2003) Although describing it as a movement suggests that it was a coordinated effort, anti-psychiatry was championed by activists, theorists, and writers for diverse reasons. Some expressed legal and ethical concerns about psychiatrists’ right to detain and coercively treat persons who were allegedly mentally ill. Others worried about the ‘medicalization’ of psychiatry. Still others doubted the very existence of mental illness and endorsed some form of anti-realism about psychiatric disorders. Though all of these concerns are worthy of critical scrutiny, I would like to address claims of the last kind. In particular, I would like to query the plausibility of the three most influential versions of anti-realism about psychiatric illness. I begin, in Section 1.2, with a general presentation of the important points of dispute between realists and anti-realists. In Sections 1.3 and 1.4, I discuss, in turn, Thomas Scheff’s labeling theory of mental illness, and Thomas Szasz’s allegation that mental illness is a myth. I argue that neither Scheff nor Szasz gives us compelling reasons to accept anti-realism about psychiatric illness. Finally, in Section 1.5, I consider efforts to promote cultural relativism with respect to mental illness, and I argue that despite the apparent appeal of this view, it is not one that is defensible.
1.2 General Disputes about Realism

In ‘The Many Moral Realisms’, Geoffrey Sayre-McCord discusses what is at stake in any debate between a realist and an anti-realist. (Sayre-McCord 1988a) Although Sayre-McCord is principally interested in moral realism and opposing views, he presents a very general account which is intended to capture the major points of disagreement in a wide range of philosophical debates about realism, including those in metaphysics, mathematics, and philosophy of science. For exegetical purposes, I shall first describe Sayre-McCord’s account as it applies to moral philosophy, and subsequently, I shall point out important parallels between debates about moral realism and those about psychiatric realism.

According to Sayre-McCord, in any debate about realism, the meaningfulness or the truth of certain claims is under dispute. For example, moral realists and anti-realists disagree about how to interpret claims about what is right or wrong and about what is good or bad. In particular, they are concerned with whether judgments about right and wrong or good and bad commit us to the existence of moral entities such as the property of being right or the fact that keeping one’s promises is right. Likewise, scientific realists and anti-realists argue about the ontological commitments which allegedly accompany acceptance of scientific claims expressed using theoretical terms. For instance, consider the claim that the demotion of an excited electron from the pi* orbital to the pi orbital of a conjugated system caused the emission of energy which we observe as fluorescence. A scientific realist would argue that by accepting this claim as part of our account of worldly happenings, we commit ourselves to the existence of entities such as electrons and molecular orbitals. In contrast, an anti-realist denies that we are so committed and endorses, instead, a more austere ontology. As an account of realism in its most general form, Sayre-McCord proposes that this view is best understood as a commitment to two theses:

**Thesis 1** The claims in question, when literally construed, are literally true or false.

**Thesis 2** Some of these claims are literally true.

Recall that, in the case of moral realism, the claims under dispute include those that are about what is right or wrong and about what is good or bad. For example, consider claims of the form given in Schema 1.
Schema 1  Action $A$ is wrong.

Taken together, Theses 1 and 2 imply that instances of Schema 1 include a few important presuppositions. According to Thesis 1, when an instance of Schema 1 is literally construed, it expresses a proposition according to which a certain property is predicated of a certain action. That is, according to the proposition expressed, the property picked out by ‘is wrong’ is predicated of the action that is the referent of an expression substituted for $A$ in Schema 1. Moreover, this proposition either does or does not represent the world accurately. As such, Schema 1 presupposes (i) that there exists a property $F$ that is picked out by the predicate ‘is wrong’, and (ii) that when one utters an instance of Schema 1, property $F$ is being predicated of action $A$. Furthermore, according to Thesis 2, since some instances of Schema 1 are literally true, Schema 1 also presupposes (i) that the property $F$, to which ‘is wrong’ refers, just is the property of being wrong; and (ii) that there is at least one action that has the property of being wrong. The presuppositions to which Theses 1 and 2 commit us when we consider instances of Schema 1 are summarized here:

1 According to Thesis 1, instances of Schema 1 presuppose:
   
   a that there exists a property $F$ that is picked out by the predicate ‘is wrong’;
   
   and
   
   b that when one utters an instance of Schema 1, the property $F$ is being predicated of an action.

2 According to Thesis 2, instances of Schema 1 presuppose:
   
   a that the property $F$ is literally the property of being wrong (simpliciter);
   
   and
   
   b that there is at least one entity that has the property of being wrong (simpliciter).

As Sayre-McCord points out, one way to be an anti-realist about morality is to deny Thesis 1. Two of the most historically significant ways to do so include either embracing emotivism (Ayer 1946) or endorsing prescriptivism (Hare 1952). According to both emotivists and prescriptivists, when one utters an instance of Schema 1, one is
not predicating a property of an action. That is, both emotivists and prescriptivists, deny Thesis 1 because they reject presupposition 1(b). Emotivists differ from the prescriptivists in their account of what one is doing by uttering instances of Schema 1 instead of predicating properties to actions. Emotivists hold that one is expressing a negative emotion or attitude as one would by uttering 'Boo to doing A!'. In contrast, prescriptivists hold that one is administering a command, making a recommendation, or voicing a prohibition as one would by uttering 'Do not do A!'.

Another way to be an anti-realist about morality involves denying Thesis 2. The most common way of denying Thesis 2 is to reject presupposition 2(a), and to embrace some form of moral relativism. (Harman 1996, Mackie 2001) An anti-realist of this sort denies that we have any good reason to think that there exist non-natural, moral properties such as the property of being wrong (simply). A standard argument for anti-realism of this sort begins with the observation that there is a great deal of cultural variability in moral codes. There are, then, two competing ways of explaining such radical differences in moral codes, judgments, and conduct: realism or relativism. According to the former, we must postulate mysterious entities, such as moral properties and moral facts, and we must explain away the observed cultural differences by arguing that many people have inadequate and badly distorted perceptions of these moral entities. In contrast, the relativist can explain widespread variability by appealing to social and anthropological differences, that we already know to exist, without extravagantly populating the universe with entities that are at best ill-understood. According to relativism, there are no mysterious moral properties or moral facts, and thus, when one utters an instance of Schema 1, say, 'Telling lies is wrong', one is not predicating the moral property of being wrong (simply) of the action-type telling lies. Rather, one is predicating the social property of being inconsistent with the social conventions for social group S, where it would be especially salient within the context of conversation which group S is.

Although much more should be said about both moral realism and the opposing views, I presented these accounts here briefly in order to locate some similarities between these debates and those about the ontological status of psychiatric illnesses. Psychiatric realism, as I shall call it, is a philosophical view about judgments concerning mental illness, which can be expressed using statements of the form given in Schema 2.¹

¹Usually, when we make psychiatric judgments, we use predicates like 'is schizophrenic' or 'is bipolar', and it may seem strange to use the predicate 'is mentally ill' to talk about someone's
Schema 2 Person $P$ is mentally ill.

As in the case of moral realism, Theses 1 and 2 when applied to psychiatric claims imply that instances of Schema 2 include a few important presuppositions. According to Thesis 1, when an instance of Schema 2 is literally construed, the predicate 'is mentally ill' picks out a property $G$, and this property is predicated of a certain person. As such, Schema 2 presupposes (i) that there exists a property $G$ to which the predicate 'is mentally ill' refers, and (ii) that when one utters an instance of Schema 1, property $G$ is being predicated of someone. Moreover, according to Thesis 2, since some instances of Schema 2 are literally true, they also presuppose (i) that the property $G$ to which 'is mentally ill' refers is literally the property of being mentally ill; and (ii) that there is at least one entity that has the property of being mentally ill. The presuppositions to which Theses 1 and 2 commit us when we consider instances of Schema 2 are summarized here:

3 According to Thesis 1, instances of Schema 2 presuppose:

a that there exists a property $G$ that is picked out by the predicate 'is mentally ill'; and

b that when one utters an instance of Schema 2, the property $G$ is being predicated of something.

4 According to Thesis 2, instances of Schema 2 presuppose:

a that the property $G$ just is the property of being mentally ill; and

b that there is at least one entity that has the property of being mentally ill.

Just as there are a variety ways to oppose moral realism, there are equally diverse strategies for endorsing anti-realism in psychiatry. One may choose to deny Thesis 1 as does Scheff. Alternatively, one may deny Thesis 2 by endorsing either a Szaszian style of argument or cultural relativism.
1.3 Scheff’s Labeling Theory

Thomas Scheff (1966/1999) proposed that uttering instances of Schema 2 was a way of engaging in social control. As he understands it, systems of social control exert pressure on members of a group to conform to social norms. According to Scheff, each society has ‘residual norms’, which are unspoken or unwritten rules, to which conformity is expected within the community. In Scheff’s favoured usage, deviance from residual norms is a violation of such rules. The violation of some residual norms is more offensive than others, and the violation of residual norms that are deemed particularly worthy of a high degree of respect are often at issue in judgments of mental illness. For example, staring at someone’s left ear while conversing with her is a violation of a residual norm governing proper conversational conduct within Western society. However, such conduct, though unusual, would be insufficient to motivate one’s societal co-members to treat one as pathologically deviant. In contrast, habitually exposing one’s genitals in public places would be a violation of a residual norm, and it is serious enough that someone who engaged in such behaviour would be treated by his co-members as someone who deviated from an important residual norm. When a member \( M \) of society \( S \) deviates from the particularly important residual norms of \( S \), \( M \)’s co-members resort to tactics of social control. Such tactics include stigmatization, segregation, and labeling.

It is not easy to identify and distinguish from each other the acts of stigmatizing, segregating, and labeling a social deviant. Perhaps, presenting an example would be the best way to illustrate how these three kinds of social responses to deviance interact. Consider for example, a prostitute whom I shall call ‘Amber’. Amber’s selling her body and sexual services is looked upon disfavourably within American society. Within this social climate, when one calls Amber ‘a prostitute’ or ‘a hooker’, one is effectively labeling her as a socially deviant member of American society. Moreover, the label ‘prostitute’ isn’t morally neutral in the same way that ‘librarian’ is. That is, by virtue of attaching the label ‘prostitute’ to Amber, one is stigmatizing her and expressing moral opprobrium. One is expressing disapproval of what she does and of the kind of person she must be in order to be working in the sex industry. In addition to stigmatizing Amber by labeling her a ‘prostitute’, one succeeds in socially segregating her. First, it’s a way of putting social distance between those of us who righteously and innocently abstain from the lurid sex industry and those who are actively engaged in that industry. In addition to stigmatizing Amber by labeling her a ‘prostitute’, one succeeds in socially segregating her. First, it’s a way of putting social distance between those of us who righteously and innocently abstain from the lurid sex industry and those who are actively engaged in that industry. Second, once Amber is labeled as such and accepts it, she is apt to congregate with others who share the same label, and she and her co-
horts will form a subculture of their own which will exist and operate at the margins of American society.

Similar considerations apply to the label 'mentally ill'. Scheff (1966/1999) writes:

When labeling first occurs, it merely gives a name to the rule-breaking, which has other roots. When (and if) the rule-breaking becomes an issue and is not ignored or rationalized away, labeling may create a social type, a pattern of 'symptomatic' behaviour in conformity of the expectations of others. Finally, to the extent that the deviant role becomes part of the deviant's self-conception, his ability to control his own behaviour may be impaired under stress, resulting in episodes of compulsive behaviour. (92)

...Much primary deviance is of short duration or of little significance in the life of the bearer. But when people become aware of their label, they may come to play the role of the mentally ill, at first inadvertently, but later, perhaps involuntarily. In other words, a group of symptoms may be stabilized, through self-consciousness and reaffirmation by others as a 'career' of mental illness. (158)

According to Scheff, when we observe someone, say, Zach, whose behaviour strikes us a deviant in the favoured sense, we attach to him the label 'mentally ill'. When we utter the sentence, 'Zach is mentally ill', we are not predicating of him any particular property, let alone the property of being mentally ill. Suppose Zach sincerely reports hearing voices that tell him to hurt himself, believing that his thoughts are being broadcast on the evening news and believing himself to be the Son of God. Suppose further Zach's behaviour is consistent with these beliefs. He cuts himself whenever he has access to sharp objects. He is nervous, agitated, and fearful when his roommate watches the evening news, and he claims to have special powers and privileges by virtue of being the Son of God. For Scheff, these eccentric behaviours count as primary deviance because they are not performed as a result of having been labeled 'mentally ill'. Presumably, self-harm, grandiose religious beliefs, and beliefs about thought broadcasting violate some social norms. When these violations are not deemed trivial enough to ignore or to explain away, we say, 'Zach is mentally ill.' In doing so, we are not predicating of Zach the property of being mentally ill. Our utterance, 'Z is mentally ill,' is an act of labeling. We have singled Zach out, expressed our moral
and social disapproval of his conduct, and initiated steps towards segregation of him from the rest of us ‘sane’ members of society. We cast Zach in a certain social role, viz. that of the mentally ill outsider who will either live at the margins of society in an uninstitutionalized way, or better yet, as a patient in a mental health facility.

According to Scheff, abnormal behaviour is not symptomatic of an abnormal mental condition, but rather, it is the result of the imposition of a social role on Zach through our acts of the labeling him ‘mentally ill’. Initially, Zach’s abnormal behaviour is merely eccentric or different. When Zach’s behaviour is finally labeled, Zach is now ‘typecast’ in his role as an insane person. The initial symptoms that first earned Zach his label fade in importance once the label has been attached to him. More importantly, Zach begins to act out the role in which he has been cast. He will more insistently and convincingly report command auditory hallucinations, grandiose beliefs, and other psychotic symptoms, not because he has a mental illness, but because he has grown into the role to which society has assigned him.

Scheff’s view counts as a form of anti-realism about psychiatric illness because he is effectively rejecting presuppositions 3(a) and (b) associated with Thesis 1. According to Scheff, when one utters, ‘Zach is mentally ill,’ one is not predicating a property $G$ of him such that it could be true or false that Zach is $G$. Rather, such an utterance does the work of attaching a label, viz. ‘mentally ill’, to Zach. Labeling is an act, the result of which is to cast Zach in a certain social role that he did not previously occupy. Behaviours such as endorsing auditory hallucinations or reporting grandiose beliefs are merely odd and unusual when exhibited by someone who hasn’t yet been labeled ‘mentally ill’, but these same behaviours become more pronounced, stable, and stereotyped when exhibited as part of one’s playing the role of the mentally ill. For Zach, such behaviours didn’t become symptoms until he was labeled ‘mentally ill’. Contrary to what surface grammar might suggest, the predicate ‘is mentally ill’ does not refer to a property, nor does the expression ‘mental illness’ refer to a psychological condition. Symptoms aren’t the effects of a physical or psychological condition which answers to the term ‘mental illness’. There is no such condition. Symptomatic behaviour is just a matter of behaving in accordance with the social role in which one has been cast.

As an account of psychiatric illness, Scheff’s view has two virtues. First, it is ontologically austere. He explains psychiatric symptoms by appealing to behavioural and social facts. As such, he doesn’t need to postulate any controversial entities, such as psychological or moral facts, in order to explain these phenomena. Second,
it can explain why those who carry a psychiatric diagnosis often are stigmatized by their societal co-members. It draws attention to how the mentally ill are singled out for (mis)treatment, and it could serve as a starting point for social change towards better, more tolerant, and more compassionate treatment of persons who strike us as odd or eccentric.

However, Scheff also grossly oversimplifies psychiatric illness and seems to under-appreciate a great deal of empirical and clinical data concerning the etiology and therapeutic treatment of mental illness. First, social patterns of labeling become implausibly powerful. Scheff would have us believe that people who have been diagnosed with a mental illness exhibit bizarre behaviours because we have cast them in that role by attaching the label ‘mentally ill’ to them and because they either involuntarily or voluntarily conform their behaviour to our expectations of someone carrying such a label. Had we not labeled them as mentally ill, they may still exhibit some of the initial mild symptoms that had inspired us to apply the label in the first place, but they would not be presenting the flagrantly pathological behaviour that constitute their playing the role of the mentally ill. There are some possible mechanisms by which symptomatic behaviours and experiences are caused or exacerbated by diagnosing a patient with a mental illness. For example, suppose that a patient, say Yves, is socially withdrawn due to fears of embarrassing himself in the eyes of others and to fears of subsequent social rejection. Suppose that Yves’s social anxieties are moderate and that, although he is able to hold a job and to make a few friends, he is distressed by social interactions at his work place and by his inability to widen his circle of friends. Suppose that Yves is diagnosed with a psychiatric illness, viz. social anxiety, and that he subscribes to certain beliefs about society’s negative attitudes towards the mentally ill. Since Yves believes that he will be devalued and rejected because he carries a psychiatric diagnosis, his social anxieties may very well worsen post-diagnosis, and he may even begin to exhibit more symptoms of social awkwardness or isolation. In this case, Yves was arguably only playing the role of the mentally ill partially pre-diagnosis because he was able to function at some minimal level though it was a struggle for him. Post-diagnosis, however, Yves seemed to take on the role of the mentally ill more completely as his symptoms became more severe and broader in scope. There is evidence indicating that in some cases attaching a diagnostic label to a patient can cause a patient to lose self-esteem. (Link, Struening, Neese-Todd, Asmussen & Phelan 2001) However, the available evidence shows that post-diagnostic loss of self-esteem is observed only in some patients, viz. those who had certain background beliefs about social stigmas associated with mental illness.
There is still a significant portion of the patient population that does not report loss of self-esteem. Moreover, Scheff’s labeling theory cannot account for single episodes of mental illness in which a patient becomes acutely symptomatic, is diagnosed and treated, and does not experience a recurrence of the symptoms. In these cases, the patients play the role of the mentally ill less and less well until their experiences and behaviours stabilize at or close to a normal baseline. Thus, Scheff’s claims about the role of labeling in the progression of symptomatic presentation accurately describes only a limited number of cases.

Second, Scheff’s account also fails to provide an adequate explanation for misdiagnoses or changes in psychiatric diagnosis. Consider, for example, patient Jones who has been diagnosed with bipolar disorder because she has had depressive episodes for which she has been hospitalized in the past and because she also reports periods during which she is elated and restless. Jones carried such a diagnosis for years, and during a recent depressive episode, Jones is readmitted. During her most recent admission interview, Jones reports that she is aware of her diagnosis of bipolar disorder and she endorses it. Jones continues to endorse past symptoms that are vaguely reminiscent of mania, which would qualify her for a diagnosis of bipolar disorder. However, more careful interviewing reveals that the episodes that she reports using words such as ‘manic’ do not qualify as genuine manic episodes because they do not meet the clinical criteria for them. Although she has experienced periods of restless and increased goal-directed activity, such episodes never last longer than a few hours. It seems, then, that Jones was misdiagnosed years ago, and although she carried a diagnosis of bipolar disorder and self-identified as a bipolar patient, she never had this disorder though she probably had some other mental illness. As reasonable an explanation as this might seem, it is not one that is available to Scheff. Since cases like Jones’s are surprisingly common in clinical practice, Scheff owes us some account of her misdiagnosis.

Perhaps Scheff could explain cases of misdiagnoses as cases of poor fit between the label and the person labeled. He could argue that some roles suit certain deviant persons better than others. In Jones’s case, the role of bipolar patient just didn’t suit her very well, and a different role fit her better. However, such a response is inadequate as it stands. One would want to know why some kinds of ‘mental illness roles’ suit some people better than they suit others and why there are still different people who don’t suit any of these roles. Scheff’s theory seems to get things backwards. Contrary to what his account suggests, many patients don’t adapt their behaviour or symptoms to suit diagnoses, but rather, patients do or don’t have certain properties that qualify them to a greater or lesser extent for certain diagnoses. The properties in question

18
are what we quite reasonably take to be features of an actual mental illnesses.

1.4 Szasz’s Myth of Mental Illness

Like Scheff, Szasz takes mental illnesses to be deviations from societal norms, but he argues for a different brand of anti-realism. He writes:

The concept of illness, whether bodily or mental, implies deviation from some clearly defined norm. In the case of physical illness, the norm is the structural and functional integrity of the body... What is the norm deviation from which is regarded mental illness? This question cannot be easily answered. But whatever this norm might be, we can be certain of only one thing: namely, that it is a norm that must be stated in terms of psychosocial, ethical, and legal concepts. (Szasz 1960, 114)

If a person does not suffer from an abnormal biological condition, we do not usually consider him to be ill. (We certainly do not consider him to be physically ill.) ...[P]hysicians are trained to treat bodily ills - not economic, moral, racial, religious, or political 'ills'. And they ... are expected by their patients to treat bodily diseases, not envy, fear and folly, poverty and stupidity, and all other miseries that beset man. Hence, there can be no such thing as mental illness. (Szasz 1974/2003, ix)

Roughly, Szasz argues that an essential property of all genuine illnesses is the property of being a deviation from a biological norm, and that since mental illnesses do not have this property, mental illnesses cannot be genuine illnesses. Unlike Scheff’s version of anti-realism, Szasz doesn’t deny Thesis 1, but rather, he rejects Thesis 2. Szasz accepts presuppositions 3(a) and (b) because he believes that the predicate ‘is mentally ill’ picks out a certain property G. However, the property G isn’t the property of being mentally ill, as one might reasonably suspect. Indeed, Szasz believes that there is no such thing as the property of being mentally ill. Instead, the predicate ‘is mentally ill’ refers to the property of deviating from social, ethical, and legal norms N. Similarly, on Szasz’s view, the referring expression ‘mental illness’ does not refer
to a genuine illness, but rather, to a deviation from social, ethical, and legal norms. Let us take a closer look at Szasz’s argument for the rejection of presupposition 4(a), which I have summarized below:

(1) In order for condition \( C \) to count as a genuine illness, \( C \) must satisfy the necessary and sufficient conditions for a genuine illness. [premise]

(2) The necessary and sufficient conditions for a genuine illness are best discovered by extracting such conditions from paradigm cases of such illnesses. [premise]

(3) Physical illnesses are paradigm cases of genuine illnesses. [premise]

(4) The necessary and sufficient condition for a physical illness is deviation from a biological norm. [premise]

(5) Condition \( C \) is a genuine illness if and only if \( C \) is a deviation from a biological norm. [1 - 4]

(6) For all mental illnesses \( M \), \( M \) is not a deviation from a biological norm. [premise]

(7) No mental illnesses are genuine illnesses. [5,6]

When Szasz discusses deviation from biological norms, he most plausibly has in mind certain biological properties, the expression of which can be averaged over a population such that one can locate deviations from the population average. For example, height is a property that admits of degrees, and all members of the human population have a measurable height. Measurement of the height of all members of the human population would yield a distribution of heights with an average of \( H \), and since there is variability in height, there will be persons who are slightly or significantly shorter (or taller) than \( H \). There might be other properties that don’t obviously admit of degrees, such as the property of having a right hand, but arguably, statistical deviation in structure or function, such as abnormally short fingers or inflexible knuckles, would count as deviations in the ways that one might have the property of having a right hand.

We can now see why Szasz thinks that a physical illness just is a deviation from a biological norm. Take for example the normal cell life cycle. Ordinarily, a cell grows and matures until a certain time at which apoptosis, i.e. cell death, is initiated. The cell growth and death occurs within an optimal range of time, but when one’s cell growth is accelerated and out of synch with normal programmed cell death, one has
cancer. It would seem then, that one has cancer if and only if the rate of cell growth, maturation, and death deviates from the average rate of these processes in members of the human population.

However, despite its initial intuitive appeal, there are two problematic premises in the argument. Premise (4) is false. Anatomical or structural deviations seem neither necessary nor sufficient for genuine physical illness. Such deviations aren’t sufficient because there are plenty of statistical deviations that aren’t plausibly classified as illnesses. For example, an adult human who is seven feet tall statistically deviates in from the average height within the human population, but it is not plausible to suppose that such a person is ill. Moreover, a lack of statistical deviation is insufficient for the absence of illness. For example, having dental cavities is quite a common condition, and it is statistically deviant for a person to lack cavities. Yet, one is not ill by virtue of lacking cavities.

In addition to the falsity of premise (4), Szasz helps himself to premise (6) without any substantial philosophical argument in its favour. In doing so, Szasz assumes that psychological conditions, both healthy and unhealthy ones, are not biological conditions. That is, he presupposes the falsity of biologicalism about the mind. While it is perhaps controversial whether biologicalism is true, it is not at all clear that it is false. In addition to the obvious defects in Szasz’s negative account of mental illness, his positive account is also found wanting. Szasz asserts that a condition is a mental illness just in case it is a deviation from psychosocial, ethical, or legal norms, but he doesn’t provide any argument for this claim. It is important to note that Szasz uses ‘deviation from norms’ in his discussion of mental illness differently than he does in his explanation of physical illness. In the latter, he seemed to be using it to mean statistical deviance, but in the former, he seems to be using it to mean violation of social, ethical, or legal rules. In the absence of a supporting argument for his positive view, Szasz’s account must be evaluated as it stands, and it is clear that it can’t be right. Deviation from social, ethical, or legal norms is neither necessary nor sufficient for mental illness. It is not sufficient because some deviations from ethical or legal norms count simply as moral or legal offenses, not mental illnesses. Chronically parking one’s car in non-parking zones constitutes behaviour that deviates from legal norms, but someone who engages in such activity isn’t necessarily mentally ill. He might just frequent parts of town where parking spaces are scarce and be willing to pay parking fines. Similarly, habitually failing to greet one’s neighbours in the morning is a way of flouting a certain social convention within one’s neighbourhood, but such behaviour isn’t plausibly taken to be indicative of mental illness. It is perhaps rude
and unfriendly, but it isn’t insane. Finally, deviation from social, ethical, or legal
norms isn’t necessary for mental illness because one could meet criteria for dysthmia,
a persistent form of moderate depression, without violating any such rules, customs,
or conventions. One need only experience depressed mood and exhibit two symptoms
which may include fatigue and insomnia. It is implausible to suppose that what
counts as a healthy mood level is governed by social, ethical, or legal rules. Moreover,
abnormal energy levels and sleep patterns don’t seem to be violations of these rules,
but rather, physiological symptoms of a psychological disturbance.

1.5 Cultural Relativism

Szasz’s intuition that psychiatric illness is a matter of deviating from social norms is
one that is widely shared, especially amongst anthropologists who are impressed by
remarkable cross-cultural diversity in the recognition of psychiatric illness. The ob-
served variability in psychiatric judgments is sometimes taken as evidence to support
the claim that such judgments are culturally relative. In this section, I shall describe
some examples of cross-cultural disagreement about psychopathology and discuss how
these cases allegedly support cultural relativism about psychiatric judgments. I also
argue that standard arguments do not provide us with sufficiently compelling reasons
to abandon realism about mental illness in favour of cultural relativism.

Silverman argued that shamanism in some primitive societies and schizophrenia in
modern Western society do not differ significantly in their course of development or in
their presenting symptoms. Rather, he claimed that the principal difference is one in
the degree of societal acceptance of an unusual solution to life crises. (Silverman 1967)
There is some disagreement within the anthropological community about what con-
stitutes a shaman, but for my purposes, a shaman is a medicine man who believes
that he communicates directly with spirits through abnormal perceptual experiences
and whose societal co-members share this belief. A shaman is accorded social pres-
tige within his community, and his or her help is often sought in efforts to heal
co-members who are allegedly afflicted with physical, behavioural, or psychological
maladies. By contrast, according to modern Western psychiatry, schizophrenia is an
illness which may be manifested in any one of a few distinct syndromes. Nonetheless,
some of the cardinal features of the illness include positive and negative symptoms.
Positive symptoms include hallucinations and delusional thinking. Some typical neg-
ative symptoms include emotional or social withdrawal, blunted affect, poor rapport, and attentional impairment. Silverman claims that these symptoms, which are taken by Westerners to indicate psychopathology, can serve an important social function within another community which is not inclined to pathologize these behaviours or experiences.

Silverman suggests that the pre-shamanic and pre-schizophrenic development are parallel until a crucial point is reached. Some pre-schizophrenics have a personal history of unresolved trauma which is intensely anxiety-provoking. The pre-schizophrenic’s inability to extricate himself from his situation leads to intolerable psychological disharmony. Since pre-schizophrenics are often hypersensitive and introverted, they cope with intolerable conditions through emotional and social withdrawal, non-reality-oriented ideation, and hallucinations. Similarly, pre-shamanics often have introverted tendencies, and the ‘shamanistic call’ may come through misfortune, crisis, or protracted illness. Silverman argues that although schizophrenic onset and the reception of the shamanistic call represent approximately comparable points within the course of illness and spiritual development, respectively, the path of the schizophrenic and that of the shaman diverge thereafter. The experiences of the budding shaman are highly valued, encouraged, and rewarded. His experiences and behaviours are legitimized, and he is accorded a social role that commands the profound respect of his societal co-members. In contrast, the similar experiences of a schizophrenic are not deemed acceptable within Western society. He is stigmatized by virtue of being diagnosed with a mental illness, and depending on how intolerable his behavioural deviance is to his societal co-members, he may be involuntarily cast into the sick role.

Although Silverman’s data concerns only schizophrenia, other data suggest similar cultural differences in the acceptability of psychiatric symptoms. For example, Obeyesekere claims that many of the distinctive features of depression are not taken by Buddhists as indicative of pathology. (Kirmayer 1994, Obeyesekere 1985) Instead, from a Buddhist perspective, these symptoms suggest a profound appreciation of the inevitability of grief, suffering, and loss. While Eastern Buddhists would regard an individual with depressive symptoms as someone who has acquired deep spiritual insight, Western mental health professionals would classify the same individual with the same presenting syndrome as mentally ill and in need of medical treatment.

Not only are there cultural disagreements about whether a certain battery of behaviours and experiences are disordered syndromes, but there are cases in which cultures agree that an observed syndrome is indicative of an illness while disagreeing
about what kind of illness is present. In Japan, *taijin kyofusho* (TKS) is a commonly diagnosed disorder whose principal symptoms include fear and avoidance of social situations and interpersonal hypersensitivity. These symptoms are typically associated with social phobia as it is understood by American mental health professionals. In North America, social anxieties revolve around fear of public embarrassment or humiliation, but in Japan, they are driven by fear of offending others with one’s ugliness, body odour, eye contact, or social behaviour. In collectivist societies like Japan, one’s conception of one’s self is determined by harmonious interpersonal relationships, and one’s self interests are subordinated to those of others. As such, TKS is an exaggeration of culturally normative concerns about interpersonal relationships. Japanese psychiatrists understand TKS to be a fundamental impairment in the patient’s conception of himself and his social relatedness, and it is accorded central importance in psychiatric diagnosis. As such, Japanese psychiatrists treat patients with delusions, anxiety, and avoidant behaviours as having a single disorder, viz. TKS, but Western psychiatrists treat each of these syndromes as fundamentally distinct kinds of psychiatric problems. (Kirmayer 1994, Millon, Grossman, Millon, Meagher & Ramanth 2004) Thus, although Japanese and Western psychiatrists would agree that a patient with exaggerated fears about his ugliness and its effect on others has a psychiatric disorder, they would disagree about whether the patient suffered from somatic delusions which may qualify him for a diagnosis of psychosis or whether he suffered from TKS which is not a psychotic disorder per se.

From cases like the ones just presented, it is suggested that psychiatric judgments are culturally relative. Suppose a person *P* exhibits a syndrome *S* which includes non-reality-oriented ideation, bizarre mannerisms, and hallucinations. On the basis of observing *P* exhibiting *S*, members of culture *C₁* judge that *P* is mentally ill while those of culture *C₂* judge that *P* is not mentally ill. Disagreement between the two cultures is evidence that there is no such thing as the property of *being mentally ill*. Rather, the predicate ‘being mentally ill’ actually refers to the property of *being mentally ill according to the standards of culture C*, where it is contextually salient which culture’s standards are in question. As such, cultural relativism about psychiatric illness is a form of anti-realism about these maladies because it rejects presupposition 4(a).

A crucial step in this line of thought is the inference from the fact that cultural groups disagree about whether a certain state of affairs obtains to the claim that whether a certain state of affairs obtains is determined by culturally variable opinions. A proponent of cultural relativism begins with a recognition of a disagreement between
two cultural groups. For example, culture $C_1$ judges that person $P$ is mentally ill, and culture $C_2$ judges that $P$ is not mentally ill. From the fact that these two judgments differ, the cultural relativist infers a claim about what does and doesn’t exist, viz. (i) that $P$ is mentally ill according to $C_1$’s standards; (ii) that $P$ is not mentally ill according to $C_2$’s standards; and (iii) $P$ can’t have or lack the property of being mentally ill (simply) because there is no such property.

However, this inferential step is problematic because it attempts to derive a claim about how things actually are from a claim about people’s opinions about how things are. A parallel line of argument could be constructed to support biological relativism. Evolutionary biologists believe that life on earth began when certain conditions obtained within a primordial soup. Creationists believe that life on earth began when God created the universe and its inhabitants, not when a postulated primordial soup happened to acquire certain properties. From these two facts, it would follow (i) that life on earth began with a primordial soup according to biologists’ standards; (ii) that life on earth didn’t begin with a primordial soup according to creationists’ standards; and (iii) that there is no fact of the matter about whether life on earth began with a primordial soup or not independently of biologists and creationists’ opinions on the matter. Nobody would accept such an argument for biological relativism, yet it has the exact same structure as the one offered in support of cultural relativism. The invalidity of the former is more striking than that of the latter, and as such, it serves to direct our attention to precisely where the latter goes wrong.

Thus far, all I’ve shown is that a certain line of reasoning for cultural relativism is invalid, but I haven’t established that there is no valid argument for this view. Although I don’t have a knock-down argument against cultural relativism, I would like to point out some of the consequences of holding such a view. I would like to show that a commitment to cultural relativism comes at a very high price and that friends of this view haven’t given us sufficiently compelling reasons to think that this account is clearly superior to realism. Three of the problems that I wish to discuss are: (i) the evidence for cultural variability is not as clear, strong, or uncontroversial as Silverman would have us believe; (ii) cultural relativism provides no basis for genuine disagreement; and (iii) cultural relativism provides no basis for both inter- and intra-cultural criticism.
1.5.1 The Ambiguity of Available Evidence

I describe here anthropological data which suggest that there are cross-cultural commonalities in the identification and classification of mental illness. I argue that consideration of all available data leads to some doubts about the strength of the claim that there are deeply divisive cultural disagreements about whether mental disorders are present when different cultural groups observe the manifestation of the same gross symptoms.

Jane Murphy used anthropological information about Siberian Inuit and Nigerian Yoruba to determine whether these non-Western cultures identify psychosis as abnormal, and if so, how they differentiate psychosis from shamanism. (Murphy 1978) For the Inuit, their concept of nuthkavihak corresponds roughly to American concepts of crazy, out of one’s mind, or getting wild. The Inuit people apply nuthkavihak to individuals who exhibited abnormal behaviours that included ‘screaming at somebody who doesn’t exist; believing oneself to be an animal; refusing to talk; getting lost; hiding in strange places; drinking urine; or going on a spree of killing dogs.’ (Murphy 1978, 6) The Yoruba have a similar concept, were, which applies to behaviours such as ‘hearing things other people do not hear; laughing to oneself; talking all the time; asking questions and then answering them oneself; tearing off one’s clothes; setting fires; and defecating in public and then mushing around in the feces.’ (Murphy 1978, 6) Moreover, a person was not considered nuthkavihak or were by an Inuit or Yoruban, respectively, unless they exhibited more than one of these behavioural symptoms. That is, both Inuits and Yorubans were only inclined to attribute mental illness to someone if they exhibited certain abnormal patterns of behaviour, not just particular, isolated mannerisms.

Murphy’s data and her interpretation of them are consistent with work reported by Edgerton. (Edgerton 1966) Edgerton interviewed members of the Kamba, Hehe, Pokot, and Sebei tribes in Eastern Africa. He asked them three questions: (1) ‘How do psychotics behave?’; (2)‘How else?’; and (3) ‘Have you actually seen a psychotic right here in [name of area]? How did this person behave?’). All four tribes seemed to distinguish between two kinds of madness: a violent kind and a more docile kind. All four tribes considered violently aggressive acts of murder and assault indicative of mental illness. Amongst the Sebei, the most common symptoms of docile madness were ‘going naked’, shouting and screaming, talking nonsense, wandering, eating dirt, and collecting trash. Similarly, the Kamba often mentioned ‘going naked’. For the
Hehe, nudity, eating and smearing feces, talking nonsense, and living alone in a bush are signs of mental illness. Finally, the Pokot most commonly mentioned talking nonsense and arson as symptoms of psychosis. Hallucinations were rarely mentioned by any of the tribe members, but it is unclear whether this is so because it is difficult for tribe members to discern whether their psychotic co-members are hallucinating or whether psychotic tribes people just don't hallucinate. Edgerton does note that patients in East African mental hospitals do hallucinate, and unless we are willing to believe that they learned to do so only after hospitalization, it is reasonable to infer that they hallucinated before admission for medical treatment.

Both Murphy and Edgerton's findings suggest that there is a great deal of overlap between what counts as a psychotic syndrome amongst disparate non-Western societies, but they seem at odds with Silverman's arguments concerning shamanism. Murphy studied shamanism amongst the Siberian Inuit, but did not pursue a similar inquiry into Yoruban shamanism because Yoruban healers are not typically discussed in the anthropology literature as likely to be insane. Murphy describes in her field notes a shamaness performing a healing ritual:

When my brother was sick, my grandmother who was a shamaness tried her best to get him well. She did all her part, acting as though a dog, singing some songs at night, but he died. While she was singing, she fell down so hard on the floor, making a big noise. After about fifteen minutes later, we heard the tapping of her fingers. Slowly, she got up, already she had become like a dog. She looks awful. My grandfather told me that he used to hide his face with his drum just because she looks different, changed and awful like a dog, very scary. She used to crawl back and forth on the floor, making noises. Even though my brother was afraid of her, he tried not to hide his face, he looked at her so that he would become well. Then my grandmother licked his mouth to try to pull up the cough and to blow it away. Then after half hour, she fell down so hard on the floor again. (Murphy 1978, 7)

Despite behaving as though she were a dog, the shamaness was not considered nuthkavihak or crazy by her tribal co-members. From interviewing the Inuit about their shamans and shamanesses, Murphy found that tribe members would acknowledge that a healer was out of his or her mind when he is curing, but that he or she is
not crazy. Murphy interpreted this and similar claims to indicate that, for the Inuit, a madman presents a pervasive pattern of abnormal behaviour, but a shaman only exhibits part of such a pattern. That is, the latter may exhibit one symptom or another some of the time, viz. during healings or seances, but not enough of symptoms enough of the time to be deemed crazy. If Murphy is right, then shamanism is not glorified psychosis or schizophrenia. Both Western and non-Western societies are only willing to attribute psychosis to individuals who exhibit a pattern of abnormal symptoms which is at or above a certain threshold level, both in time and in expression. The only difference between Western and non-Western societies is their treatment of individuals who exhibit sub-threshold behavioural deviations. The former consider such people odd or eccentric, but not pathologically so, and it may marginalize such persons by attaching to them labels such as 'freak'. The latter reveres such individuals and accords them especially prestigious social status within their communities. However, this difference is not one that poses any threat to the universality of psychiatric classification because it is a difference between how sane individuals are regarded and treated within their respective societies.

While Murphy and Edgerton's works suggest that cultural variability is perhaps exaggerated, they are not enough to show that there is no divergence at all between psychiatric judgments in one culture and those in another. That is, certain core features of a psychiatric syndrome may provide a basis for universal agreement about the presence of a mental illness, but not all syndromes will elicit cross-culturally uniform psychiatric judgments. Thus, it is not entirely clear whether anthropological data strongly support a claim about cultural variability or one about cultural uniformity. Since cultural relativists build the case for the view on a claim about cultural variability, the force of their argument is diminished by the lack of clear evidence for this crucial claim.

1.5.2 The Disappearance of Genuine Disagreement

I now consider the status of cultural disagreements from the perspective of cultural relativism. I argue that on this view, there can be no genuine disagreements about the presence or absence of mental illness. I argue further that this is a serious disadvantage of this account.

Let us take a closer look at how genuine cross-cultural disagreements vanish if cultural
relativism is right. Let us take sentence 1 to express the judgment of a Westerner and sentence 2 to express the opinion of a non-Westerner.

**Sentence 1** Anyone who exhibits non-reality-oriented ideation, abnormal perceptual experiences, social withdrawal, and bizarre mannerisms is mentally ill.

**Sentence 2** Anyone who exhibits non-reality-oriented ideation, abnormal perceptual experiences, social withdrawal, and bizarre mannerisms is not mentally ill.

Under ordinary conversational circumstances, when sentence 1 is uttered by a modern Westerner and sentence 2 by a primitive non-Westerner, there would seem, on the face of it, to be a genuine disagreement between the two speakers precisely because we assume that there is a fact about whether the symptoms mentioned are indeed indicative of mental illness. As such, one of the speakers utters a truth, and the other a falsehood. The disagreement may not be readily resolved because one group does not know the relevant fact and the other can’t easily make that fact known. However, there is no reason to think that all facts must be known by all parties who might have an opinion about them.

If cultural relativism is right, facts about mental illness are represented by sentences with the form of sentence 3. Recall that cultural relativism is a form of anti-realism according to which presupposition 4(a) is false. (See Section 1.2.) That is, the predicate ‘is mentally ill’ doesn’t refer to the property *is mentally ill*, but rather, it refers to the property *is mentally ill according to the standards of culture C*. If we reframe the argument between the Westerner and the non-Westerner on the supposition that cultural relativism is true, sentences 1 and 2 express propositions that are more precisely expressed by sentences 4 and 5, respectively.

**Sentence 3** Anyone who exhibits syndrome $S$ is mentally ill according to the standards of group $G$.

**Sentence 4** Anyone who exhibits non-reality-oriented ideation, abnormal perceptual experiences, social withdrawal, and bizarre mannerisms is mentally ill according to the standards of Western culture.
Sentence 5 Non-reality-oriented ideation, abnormal perceptual experiences, social withdrawal, and bizarre mannerisms is not mentally ill according to the standards of non-Western culture.

Let us suppose that cultural relativism is true. When sentences 4 and 5 are uttered by a Westerner and a non-Westerner, respectively, they now each speak truly. It’s true that according to Western standards the mentioned symptoms are devalued and associated with psychiatric illness, and it’s equally true that according to non-Western standards they aren’t devalued or associated with mental illness. There is no genuine disagreement between the two speakers because there is no fact of the matter about whether these symptoms really are associated with mental illness.²

Cultural relativists and realists have two different explanations for what is happening when two cultural groups differ in their psychiatric judgments. The latter take disagreements at face value because they presume that there is a fact of the matter about whether a hallucinating, delusional person is mentally ill. The former must suppose that all disagreements about psychiatric judgments are merely apparent. Although different cultural groups may act and speak as though they are contradicting each other, there is no genuine disagreement, and indeed, they are often talking past each other. Thus, one of the consequences of cultural relativism is that we are all systematically and pervasively mistaken about the import of our (apparent) disagreements. If we had strong, overwhelmingly independent reasons for accepting cultural relativism, living with such a consequence would be forced on us, and perhaps, we would become less invested in cross-cultural pseudo-disagreements about mental illness. However, proponents of cultural relativism have yet to present a convincing case for their view, and in the absence of such, we shouldn’t feel any pressure to accept their view and consequently to deny that cultural disputes about mental illness are genuine ones.

²Moral relativists have made some effort to sketch an account of how one might express disagreement with a relativistic moral language. Psychiatric relativists might model a response to this objection after these proposals. However, it is not entirely clear whether such accounts succeed in enabling us to express genuine disagreement, or whether they simply allow us to talk as though we disagreed. See Harman (1996).
1.5.3 The Impotence of Criticism

Another unacceptable consequence of cultural relativism is that both inter- and intra-cultural criticism would be impossible even when some behaviours and experiences are not mental illnesses at all. The most striking examples include historical embarrassments and contemporary controversies in Western psychiatry. I shall consider two such cases.

In 1851, Samuel Cartwright, a physician from Louisiana published a paper entitled ‘Report on the diseases and physical peculiarities of the Negro race’ in The New Orleans Medical and Surgical Journal. (Cartwright 1851) In it, Cartwright discusses alleged mental illnesses that ‘afflicted’ black slaves, and he recommended treatments which purportedly cured these slaves of their maladies. One of the conditions presented by Cartwright is drapetomania which allegedly afflicted slaves who fled from their owners. Cartwright argued that this was a treatable medical disorder, for which he prescribed whipping. Cartwright also discussed dysaethesia aethiopica, an alleged disorder that afflicted slaves who exhibited lassitude and lack of motivation, and he suggested that whipping would cure slaves of this illness as well.

Within the cultural context of nineteenth century Louisiana, the slavery of black persons was considered the norm. Thus, when slaves deviated from the norm by seeking freedom or refusing to play the social role of a slave, these behaviours struck Cartwright and his white Louisianan contemporaries as deviant. Moreover, since slave-owners had a vested interest in the continued enslavement of black individuals and some even believed that they were improving the lives of their captives, such deviant behaviour was socially devalued and perhaps even incomprehensible.

If cultural relativism was right, there would be no way for anyone to criticize Cartwright or any slave-owner who administered the prescribed ‘treatments’ to slaves bearing the diagnosis of drapetomania or dysaethesia aethiopica. According to cultural relativism, when Cartwright claims that slaves who flee from captivity are mentally ill, he is really saying that slaves who flee from captivity are mentally ill according to his culture’s standards. Likewise, a critic of Cartwright, who claims that these slaves are not mentally ill, is actually saying that they are not mentally ill according to the critic’s social group’s standards. Both Cartwright and his critic make true claims, and since there is no fact that would make it the case that Cartwright is mistaken, Cartwright’s critic is in no position to show that Cartwright is wrong. There is no basis on which to distinguish between cases in which individuals deviate from unjust
societal norms and cases in which individuals suffer from a genuine mental illness. Furthermore, if cultural relativism were true, there would be no measure of progress for the classification of psychiatric illness. A taxonomy that excludes drapetomania as a mental illness is no less distorted than one that includes it. Thus, when the former replaces the latter, there is no basis on which to claim that such a change is an improvement.

More recent controversies concerning American psychiatric taxonomy are primarily debates about the utility, reliability, and validity of the Diagnostic and Statistical Manual of Mental Disorders (DSM). The DSM is a catalogue of mental disorders that is published by the American Psychiatric Association. The DSM-I was first published in 1952, and it has since undergone three major revisions (DSM-II, DSM-III, and DSM-IV). Although there is an ongoing debate about whether the DSM is an adequate classificatory system for mental illnesses, it is the closest approximation to a complete taxonomy of psychiatric conditions. It is the only document in widespread clinical use that has been endorsed by the mental health profession at large. It sets the standards for diagnosis, research, and insurance reimbursement. As such, some of the debates about past and present DSM categories are germane to a discussion of disagreements about psychiatric categories within American society.

In response to criticisms from gay activists, homosexuality per se was deleted from the DSM-II, the first revision of the DSM, and hence, it was no longer classified as a mental disorder. (Spitzer 1981) It was replaced by a category that would later become ego-dystonic homosexuality in the DSM-III. Persons with the view that homosexuality is invariably pathological considered this a victory of politics over science, and persons who considered homosexuality a normal variant of human sexuality were also dissatisfied because the criteria for ego-dystonic homosexuality included: (i) a persistent lack of heterosexual arousal, which the patient experienced as interfering with initiation or maintenance of wanted heterosexual relationships, and (ii) persistent distress from a sustained pattern of unwanted homosexual arousal. According to gay-friendly critics of the DSM, criterion (i) suggests that, in at least some cases where an individual ‘complains’ of absent heterosexual arousal and a sustained homosexual arousal patterns, an appropriate therapeutic activity is to encourage that person to develop a heteronormative sexual arousal pattern, and not merely to become more comfortable with his or her own homosexuality. Such critics also argued that psychological problems that are related to ego-dystonic homosexuality could be treated by other general diagnostic categories, and the existence of a special diagnosis singling out homosexuals reflected antigay prejudices and perpetuated an antigay stigma. In
response to these criticisms, the DSM was further revised, and in 1986, ego-dystonic homosexuality was removed as a diagnostic category. Presently, the only vestige of this putative disorder is a category labeled ‘Sexual Disorders Not Otherwise Specified’, according to which an individual experiences ‘persistent and marked distress about sexual orientation.’ (APA 2000)

For those who sympathize with cultural relativism, the controversy over homosexuality’s status as a psychiatric illness is taken as evidence that arbitrary decisions based on personal or social prejudices determine whether a condition counts as a mental disorder. In pre-DSM-II American society, heterosexuality was a sexual norm, and homosexuality was a devalued deviation from the norm. In post-DSM-II American society, this is no longer the case. As in the case of dismissal of drapetomania as a mental illness, the declassification of homosexuality as a psychiatric disorder could not be considered a mark of progress under the presumption of cultural relativism. There would be no sense in which the current version of the DSM-IV, better represents actual psychiatric categories or is less contaminated by personal prejudices, at least in its treatment of sexual orientation. If this were true, however, it would be hard to understand why we should be at all motivated to accept the DSM-IV over the DSM-I because according to cultural relativism, presupposition 4(a) is false (See Section 1.2.), i.e. there is no fact that makes it the case that the DSM-IV better represents what counts as a psychosexual disorder than the DSM-I.

Just as cultural relativists and realists offer different explanations of disputes about psychiatric judgments, they disagree about what the DSM is actually documenting. Under the presumption of cultural relativism, revisions of the DSM are, at best, merely reports documenting social attitudes as they change over time, and the DSM-IV better describes contemporary attitudes while the DSM-I more accurately records the attitudes of American society in the 1950s. However, realists take the DSM at face value, and they do not consider it a historical or sociological document. Rather, they take the DSM to be an attempt at a psychiatric taxonomy that best represents actual mental illnesses, albeit one that may be crude, approximate and subject to revision in light of new data.

As we saw in Section 1.5.2, a commitment to cultural relativism comes with certain costs, and in addition to those described earlier, we see here that if this view is taken seriously, the force of cross-cultural criticism of groups, who judge homosexuals or freedom-seeking slaves to be mentally ill, is lost. This is so because, according to cultural relativism, there is no fact of the matter about whether homosexuals or
freedom-seeking slaves are mentally ill. Instead, there are only facts about whether certain social groups judge them to be mentally ill. If we accept cultural relativism, we would not have the leverage to demand revision or reform of grotesque taxonomic schemes that are used as tools for social and political oppression. While this does not show that cultural relativism is false, it is a consequence of a view that we should not embrace without overwhelming independent reasons to do so, and such reasons have yet to be supplied by the proponents of cultural relativism.

1.6 Conclusion

In this chapter, I have surveyed three major anti-realist views about mental illness, viz. Scheff’s labeling theory, Szasz’s myth of mental illness, and cultural relativism. I have argued that although diagnostic labeling can influence a patient’s course of illness in a limited number of cases, Scheff’s theory cannot provide an adequate account of mental illness more generally. I have also argued that Szasz’s version of anti-realism presupposes an analysis of illness that cannot be correct and a controversial theory of mind for which he gives no compelling argument. Finally, I have presented arguments that expose some of the weaknesses and disadvantages of cultural relativism. I have argued that, given the costs of endorsing cultural relativism, we should only accept this view if we are given overwhelmingly persuasive reasons to do so, but that the cultural relativists have yet to present us with such reasons. Thus, problems that plague anti-realist accounts of mental illness are sufficiently serious that we should be hesitant about abandoning realism.

Of course, difficult questions also arise in realism about mental illness. One of the thorniest questions is What is a mental disorder?, or put slightly differently, What kind of property is having a mental disorder? I consider possible answers to this question in Chapter 3. Before we take up these metaphysical questions, I would like to address epistemological worries about whether we could have scientific knowledge of psychiatric properties, like the properties of being schizophrenic or being bipolar. Accordingly, in Chapter 2, I discuss a commonly articulated challenge to the possibility of a scientific psychiatric taxonomy, and I show that, with a proper understanding of scientific inquiry, there is no obvious reason to doubt that a scientific classification of mental disorders is beyond our reach.
2.1 Introduction

In Chapter 1, I argued that we have good reason to take seriously the claims that there are facts about mental disorders and that these facts are independent of our personal judgements about these psychological conditions. Supposing then that this line of thought is on the right track, there remain important questions about the metaphysics and epistemology of psychiatry. In the next chapter, I shall address the metaphysical question that concerns many psychiatric realists, \textit{What kind entity is a disorder?}. Presently, I am primarily concerned with epistemological questions about how we conduct inquiry into mental disorders and how we can arrive at knowledge of them. More specifically, I shall explore some common worries about the epistemology of psychiatric nosology, i.e. the taxonomy of psychiatric disorders. By way of introducing some of these concerns, let us briefly consider some criticisms of the Diagnostic and Statistical Manual of Mental Disorders (DSM), currently the best attempt at a psychiatric nosology developed and published by the American Psychiatric Association (APA).

Between the publication of the first edition DSM-I in 1952 and that of the latest edition DSM-IV-TR in 2000, the number of psychiatric disorders recognized by the APA skyrocketed from a modest 106 to an arguably implausible 400, and not surprisingly, the alarming rate at which psychiatric disorders have multiplied has attracted
the attention of many critics. Kutchins & Kirk (1997) argue that the authors of the DSM-IV-TR are simply creating mental disorders by pathologizing everyday behaviour. They claim that greed motivates the expanding list of recognized disorders for which mental health professionals may legitimately bill insurance companies and for which they may prescribe medications. According to Kutchins and Kirk, the personal financial interests of mental health providers and the commercial interests of pharmaceutical companies introduce pressures and biases into psychiatric nosology so that the boundary between normal and pathological human psychology and behaviour has become distorted.

Not only can greed influence the development of a classification system of mental disorders, but social interests can also be distorting forces. Kaplan (1983) notes that treatment rates for disordered behaviour are higher for women than they are for men, and that these statistics can be explained by gender biases that are reflected in the DSM. She claims that ‘masculine-biased assumptions about what behaviours are healthy and what behaviours are crazy are codified in diagnostic criteria; these criteria then influence diagnosis and treatment rates and patterns.’ (786) According to traditional societal attitudes, women are expected to be more emotional, more excitable in minor crises, less independent, and more submissive. Moreover, these are some of the distinguishing features of Histrionic or Dependent Personality Disorder as they were described in DSM-III and as they are still characterized in DSM-IV-TR. Kaplan suggests that women are diagnosed with one of these two personality disorders when they overconform to feminine stereotypes. Furthermore, healthy women who refuse to play traditional feminine roles are also labeled disordered. Thus, gender biases about what is socially acceptable behaviour in a woman influence which diagnostic categories are included in the DSM and likely distort which conditions are represented as disordered.

Personal or commercial ambitions and social attitudes can distort our representations of mental disorders and can determine which psychosocial conditions are included in, or excluded from, a psychiatric taxonomy. Although the arguments offered by Kutchins and Kirk and by Kaplan may well give us reason to be critical of some of the diagnostic categories proposed in the DSM-IV-TR, skeptics about psychiatric nosology often use similar arguments to establish more generally that a scientific psychiatric nosology is impossible precisely because personal, social, political, and economic values will always and everywhere have a distorting role in the construction of a psychiatric taxonomy. One such skeptic is Kendler (1990) who presented an argument of this kind in his paper ‘Towards a Scientific Psychiatric Nosology’.
In this chapter, I shall argue that Kendler’s argument presupposes an oversimplified understanding of science and that, once a more realistic account of science is accepted, it is clear that a scientific psychiatric nosology is within our reach. In Section 2.2, I present what I call ‘the Traditional View’, which is a slightly outdated, but surprisingly common, account of science, its goals, and how it purports to accomplish these goals. I argue that this is the model of science that Kendler presupposes, and I discuss how understanding science in this way leads him to the conclusion that psychiatric nosology could never be a purely scientific enterprise. In Section 2.3, I present reasons for rejecting the Traditional View, and I argue that, once we accept a more realistic social model for scientific knowledge, like the one proposed by Longino (1990), it is no longer plausible to claim that psychiatric nosology is inevitably unscientific. Finally, in Section 2.4, I argue that we should not feel any pressure to embrace subjectivism or unconstrained pluralism about psychiatric taxonomy just because we have found a place for contextual values in this field.

2.2 Science, Objectivity and Value-Neutrality

On the Traditional View, two propositions about science have a central role: (i) that science is objective in some important sense, and (ii) that this objectivity is ensured by the value-neutrality of science. Indeed, these are taken to be self-evident truths about science which proponents of the Traditional View often assert without further argument. However, in order to understand where the Traditional View of science goes wrong, it is worthwhile to examine more closely propositions (i) and (ii).

2.2.1 Objectivity

According to the Traditional View, science provides us with an objective understanding of the world in two importantly distinct senses of the term ‘objective’. In one sense, scientific theories and explanations are objective insofar as they are approximately true representations of a genuine domain of fact. On this understanding, questions about the objectivity of the domain in question are intertwined with realism about that domain. For clarity, I use the term ‘representational objectivity’ to refer to objectivity in this first sense. We can define representational objectivity thus:
Definition 1 A theory $T$ is representationally objective just in case:

i sentences $S_1, S_2, \ldots, S_n$ which comprise $T$ are meaningful representations of the world; and

ii facts that are independent of our personal opinions or judgements determine whether $S_1, S_2, \ldots, S_n$ are true.\(^1\)

For example, when one says that electronic structure theory is representationally objective, one asserts that it makes sense to claim that there exist electrons with certain physical properties and that this theory describes their actual nature and causal behaviour more or less accurately. While representational objectivity is a feature of theories, explanations, or claims, objectivity in the second sense is primarily a characteristic of scientific inquiry. I use 'methodological objectivity' to refer to objectivity in this latter sense in order to distinguish it from representational objectivity. Following Rosen (1994), we can define methodological objectivity thus:

Definition 2 An inquiry is methodologically objective just in case its trajectory and its products are uninfluenced, or minimally affected, by the peculiar biases, prejudices, ambitions, or ideological commitments of those conducting it.

Since representational objectivity applies to theories, explanations, or claims and methodological objectivity to inquiry, it is clear that they are distinct concepts. But, conventional wisdom about science and objectivity would have us believe that the two are intimately related. This relationship is described in Claim 1.

Claim 1 Theories and hypotheses are representationally objective because inquiry leading to the development of theories and acceptance of hypotheses is methodologically objective.

Since methodological objectivity is thought to be responsible for representational objectivity and since representational objectivity is the principal goal of scientific inquiry, it is of paramount importance that no instance of scientific inquiry involve any factor that could threaten methodological objectivity.

\(^1\)Such facts are sometimes called 'objective facts'.

38
2.2.2 Value-Neutrality

Some of the greatest threats to methodological objectivity are those mentioned in Definition 2, i.e. biases, prejudices, ambitions, and ideological commitments. Personal, social, and cultural interests are considered corrupting pressures that can influence an inquiry, but if this inquiry is to count as a scientific one, these pressures must be minimized, if not completely eliminated. It is common to use the word ‘values’ to refer to the personal, social, and cultural interests that threaten scientific objectivity. Thus, the claim that science is methodologically objective is taken to be equivalent to the claim that science is value-free.

It is important to recognize that the word ‘value’ is often used to mean many different things. Some things are valuable independently of our beliefs and desires. For example, it is arguable that knowledge and beauty are valuable whether or not we judge them to be so. When we use the word ‘value’ narrowly, we use it to refer only to things that are valuable whether or not we actually recognize that this is so and whether or not we actually desire these things. I shall adopt the notation ‘value\text{Narrow}’ to refer to things that are valuable in the narrow sense. In contrast, some people, including many of those participating in debates about the objectivity of psychiatry, use the word ‘value’ much more loosely to refer to anything that we believe to be valuable, anything that we desire, anything that we prefer to other things, and anything in which we take an interest, including things that may not be valuable in the narrow sense. I shall use ‘value\text{Loose}’ to distinguish value in the loose sense from value\text{Narrow}. For example, in a society in which a dominant subgroup will benefit from policies or technology that would contribute to the oppression of a minority group, the interests of the dominant group are values\text{Loose}. In this scenario, the dominant group’s values\text{Loose} may not be values\text{Narrow} because they may be unjust, assuming quite plausibly that justice is a value\text{Narrow}. Thus, in the wider usage of the term ‘value’, we lose the distinction between (i) things that are valuable independently of our opinions about them and (ii) things that we desire or believe to be valuable, regardless of whether they are valuable in the sense described in (i). Indeed, proponents of the Traditional View, which include many of the participants in the debate about the objectivity of psychiatry, presuppose without argument that all things are valuable only insofar as we judge them to be so, and hence, that there is no distinction to be made. For them, all values are values\text{Loose} because values\text{Narrow} don’t exist. This presupposition is summarized in Claim 2. Moreover, such writers are convinced that, because values are just a matter of opinion, judgment, or preference, they are
precisely the kinds of things that reflect biases and prejudices, and hence, that are likely to pose a threat to the objectivity of science.

**Claim 2** Values are nothing more than (i) things we judge to be valuable, (ii) things we desire; (iii) things in which we take an interest; or (iv) things which we prefer to other things.

Of course, Claim 2 is a substantive philosophical proposition, and proponents of the Traditional View owe us an argument to justify their commitment to this claim. Although I think the claim false, I shall not press the point further here, and instead, I will save discussion of the acceptability of Claim 2 for Section 2.3.2. Presently, I shall use the word ‘value’ loosely to refer to our goals, interests, and preferences, some of which may be mere idiosyncratic preferences or idle wishes, but others of which may not.

Even for proponents of the Traditional View, it is not entirely true that science is value-free. Certain norms and values are intimately related to the goals of science, which includes generating theories that have certain features, like representational objectivity and explanatory and predictive power. Features like explanatory power can be associated with still other characteristics that constitute a good explanation, such as simplicity and breadth. Values associated with the declared goals of science are what Longino (1990) calls ‘constitutive values’. Constitutive values guide the formulation of rules that govern scientific practice and method. Of course, when friends of the Traditional View claim that science is value-free, they do not wish to deny that constitutive values may guide the practice of science. Rather, they claim that science is free of what Longino calls ‘contextual values’. Contextual values are those that belong to the social and cultural environment in which science is done. They may include, for example, certain commercial or political interests. Proponents of the Traditional View contend that there is a clear distinction between constitutive and contextual values, that they are separable in principle and in actuality, and that contextual values threaten the methodological objectivity of science.

Even the claim that science is free of contextual values requires further clarification. Longino (1990) distinguishes the attributes of autonomy and integrity. The claim that contextual values do not affect the autonomy of science is equivalent to the claim that scientific inquiry proceeds completely unaffected by social, cultural, political, or moral
interests, and this is false. Science, as it is currently practiced, depends heavily on both public and private funding, and as such, political, social and commercial interests determine to a large extent which scientific questions are posed and pursued. For example, when faced with a scarcity of funding, it is not uncommon to divert funds from projects aimed at understanding obscure features of a disease to those aimed at developing marketable cures. While it is obvious that contextual values can help to determine which questions get posed and subsequently explored by the scientific community, friends of the Traditional View find this kind of role for contextual values rather harmless. They are far more concerned to defend the thesis that contextual values do not affect the integrity of science, i.e. that contextual values do not influence the internal practice of science, including observation, experimentation, theory construction, and induction. Thus, according to the Traditional View, we should also accept Claim 3.

**Claim 3** Any inquiry in which contextual values can influence observation, experimentation, theory construction, or induction is pseudoscience, bad science, or non-science.

The plausibility of Claim 3 rests principally on the belief that if contextual values are permitted to influence not only which questions are worth pursuing in inquiry $I$, but how these questions are answered, then bias, prejudice, and other distorting factors will determine which claims are accepted. Consequently, the methodological objectivity of inquiry $I$ would be compromised, and by Claim 1, we would have no way of ensuring the representational objectivity of the theories and explanations generated by $I$. As such, inquiry $I$ would fail to meet the goals of science and would not be worthy of the name ‘science’.

### 2.2.3 The Scientific Method

According to the Traditional View, the best way to secure methodological objectivity in scientific practice is to require that scientists employ the scientific method when they conduct their studies of the world. On the broadest understanding of it, the scientific method is a set of rules that govern how to determine whether a hypothesis is confirmed by evidence, but in most discussions of psychiatric taxonomy, a positivistic
understanding of the scientific method is presupposed. Historically, positivism about science was a view according to which meaningful sentences are only those whose content is observable or verifiable. Sentences express known propositions just in case: (i) the propositions expressed are true and (ii) our belief in them is justified by experience; or they are derived in a rule governed way from propositions known on the basis of experience. Some of the most basic rules that govern whether a statement expresses a piece of scientific knowledge are those that determine when a hypothesis has been confirmed by empirical data.

In ‘Studies in the Logic of Confirmation’, Hempel (1965b) attempted a purely formal analysis of confirmation. On Hempel’s view of science, its primary goals are the prediction and explanation of physical phenomena. Moreover, a scientific explanation has the structure of a deductive argument where (i) the conclusion is a statement describing a state of affairs to be explained, and (ii) the premises are statements describing initial conditions and covering laws. Just as science was conceived primarily as a deductive enterprise in which there existed a formal relationship between the explanada and the explanans, the confirmation of a hypothesis that may have important explanatory roles was analyzed as a formal relationship between a hypothesis-sentence and sentences describing allegedly confirmatory states of affairs. According to Hempel, just as one could tell whether a conclusion $C$ followed from premises $P_1, P_2, \ldots, P_n$ by considering (i) the form of $C$, (ii) the form of the premises, and (iii) the formal rules of deductive logic, one should be able to determine whether a hypothesis $H$ is supported by a statement $S$ describing a state of affairs that purports to be evidence for it by attending to (iv) the form of $S$, (v) the form of $H$, and (vi) the syntactic relationship between them. Hempel’s account of confirmation included the analysis of two relations: the relation of direct confirmation, and the relation of confirmation (more generally). The latter was analyzed in terms of the first. For our purposes, I shall limit my discussion to the relation of direct confirmation which Hempel characterizes thus:

**Definition 3** An observation statement $S$ directly confirms a hypothesis $H$ if $S$ entails the development of $H$ for the class of objects mentioned in $S$.

The development of a hypothesis $H$ for some class $K$ is what $H$ would assert if there existed only those objects that are members of $K$. For example, suppose that we are entertaining hypothesis $\mathcal{H}$, ‘all bodies falling from rest move at a uniformly
accelerated rate'. Suppose also that we wish to know if $\mathcal{H}$ is directly confirmed by the observation statement $S$, ‘this apple, released from rest at time $t$, moves at a uniformly accelerated rate.’ The development of $\mathcal{H}$ for the class of objects consisting of this apple released from rest at time $t$ is the sentence, ‘this apple, released at time $t$, moves at a uniformly accelerated rate.’ The observation statement $S$ entails the development of $\mathcal{H}$ by being identical to it. Since the relation of direct confirmation is a purely formal one, determining whether acceptance of a hypothesis is justified amounts straightforwardly to determining whether this relation holds between the hypothesis and a report of an observed state of affairs.

A positivistic account of confirmation, such as Hempel’s, illustrates well how a constitutive value might guide scientific practice. Recall that one of the goals of science is the construction of an approximately true theory that is representationally objective. (See Section 2.2.1.) Thus, truth or accuracy is a constitutive value of science. The formal rules of confirmation that Hempel suggests are supposed to be truth-preserving in some (rather loose) sense. That is, if we suppose that the observation statements are true, accurate, or undistorted, then a hypothesis confirmed in the way suggested in Definition 3 is also true, accurate, or undistorted. As we shall see in Section 2.3, it is no simple matter to rid observations or hypothesis acceptance of potential distortions, and the Traditional View does not provide adequate mechanisms by which to do so. Before we turn to arguments purporting to establish this, let us take a closer look at how presupposing the Traditional View of science leads some writers to reject psychiatric nosology as a scientific enterprise.

### 2.2.4 Psychiatric Taxonomy as Pseudoscience

In his article, ‘Toward a Scientific Psychiatric Nosology’, Kendler presupposes the Traditional View of science. He writes, ‘The essence of the scientific method is hypothesis generation and hypothesis testing. Therefore, the critical initial step for a scientific nosology is to form empirically testable hypotheses.’(Kendler 1990, 970) In the case of psychiatric taxonomy, the kinds of hypotheses that are most relevant are those of the form given in Schema 3

**Schema 3** Disorder $D$ is better captured by diagnostic criteria set $A$ than it is by diagnostic criteria set $B$. 
For Kendler, hypotheses of this form are empirically testable just in case they are supported by the typical validators for psychiatric disorders. Kendler uses the term ‘validator’ to refer to a certain kind of evidence. Typical psychiatric validators were first proposed by Robins & Guze (1970), and they include family history, demographic correlates, biological and psychological tests, environmental risk factors, concurrent symptoms, treatment response, diagnostic stability, and course of illness. For example, let us consider family history as a validator for disorder $D$. And, let us consider the hypothesis $H$, ‘Any individual who carries diagnosis $D$ meets criteria $X$, $Y$, and $Z$ and has a significant number of first degree relatives who also meet criteria $X$, $Y$, and $Z$.’ Suppose that our observations can be summarized in statement $S$, ‘Patient $P$ who carries diagnosis $D$ meets criteria $X$, $Y$, and $Z$ and has three siblings, two of whom also meet criteria $X$, $Y$, and $Z$.’ On a positivistic understanding of hypothesis confirmation, as described in Section 2.2.3, hypothesis $H$ is directly confirmed by observation statement $S$. Although only one possible validator for disorder $D$ was considered in this case, psychiatric researchers commonly endorse Assumptions 1 and 2.

**Assumption 1** All validators are equally important for hypotheses concerning any given psychiatric disorder $D$.

**Assumption 2** All validators for $D$ will unambiguously support the hypothesis that a certain diagnostic criteria set for $D$ is superior to all alternatives to it.

Kendler rightly points out that there is no obvious reason to suppose a priori that Assumptions 1 and 2 are true. Indeed, as a matter of fact, Kendler claims that the available evidence about schizotypal personality disorder (SZT) suggests that both assumptions are false.

SZT is a disorder in which sufferers exhibit many schizophrenia-like symptoms without showing any signs of psychosis, i.e. signs of hallucinations or delusions. Sufferers are typically characterized by eccentric behaviour or appearance, magical thinking or superstitiousness, and social isolation. According to Kendler (1985), there are two major traditional schools of thought about the nature of SZT, viz. the familial camp and the clinical camp. Although both camps acknowledge that sufferers of SZT can exhibit a range of symptoms and that precisely which symptoms are present in a given individual with SZT can vary, each camp attempts to identify core defining features.
of SZT and proposes criteria for the best evidence for SZT. According to members of the familial camp, SZT is a primarily, though not exclusively, a familial disorder, and its defining features include the abnormal non-psychotic symptoms exhibited by non-schizophrenic relatives of schizophrenia patients. In contrast, members of the clinical camp hold that SZT is a non-psychotic disorder in which a sufferer presents clinically many stereotypically schizophrenia-like symptoms, regardless of the incidence of schizophrenia in one's pedigree.

Since members of the familial and clinical camps disagree about what SZT is, they also disagree about which validators are the best evidence for SZT. For the familial camp, the most important validator for SZT would be family history, and the accepted diagnostic criteria set for SZT would be the one which best distinguishes relatives of schizophrenics from relatives of non-schizophrenics. According to the clinical view, the key validator for SZT would be the degree to which a proposed diagnostic criteria set agrees with a 'gold-standard' clinician who had extensive experience diagnosing SZT. Assumption 1 fails because two validators have been singled out as being more important than all others in determining which diagnostic criteria ought to be accepted for SZT. Assumption 2 also fails because as a matter of fact, familial data support one diagnostic criteria set while clinical data support a different set. (Kendler 1985, Kendler 1990)

It is also important to recognize that disagreements about the nature of various psychiatric disorders, and derivatively, what counts as good evidence for them, abound. For example, it is not clear whether to regard schizophrenia as primarily a neurocognitive disorder or an emotional disorder. (Aleman & Kahn 2005) Although psychotic symptoms are often the most striking, negative symptoms such as affective flattening are often associated with impaired social functioning which, in turn, strongly influences how well or poorly a schizophrenic can adjust to living independently within a community. If schizophrenia is principally a cognitive disorder, the best validators for schizophrenia would be psychotic symptoms and disordered thought, but if it is primarily an affective disorder, the relevant validator would be blunted or inappropriate affect.

The falsity of Assumptions 1 and 2 troubles Kendler because the available evidence seems insufficient to confirm or disconfirm competing hypotheses, and the acceptance or rejection of a hypothesis is no longer determined by application of the scientific method, as it is traditionally understood. Empirical data alone cannot determine whether (i) a diagnostic criteria set supported by familial data ought to be accepted
for a disorder $D$; (ii) a diagnostic criteria set supported by clinical data ought to be accepted for $D$; (iii) there are two distinct, but overlapping, disorders captured by each of the competing criteria sets; or (iv) there are two distinct, but overlapping, subtypes of $D$ captured by each of the competing criteria sets. According to Kendler, value judgements about whether it is more important to have diagnostic criteria that track treatment outcome, familial aggregation, or course of illness will ultimately determine whether members of the mental health clinical and research communities take (i), (ii), (iii), or (iv) as the appropriate response to the available data. Kendler (1990) writes:

The application of increasingly rigorous empirical methods to the evaluation of nosologic proposals will provide many important benefits to our diagnostic process, especially increased rigour and objectivity. At the same time, we need to guard against exaggerating the potential impact of science on our nosology. Many important nosologic questions in psychiatry are fundamentally nonempirical... [W]e should not in our enthusiasm overlook the inherent limitations of the empirical method. There is a danger that this process will degenerate into pseudoscience, in which we pretend to be ‘objective’ and ‘empirical’ when, in reality, we are making informed value judgements. (972)

Throughout the discussion of his view, Kendler seems to pit the application of empirical methods against the formation of value judgments without distinguishing between constitutive and contextual values. He seems not to recognize that, even on the Traditional View, some value judgments, viz. those involving constitutive values, are permissible in scientific inquiry. Kendler seems to think that raw empirical data and the application of the scientific method, as described in Section 2.2.3, should be sufficient for the construction of a unique set of diagnostic criteria that tracks actual psychiatric disorders. And, he seems unfriendly to the possibility of appealing to constitutive values, besides truth, in deciding between alternatives that are supported equally well by available data. Thus, Kendler’s understanding of science is a particularly restrictive version of the Traditional View since he allows only truth to guide scientific inquiry and excludes contextual and other constitutive values. Of course, if science is understood on such a demanding interpretation of the Traditional View, it is easy to arrive at the conclusion that psychiatric nosology is not scientific in Kendler’s favoured sense. However, someone who endorsed a more liberal understanding of the
Traditional View might be inclined to defend the thesis that psychiatric taxonomy is scientifically legitimate as long as one embraces a more reasonable version of the Traditional View.

While it might be true that in some domains of inquiry, appealing to one or more constitutive values besides truth will be sufficient to arrive at a single theory, explanation, or classification system, this may not always be true. For example, we might try allowing an appeal to simplicity as a scientifically legitimate way to decide between competing classification schemes. Appeal to simplicity alone might enable us to rule out taxonomic schemes which postulate multiple diagnostic categories or subtypes, but it won't be sufficient to help us choose between two competing classification systems in which a single diagnostic category differs. Worse still, one scheme $S$ might be simple in one respect $R$, and another scheme $S'$ might be simple in a different respect $R'$. So, we would need something besides simplicity to decide between $S$ and $S'$. Appealing to yet another constitutive value, say predictive power, is still not likely to yield a taxonomy that is obviously superior to available alternatives because one taxonomic scheme might be better at predicting the incidence of disease within a family while another might be better at predicting the course of illness in an individual afflicted with a certain disorder. Given the complexity of psychiatric disorders and the multitude of ways in which a psychiatric classification system might be simplified or used to predict or explain various kinds of behaviours, courses of illness, treatment outcomes, etc., it is highly unlikely that empirical data, scientific method, and judgments guided by constitutive values will lead to the emergence of a single psychiatric taxonomy that is obviously superior to all other alternatives.

2.3 The Traditional View Rejected

At this point, we are at an important juncture in our own inquiry into the scientific status of psychiatric nosology. Someone who firmly embraces the Traditional View might feel forced to concede that psychiatric taxonomy is unscientific after all. However, I argue in this section that we are not driven to this conclusion because the Traditional View of science is flawed. A commitment to the Traditional View is associated with two assumptions about contextual values and their impact on the scientific status of inquiry. First, friends of the Traditional View assume that any inquiry in which contextual values have a role in observation, experimentation, the-
ory construction, and induction is unscientific. The reader may recognize this as a re-statement of Claim 3 which first appears in Section 2.2.2 and which I re-write below as a reminder. In addition to Claim 3, proponents of the Traditional View also accept Claim 4. That is, they assume without much argument that contextual values are always distorting.

**Claim 3** Any inquiry in which contextual values can influence observation, experimentation, theory construction, or induction is pseudoscience, bad science, or non-science.

**Claim 4** Inquiry in which contextual values guide observation, experimentation, theory construction or induction always leads to distorted theories, explanations, and classification systems.

In Section 2.3.1, I argue that Claim 3 is false because many domains of scientific inquiry require an appeal to contextual values, and the role played by these values does not undermine their scientific legitimacy. In the following section, I challenge the truth of Claim 4, and I discuss two cases in which properly regulated contextual values guide theory construction in a way that minimizes distortions.

### 2.3.1 Is inquiry guided by contextual values always unscientific?

There are three different ways of thinking about the relationship between a hypothesis and evidence for it. First, Hempel found it easiest to provide an analysis of the relation of confirmation as one between two sentences, viz. a hypothesis-sentence and an observation sentence describing a state of affairs that purportedly supported the hypothesis. Hempel preferred to analyze confirmation in this way because he was committed to a certain positivistic conception of science and to a deductive-nomological model of scientific explanation and prediction. However, the failure of the deductive-nomological model of science and the shortcomings of Hempel's logical analysis of confirmation leaves it open to us to discuss the hypothesis-evidence relation in terms that are more familiar and intuitive. The other two ways in which one may think of the hypothesis-evidence relation include the following: (i) a relation between
a state of affairs and a sentence or proposition; and (ii) a relation between two states of affairs. For example, in the first case, one might say that the muddy footprints at the scene of the crime is evidence for the hypothesis that the gardener is the murderer. In the second case, one might say that the line of condensation in the cloud chamber is evidence for the presence of a particle moving through the chamber. The latter two ways of thinking about confirmation are the more fundamental way since it is how we think of evidence in both ordinary and scientific cases of confirmation. In what follows, I discuss confirmation as a relation between a state of affairs and a sentence or proposition.

Longino (1990) argues that whether or not a state of affairs is evidence for a hypothesis $H$ depends on the background assumptions of the individual who is seeking confirmation or disconfirmation of $H$. Her argument proceeds in two stages. In the first stage, she shows that whether or not a state of affairs $A$ is taken as evidence for a hypothesis $H$ depends on one's background assumptions, and the in the second, she argues that there is often no way to justify our background assumptions on a purely empirical basis.

Longino takes seriously Quine's suggestion that theory is underdetermined by evidence. (Quine 1980b) She writes:

[A] given state of affairs can be taken as evidence for the same hypothesis in light of different background beliefs, and it can be taken as evidence for quite different and even conflicting hypotheses given appropriately conflicting background beliefs. Similarly, different aspects of one state of affairs can be taken as evidence for the same hypothesis in light of different background beliefs, and they can serve as evidence for different and even conflicting hypotheses given appropriately conflicting background beliefs. (43)

Longino illustrates her point using two different kinds of examples. In her first example, she considers how one reasons from the fact that a child has a stomach rash to a hypothesis about her condition. One might take the stomach rash to be evidence that the child has measles because one believes that such a rash is a symptom of measles. However, even if one lacked this last belief, one might still take the rash to be evidence for measles if one had quite a different background belief. For example, one
might believe that his neighbour has special clairvoyant powers and that any of this neighbour’s claims is true. Moreover, one might have been told by this neighbour that if the child had a stomach rash, she would be afflicted with a disease called ‘measles’. In these cases, one and the same state of affairs was taken as evidence for the same hypothesis for two different reasons. In the first case, the child’s stomach rash was taken as evidence for the hypothesis that she had measles because of a belief about the causal connection between the rash and the disease. In the second case, the rash was taken as evidence for the illness because of a belief in someone’s abilities to access occult information and reliably report it. One might also take one and the same state of affairs to support two conflicting hypotheses if one had sufficiently conflicting background beliefs. If, for instance, one believed that a stomach rash was a sign of good health, then one would take the child’s rash to be evidence for the hypothesis that she is perfectly healthy. Though it seems unlikely that someone should have such background beliefs, it is clear that such a case is possible.

It is also possible to attend to different features of a certain state of affairs and to reason either to the same or to competing hypotheses as a result of different background assumptions. Consider the state of affairs in which a gray hat sits on a bannister. Jones might have the belief that Nick’s hat is gray and reason that this state of affairs is evidence that Nick is in the house. In contrast, Smith might have the belief that Nick has a habit of leaving his hat on the bannister just so, and he might well reason from this state of affairs that Nick is in the house. In this case, Jones and Smith attend to different aspects of the same state of affairs, i.e. Jones to the hat’s colour and Smith to the hat’s location, and because they each have appropriate background beliefs, they coincidentally arrive at the same hypothesis about Nick’s whereabouts. Of course, it is also quite possible to arrive at different hypotheses by attending to different features of one and the same state of affairs. If, for example, Smith thought that only James tosses his hat onto the bannister just so, he might take the presence of the gray hat on the bannister to be evidence that James, not Nick, is in the house.

According to Longino, background beliefs function as they do in evidential reasoning because they are about actual or presumed connections between certain kinds of states of affairs and other kinds of states of affairs. Without such beliefs, no state of affairs could be taken as evidence for another. In one context of a given set of assumptions, state of affairs $A$ will be taken as evidence for hypothesis $H$, but in a different context with a different set of assumptions, $A$ will be taken as evidence for a different hypothesis $H'$, or for no hypothesis at all. Finally, Longino (1990) writes:
If a set of data $E$ is taken to be evidence for hypothesis $H$ in light of background assumption $B$, then we cannot with any finality determine whether it is correctly so taken by examining the evidential support for $B$, since whatever data $E'$ is taken for evidence $B$ is so in light of some further assumption $B'$. To maintain that there is a distinction between what is taken to be evidence and what is really evidence is to suppose that there is some nonempirical way to discover the truth or falsity of background assumptions. Even though observational and experimental data relevant to the direct confirmation of auxiliary hypotheses may not be available (in fact or in principle), we are not precluded from inquiring as to what kinds of reasons can be offered in their support. That is, we may have reasons to accept a set of background assumptions that are short of reasons demonstrating the truth of those assumptions. (52)

According to Longino, we are not always in a position to determine on the basis of empirical data alone the truth or falsity of our inquiry-guiding background assumptions. Since these assumptions are crucial for conducting scientific inquiry, we must have reasons besides truth for accepting these assumptions. The reasons that we cite for favouring some assumptions rather than others will reflect certain contextual values.

The underdetermination of scientific hypotheses by available empirical evidence is not peculiar to psychiatric taxonomy, but rather, it is a general problem. Moreover, the role of 'fundamentally nonempirical' factors is something with which all sciences must contend. Contrary to what Kendler would have us believe, the fact that these factors have such a crucial role in psychiatric taxonomy makes it more, not less, plausible that psychiatric taxonomy is a scientific enterprise after all.

### 2.3.2 Do contextual values always introduce distortions and biases into observation and theory construction?

Friends of the Traditional View assume that an undistorted theory is ensured by attending only to empirical data and strictly applying a scientific method that adequately serves only constitutive values. However, there are cases in which these elements alone do not yield a maximally informative theory, but rather, a theory of
only partial truths that can be misleading. I describe here two cases in which the methodological objectivity of inquiry is enhanced either by the scrupulous integration of contextual values in the inquiry-guiding background assumptions or by the social regulation of contextual values through critical public discourse.

Anderson (1995) discusses, for example, the portrayal of Jews in that Atlantic slave system in the Nation of Islam’s book *The Secret Relationship between Blacks and Jews*. In this book, many truths about Jews and their involvement in the slave trade are emphasized like the following:

1. that Jews invested considerably in the Dutch West India Company which had a significant role in the slave trade during the seventeenth century;
2. that there were a significant number of Jews who colonized Dutch Brazil and bought a considerable percentage of slaves traded by the Dutch between the 1630s and 1650s; and
3. that a larger percentage of Jews living in the U.S. South owned slaves than did white Southerners as a whole.

Anderson points out that the problem with how Jews are represented in *The Secret Relationship* is not that it contains many falsehoods, but rather, it does not include all the facts that would be relevant to making the proper comparisons and to assessing the significance of Jewish involvement in the slave trade. Part of the problem is that the question that *The Secret Relationship* purports to answer is not made explicit. One might think that the question is a seemingly value-neutral question such as:

**Question 1** What was the role of the Jews in the Atlantic slave trade?

However, Anderson contends that the question with which *The Secret Relationship* is concerned is better articulated as:

**Question 2** Do Jews deserve special moral blame for their roles in the Atlantic slave system?
Relative to Question 2, *The Secret Relationship* fails to give an unbiased account of Jewish involvement in the slave trade because it neglects to mention that the share of Jewish investment in the Dutch West India Company was rather small, that Jews only owned slaves in Dutch Brazil for a few decades before they were expelled by the Portuguese, that Jewish slaveowners in the U.S. owned fewer slaves per household than the average slaveowner, and that a vast majority of U.S. Jews lived in the non-slaveholding North. Thus, *The Secret Relationship* is biased because it does not include facts that are relevant to answering the Question 2, i.e. facts about whether Jews behaved any differently than anyone else who profited from the slave system.

If one is sympathetic to the Traditional View of science, one might be inclined to diagnose the problem here as the mere presence of contextual values and to solve the problem by requiring that contextual values not have any role at all in shaping the formation of scientifically interesting questions. Along this line of thought, Question 2 is illegitimate precisely because it rests on the presupposition that it is permissible to single out an ethnic group and to pass judgements of collective guilt on it. And, pushing this line of thought further, one might claim that the only scientifically legitimate questions are those which rest on no moral, social, or political presuppositions whatsoever.

Before assessing the adequacy of this suggestion, we need to address the question of whether or not Claim 2 is true or acceptable. I re-state this claim below:

**Claim 2** Values are nothing more than (i) things we judge to be valuable, (ii) things we desire; (iii) things in which we take an interest; or (iv) things which we prefer to other things.

Consider, for example, disjunct (ii), according to which values are nothing more than things that we desire. There exist things or states of affairs that we desire, but that aren't valuable or good or right. I may want to sit in front of the television gorging on cake and ice cream for months on end, but it isn't good for me to do it. I may want to break my promise to drive my friend to the airport, but it isn't right for me to do so. Furthermore, disjunct (iv), according to which values are nothing more than things we prefer to other things, does not seem correct. I may prefer posters of dogs playing poker to canvases by Manet, but that doesn't make the former fine art. Disjuncts (i) and (iii) can also be shown to be inadequate using similar examples, but
I leave that exercise to the reader. In the end, although it is important to recognize that people’s tastes and opinions will vary, it is equally important to acknowledge that in some cases, people’s judgments about what is valuable, good, or right can be mistaken so that we are in a position to criticize erroneous judgments and to correct them. Insofar as people can be mistaken about something’s aesthetic value or about an action’s moral value, there are some facts about values which aren’t ‘up to us’, so to speak. Some of those values will include beauty and justice.

Returning now to Anderson’s case, when the friend of the Traditional View fingers moral, social, or political presuppositions as the source of the bias and prejudice in *The Secret Relationship* and consequently forbids them from playing any role in the formulation of scientific questions, he misunderstands what is genuinely problematic about that manuscript. The authors of *The Secret Relationship* seem to take for granted the moral inferiority of Jewish persons. Although they appear to leave it an open scientific question whether, during the seventeenth century slave trade, Jewish persons behaved in a way that is consistent with this background assumption, they attended only to facts that were consistent with their background assumption so that they arrived at a biased and distorted answer. The problem isn’t that some moral proposition or other was assumed, but rather, that a false moral proposition was assumed. Furthermore, authors of *The Secret Relationship* conducted their inquiry in a way that was unfair or unjust. Thus, there are two problems in the work presented in *The Secret Relationship*, viz. one concerning the inclusion of a false moral proposition in the background assumptions and another concerning the way in which an answer was sought to the question posed at the start of the inquiry.

Correcting these problems is not a matter of excluding all moral propositions from the background assumptions or of somehow conducting inquiry in a morally neutral way. First, proper revision of the background assumptions in *The Secret Relationship* is not the mere deletion of the offending proposition, but rather, it requires the replacement of the erroneous moral proposition that Jews are morally inferior to people belonging to other ethnic groups with the true moral proposition that ethnicity alone is morally irrelevant. In the former approach, the background assumptions would leave open the relevance of ethnicity per se in moral judgments when this matter is not at all an open one, but in the latter, the background assumptions would constrain answers, theories, and explanations to those that don’t presuppose any moral falsehoods. Second, the authors of *The Secret Relationship* were rigging the inquiry so that the answer to Question 2 was not only misleading, but also unfair in its portrayal of Jewish persons. Whether or not an inquiry is conducting in a way that is just or fair is not determined
merely by whether or not it uncovers truths. Indeed, in this particular case, the claims made in the *The Secret Relationship* were mostly true. However, the fact that an inquiry is conducted unjustly is good reason to find it suspect and to doubt the legitimacy of its results even if those results contain only truths.

Anderson’s case of *The Secret Relationship* shows that contextual values do not always introduce biases into inquiry. First, when value-based propositions are included as background assumptions, we must be careful to exclude those that are false and to include relevant truths. When the latter are included, false or misleading theories are eliminated as legitimate candidates in response to questions of interest. Second, when we study other people and their actions, we ought to conduct our inquiry in a way that promotes not only truthfulness, but also fairness and justice, in our representations of them and their behaviour. In this way, we require the integration of certain contextual values, viz. fairness and justice, into scientific practice.

So far, I have only considered contextual values whose inclusion in background assumptions or in determining acceptable scientific conduct could operate to minimize distortions at the level of individual scientific practice. That is, if a particular scientist included certain contextual values as background assumptions and he also conducted inquiry in a way that is consistent with certain contextual values, he could minimize some biases and distortions in his own work. However, in many cases, it is not easy, and perhaps not even possible, to purge one’s self of personal biases. In such cases, our own personal and social interests can lead to distortions in results that we arrive at individually, but this does not render our projects hopelessly unscientific. It just means that we should welcome the opportunity to engage others whose perspectives differ from our own, and their findings which may include biases peculiar to them can be compared with ours so that we can expose each other’s prejudices and arrive together at claims, theories, explanations, or classification schemes in which each individual’s biases have been minimized. In this way, methodological objectivity of inquiry is enhanced not by constraining assumptions and inquiry at the level of individual scientists, but by regulating the influence of biases and distortions through critical discourse at the social level. An example in which the scientific community as a whole improved the methodological objectivity of their studies through communal regulation of personal and social biases is one described by Hrdy (1986).

Hrdy argues that theories of sexual selection have long portrayed animal mating practices in a distorted way because male-biased assumptions influenced the way in which data was collected. Hrdy claims that until relatively recently, there have been two
important stereotypes that provided the basis for how animals select their mates. First, there is the Victorian stereotypes about sexually undiscriminating males and highly discriminating ‘coy’ females, and second, there are stereotypes about nurturing females who are heavily invested in their offspring in contrast with those about relatively uninvolved males. Hrdy claims that these stereotypes led to assumptions about male and female sexual behaviour which were misleading. Early theories about male and female mating behaviours predicted that males would attempt to mate often with as many females as possible, and that females would select their mates carefully. While some findings were consistent with these predictions, starting in the 1970s there has been accumulating evidence that females engaged in polyandrous mating and that males often played complex and critical roles in the upbringing of offspring. Hrdy attributes the shift in the available evidence to the entrance of women in the fields of primatology and sociobiology. As women researchers engaged in sociobiological research, they took a greater interest in female mating behaviour than their male colleagues did. As a result, they made observations and collected data that their male colleagues neglected. A consequence of increased participation by women in sociobiology and primatology, theories of sexual selection are being revised in an attempt to represent more accurately sexual behaviour amongst animals.

In this case, complex social expectations of men and women living and interacting in human society were imported into research on the social behaviour of other animals. Men expected coy behaviour in women and had trouble envisioning women, or females of any other species, acting any differently. Consequently, they had little interest in studying female mating behaviour that was not consistent with their expectations. Male gender-bias led to distorted theories of sexual selection. As a matter of historical fact, these biases were removed, not by first disabusing male sociobiologists of their expectations of women and gathering new unbiased data, but rather, by the inclusion of data gathered by women whose biased interests led them to take a keen interest in female mating behaviours of non-human animals. Neither the men nor the women working in sociobiology conducted their studies in a way that was entirely free of their own gender specific interests. Indeed, it is difficult to imagine how one might go about completely purging someone of such interests. A theory of sexual selection need not, and arguably could not, result from requiring the elimination of all gender specific interest of each man or woman working in this field. Rather, an unbiased theory requires only that the men and women whose work contributes to it engage each other in critical discourse so that they may expose each other’s prejudices and correct them together so that the community of sociobiologists, as a whole, arrives at
a theory in which distortions originating from any one scientist have been minimized.

2.3.3 Summary

In this section, I have argued that the Traditional View is flawed. This view assumes that whenever contextual values have a role in guiding observation and theory construction in a particular field, that field is unscientific, and studies within it are rife with biases. I have challenged these assumptions by pointing out that contextual values often have an important role in scientific inquiry and by arguing that appealing to appropriate contextual values can actually reduce the degree to which a theory is distorted as long as we carefully regulate the role of these values. Managing the role of contextual values occurs at both the individual and social levels. In some domains of inquiry, the inclusion of some moral truths in one's background assumptions is appropriate and necessary to minimize distortions due to unfair representation of one's subject. In other fields of study, social interests are deeply ingrained, and inquiry would come to a standstill if we were each required to purge ourselves of them. Yet scientific inquiry is still possible because unbiased study depends not on the absence of bias in each individual researcher, but rather, on communal efforts to recognize and correct each others' prejudices so that we can arrive at a distortion-free theory together as a group of scientists. The importance of social regulation of contextual values in scientific inquiry cannot be understated. For, even in cases like Anderson's, the correction of problematic background assumptions and the rehabilitation of unjust practices are often only possible if one's colleagues scrutinize one's presuppositions and methods and one is responsive to their criticisms.

2.4 Anything Goes?

Given that contextual values have an important role in the practice of science, we must recognize that this needn't force us to conclude that anything goes in the epistemology of science. Upon recognizing the crucial role that contextual values play in science, it may be tempting to embrace either subjectivism or liberal pluralism about science and psychiatry, but such a move would be hasty and rash.
The objectivity of an inquiry is less a matter of eliminating value judgments than it is a matter of appropriately regulating which contextual values have guiding influence. It is important to recognize that such regulation is a social endeavour because the best way to limit the influence of personal biases and prejudices is to subject evidence, hypotheses, and background assumptions to critical scrutiny from many alternative perspectives. Moreover, the mere presence of intersubjective criticism is not enough to secure methodological objectivity in science. Scientific practices and judgements must be responsive to criticisms. To ensure an appropriate level of responsiveness, Longino (1990) suggests four criteria for critical discourse within the scientific community: (1) there must be recognized forums for criticism of evidence, methods, assumptions, and reasoning; (2) there must be shared standards that critics may invoke; (3) the community must be responsive to criticism; (4) intellectual authority must be shared equally amongst qualified practitioners. Recognized forums for criticism include peer-reviewed journals and professional conferences in which experts scrutinize each other’s work. There should also be standards which one may criticize another scientist’s work by appeal to standards to which all members of the community feel bound. And, once it is shown that certain scientific practices or claims have not withstood sustained criticism, there should be appropriate reform. Finally, all qualified scientists should be treated as equals within the community, and no one’s claims or criticisms should be discounted just because they are not affiliated with a dominant social group.

Subjectivism, according to which scientific claims are whatever we wish or believe them to be, seems to presuppose that Claim 2 is true. For, unless we take value judgments to be mere opinions, it is implausible that the contextual values which help to justify scientific claims could be anything we want them to be. However, as I have argued in Section 2.3.2, we should reject Claim 2. The truth of a value judgment is not secured merely by our believing it to be so. Value judgments are subject to scrutiny and are revisable in light of evidence and argument.

Some writers who do not want to endorse subjectivism have instead championed pluralism, according to which many different psychiatric nosologies are equally legitimate. Ghaemi (2003) has argued that the greatest flaw of the DSM as a nosologic project is that it serves too many masters. It is a taxonomy of psychiatric illnesses that is supposed to serve the interests of clinicians, educators, students, insurance companies, government officials, and researchers, and these interests may not always coincide. For example, a researcher may well be satisfied with a diagnostic category whose features are subject to revision in light of future evidence, but a clinician whose fee depends on an insurance company’s willingness to pay for treatment of a disorder...
will be dissatisfied with fluid diagnostic categories which may or may not apply to a patient from year to year. As such, decisions about which disorders to include in the DSM become unnecessarily difficult, and in the end, by trying to satisfy everyone, the DSM satisfies no one. Ghaemi's solution to the problem is the endorsement of pluralism with respect to psychiatric taxonomy. Each group with an interest in psychiatry will have its own taxonomy to serve its goals and purposes, and there is no basis for favouring any one classification system over another.

Appealing to contextual values can result in different competing theories or taxonomies that are equally acceptable when assessed relative to the relevant guiding values, but it doesn't follow that we should accept unconstrained pluralism about psychiatric disorders, or any other domain of scientific inquiry. Not all political, economic, social, and moral interests are equally important, or even legitimate. It is absurd to assert the legitimacy of a taxonomic system in which few people can meet criteria for a disorder because insurance companies are loath to pay for treatment of that condition or to claim that such a classification is no better or worse than a taxonomy based on pathophysiological and epidemiological data. However, it is reasonable to suppose that a taxonomy founded on etiological data could be different than one based on clinical treatment outcomes, and that neither is any more or less legitimate than the other. In the former, one might be more interested in preventing the onset of disordered behaviour and experience, while in the latter, one might be more interested in treating disorders that have already been realized and choosing effective treatment regimens. Some disorders that have similar etiological patterns may be more or less responsive to different treatments, and other disorders that have different etiological features may respond to one and the same kind of treatment. Thus, both etiologically-based and treatment-based taxonomies may differ, but be equally legitimate. And, neither an etiologically-based nor a treatment-based classifications need be placed on equal footing with insurance-reimbursability-based taxonomies.

2.5 Conclusion

In this chapter, I have argued that psychiatric nosology can be a legitimate scientific enterprise as long as we recognize the deficiencies of the Traditional View and acknowledge the role of contextual values and their social regulation in science. While there is room for moderate pluralism within psychiatric nosology, the fact that some
goals, values, and interests are more important and defensible blocks a slide to extreme pluralism in which any possible taxonomy of disorders could claim legitimacy. Contextual values that earn the right to guide psychiatric taxonomy are justified by their ability to withstand critical scrutiny by members of the mental health community.
Chapter 3

What kind of thing is a disorder?

3.1 Introduction

In Chapter 1, I argued that anti-realists about psychiatric illness have not succeeded in giving us compelling reasons to embrace their view. Let us briefly re-visit the debate between the realists and the anti-realists about mental illness. Recall that, according to Sayre-McCord (1988a), realism about any domain is best understood as a commitment to two theses:

Thesis 1 The claims in question, when literally construed, are literally true or false.

Thesis 2 Some of these claims are literally true.

In the case of psychiatric realism, judgments about the presence or absence of a mental disorder are the claims of interest. Although I formulated the relevant judgments in terms of the predicate ‘is mentally ill’ in Section 1.2, the psychiatric claims that interest us here have the form given in Schema 4.1

1I take a claim of the form Person $P$ is mentally ill and one of the form Person $P$ has a mental disorder to be approximately equivalent. While there might be an interesting distinction between being ill and having a disorder, I shall not pursue that issue here.
Schema 4 Person \( P \) has a mental disorder

When Theses 1 and 2 are applied to psychiatric claims, instances of Schema 4 include important presuppositions that I state below:

5 According to Thesis 1, instances of Schema 4 presuppose:

a that there exists a property \( H \) that is picked out by the predicate ‘has a mental disorder’; and

b that when one utters an instance of Schema 4, the property \( H \) is being predicated of something.

6 According to Thesis 2, instances of Schema 4 presuppose:

a that the property \( H \) just is the property of having a mental disorder; and

b that there is at least one entity that has the property of having a mental disorder.

In Chapter 1, I discussed arguments for rejection of the analogs of presuppositions 5(a), 5(b), and 6(a), and I argued that they were not persuasive enough to compel us to abandon realism in favour of anti-realism. (See Sections 1.3 - 1.5.)

Presently, I would like to explore some common realist strategies for providing an informative account the property \( H \), reference to which is made in presuppositions 5 and 6. That is, I seek an account for the property of having a mental disorder. One might wonder what more a realist needs to say besides that which is stated in presuppositions 5 and 6. However, even after one accepts the presuppositions, there remain interesting metaphysical questions. Consider, for example, realism about tigers. Even after we all agree that the predicate ‘is a tiger’ picks out a property and that the relevant property just is the property of being a tiger, we might still wonder what having that property really amounts to. That is, we might still, in all seriousness, ask What does it take to be a tiger? or What kind of thing is a tiger?. In this chapter, I am particularly interested in answers to a question analogous to the latter, i.e. the question What kind of thing is a mental disorder?. Although I am ultimately concerned with an account of mental disorders, the more general problem
of providing an account of disorders has proved challenging for many writers with realist sympathies, so I too will discuss general accounts of disorders here.

In this chapter, I survey various accounts of disorders as Lockean kinds, and I argue that these strategies fail. In Section 3.2, I describe Lockean kinds generally. In Sections 3.3 and 3.4, I discuss various views according to which disorders are Lockean kinds, and I argue that they are all unsuccessful. In Section 3.5, I sketch an account of disorders as homeostatic property clusters that crucially involve an impairment of functioning in an Aristotelian sense, and I claim that this view is more promising than Lockean accounts because it better captures the heterogeneity of disorders.

### 3.2 Lockean Kinds

One common strategy for distinguishing disorders from non-disorders is to treat disorders as Lockean kinds, which are special sorts of objective types. (Armstrong 1989) Consider all of the things that populate the universe, including the smallest subatomic particles, ordinary physical objects like chairs and tables, and even more abstract things like events. We can group these things into sets that are more or less unified. An objective type is a set $S$ in which each member of $S$ is similar to all other members of $S$ in some respect $R$ and dissimilar to all non-members of $S$ in respect $R$. An objective type is objective in the sense that whether or not the members of $S$ instantiate the properties that unify them is independent of our opinions, judgments, and theories about them. The unity of objective types admits of degrees. Some types are highly unified, e.g. the set of all hydrogen atoms, while others are only weakly unified, e.g. the set of all chairs. In some cases, we may have a set of things which is not unified in any respect, and this set of things would not constitute an objective type.

Lockean kinds come in two varieties: real kinds and nominal kinds. Membership to a real kind is determined by a real essence, and membership to a nominal kind by a nominal essence. The real essence of a kind $K$ includes the properties that have a fundamental role in the world’s causal structure, and explanations of the behaviour or superficial properties of $K$’s crucially include reference to their real essence. The nominal essence of a kind $K$ usually includes the most salient superficial properties of $K$’s, and we take these properties to determine the kind because they are interesting or useful to us for the purposes of classification.
Consider how gold might be a Lockean kind. Let's begin by treating it as a real kind. Arguably, the real essence is its atomic structure which is encoded in its atomic number, 79, and all and only things with atomic number 79 are samples of gold. Moreover, the atomic structure of gold explains its other chemical and physical properties, e.g. its malleability. If we didn't have access to the real essence of gold, say, because modern chemistry hadn't yet been developed, then we could only group together gold things using its nominal essence. Presumably, some of the properties of gold that most interest us include its colour, lustrousness, and pliability, so we might take gold to be all and only those things that are yellow, shiny, and malleable. In this example, the sets $S$ and $S'$ which are determined by the real and nominal essence of gold, respectively, may not be co-extensive because there might be yellow, shiny, malleable things that aren't gold, or vice versa. It is not clear whether it is possible to define the nominal essence of gold, or any other kind, so that the nominal kind and the real kind are co-extensive.

Although the terms 'Lockean kind' and 'natural kind' are often used interchangeably, I wish to distinguish them. I reserve the term 'natural kind' to refer to types that are unified by natural properties, i.e. physical, chemical, or biological properties. Although some Lockean kinds, such as gold, may also be natural kinds, it is compatible with the Lockean view that either the real or nominal essence of a kind include non-natural properties, e.g. psychological or social properties. For example, suppose one thought that there are necessary and sufficient conditions for one's having the intention to do X. Suppose further that these conditions included reference to one's beliefs and desires about X. If one didn't think that beliefs and desires could be naturalized, one would hold the view that intentions are a Lockean psychological kind. Similarly, suppose all and only tribes have in common certain social properties. Then, tribes would constitute a Lockean social kind.

Before turning to specific accounts of disorders as Lockean kinds, I would like to make a few clarificatory remarks about how these views will be presented. Since the proponents of the views that I will shortly be presenting take themselves to be revealing the nature of disorders, I take them to be treating the set of all disorders as a real kind, and henceforth, I shall use 'essence' and 'Lockean kind' to refer more specifically to a real essence and a real kind, respectively. Since an essence determines membership to a Lockean kind, giving an account of such a kind amounts to giving a description of the essence. Moreover, identifying an essence is just a matter of providing necessary and sufficient conditions for membership to the Lockean kind. These conditions for kind membership are often articulated as a definition or analysis.
of the kind concept of interest. In what follows, I present accounts of disorders by stating proposed definitions or analyses of the concept of disorder. Although this way of presenting the views has the disadvantage of giving the impression that the disagreements between competing accounts are merely semantic, I wish to emphasize that the proposed analyses are a way of articulating or describing the essence of disorders, and hence, to be providing answers to the metaphysical questions What is it to be a disorder? and What kind of thing is a disorder?.

3.3 Disorders as Purely Natural Lockean Kinds

Many scientifically minded writers favour treating the concept of disorder as a Lockean natural kind concept. (Cohen 1981, Kendell 1975, Scadding 1990) In what follows, I shall discuss two different purely naturalistic accounts of the concept of disorder: one based on statistical deviation and another based on biological disadvantage.

3.3.1 The Statistical Deviation Account

According to the statistical deviation account favoured by Cohen (1981), the concept of disorder can be analyzed thus:

Definition 4 D is a disorder iff D is a quantitative deviation from a relevant statistical norm.

The virtue of this account is that it can capture our intuitions about some disorders. For example, hypertension would rightly count as a disorder because it is characterized by blood-pressure that deviates from the statistically normal range of blood-pressures for an adult human. However, this account of disorder cannot be right because deviation from a statistical norm is neither necessary nor sufficient for a disorder.

Since there exist deviations which aren't disorders, statistical deviation from normal is not a sufficient condition for a disorder. Olympic weightlifters who are much stronger than normal adult humans aren't disordered, nor are geniuses whose intelligence
deviates greatly from normal. Even if we tried to save the statistical deviation account by requiring that the deviations be undesirable, there still exist counterexamples to the sufficiency claim. For example, adult men who are five feet tall are statistically deviant from the normal height for adult male humans, and it is arguably undesirable for them to be so. Yet, these men are not obviously disordered. Besides being an inadequate modification to the statistical deviation account, the inclusion of an undesirability condition would be unsatisfactory to proponents of this view because their endorsement of it is motivated principally by their conviction that disorders are Lockean natural kinds, and it is unclear whether undesirability could be naturalized to their satisfaction.

Just as there are counterexamples to the sufficiency condition, there are problem cases for the necessity claim. For example, atherosclerosis and dental caries are statistically normal in American society, but this doesn’t mean that neither counts as a disorder. Even if we imagined a world in which all humans lived with AIDS, this condition would not cease to be a disorder. Thus, there are disorders whose status as such is not compromised by its actually or possibly being the statistical norm.

3.3.2 The Biological Disadvantage Account

In an attempt to overcome the difficulties associated with the statistical deviation account, some naturalistically minded writers have proposed an account, according to which disorders are conditions that place organisms at a biological disadvantage. (Kendell 1975, Scadding 1990) In order for this account to remain a naturalistic one, its proponents must interpret ‘disadvantage’ in natural terms. Kendell favours the following analysis

Definition 5 Organism $O$ has a disorder iff:

1. there is a condition $C$ such that $O$ has $C$; and
2. $C$ reduces the fertility or longevity of $O$. 
While a reduction in longevity can be straightforwardly understood as anything that shortens an organism’s life, a reduction in fertility might be understood either narrowly or broadly. On the narrow interpretation of a reduction in fertility, it is the inability of the organism’s reproductive system to perform its function properly. On the broader understanding, it is best taken to be a decrease in reproductive fitness due to any number of factors including, amongst others, the inability to attract mates and the inability to mate. There is some textual evidence that Kendell favours the broader interpretation of the reduction in fertility because he wants to accommodate psychiatric disorders, like schizophrenia, in his account. Since schizophrenia is not a disorder of the reproductive system, it wouldn’t count as a disorder unless it was associated with a reduction in fertility through mechanisms other than the dysfunction of reproductive organs. Indeed, the kinds of evidence the Kendell cites for the decrease in fertility of schizophrenics includes lower rates of marriage and higher rates of childlessness.

Like the statistical deviation account, the biological disadvantage view seems to get some paradigm cases of disorders right. For example, ataxia telangiectasia (AT) is an immunodeficiency disorder with a very poor prognosis. Sufferers of AT typically die in their late teens or early twenties although some have survived until their forties. As such, AT reduces fertility and longevity of its sufferers because many die just as they reach their prime child-bearing years.

However, the biological disadvantage account has its fair share of problems, too. Some conditions, like sciatica, are such that afflicted persons can experience a great deal of pain and a considerable reduction in their quality of life without having their fertility or longevity seriously compromised. Yet, it is plausible and reasonable to include such conditions as disorders. Since there exist disorders that don’t reduce fertility or longevity, biological disadvantage, in the favoured sense, is not necessary for a disorder. Moreover, biological disadvantage isn’t sufficient for a disorder. For example, homelessness or unattractiveness presumably reduces one’s chances of successfully mating, but it is not reasonable to suppose that homelessness or unattractiveness is a disorder. The obvious way to block homelessness from counting as a disorder is to stipulate that we should understand the reduction in fertility criterion narrowly. However, this move would exclude conditions like schizophrenia from disorder status.
3.4 Disorders as Partially Natural Lockean Kinds

Although Wakefield is critical of purely naturalistic Lockean accounts of disorder, he does endorse a view according to which disorders are Lockean kinds with both natural and non-natural properties. (Wakefield 1992) According to Wakefield, disorders are harmful dysfunctions, and the concept of disorder can be analyzed thus:

**Definition 6** Organism $O$ has disorder of structure or mechanism $X$ iff:

i. there is a function $F$ such that $F$ is a natural function of $X$;

ii. $X$ fails to perform $F$; and

iii. (ii) causes some harm to $O$ or deprives $O$ of some benefit as judged by the standards of $O$'s culture.

Conditions (i) and (ii) together constitute the dysfunction criterion, and condition (iii) the harmfulness criterion.

Wakefield’s dysfunction criterion presupposes an evolutionary theory according to which an anatomical structure or physiological mechanism can acquire a natural function. Roughly speaking, a natural function of a structure or mechanism is an adaptation. An adaptation is an ‘inherited and reliably developing characteristic that came into existence as a feature of a species through natural selection because it helped, directly or indirectly, facilitate reproduction during the period of its evolution.’ (Buss, Haselton, Shackelford, Bleske & Wakefield 1998, 535) A piece of anatomy might have many functions, but not all of those functions are its natural function. A nose can, and often does, function to support spectacles, but this is not its natural function. It is not the case that the human nose was naturally selected as a support for spectacles because human reproductive fitness was enhanced by humans’ ability to rest spectacles on the bridge of their noses. However, it is plausible to suppose that the nose was naturally selected for its ability to detect odours of potential foodstuffs and potential mates, and hence, that the natural function of the nose is to detect odours. In this case then, failure to support spectacles is not a nasal disorder, but failure to smell could be such.
Since not all dysfunctions are disorders, satisfaction of the dysfunction criterion alone is not sufficient for a condition to count as a disorder, but Wakefield suggests that satisfaction of this criterion together with a harmfulness criterion would be sufficient. He invites us to consider a speculative case in which we are to suppose that the mechanism by which humans age is naturally selected. He argues that if this mechanism malfunctioned so that the aging process occurred at a slower rate in a given person P, we would not consider P disordered, but rather, very fortunate. So, he proposes that dysfunctions are only disorders if they also meet a harmfulness criterion which he articulates in non-natural terms. According to the harmfulness criterion, whether or not a condition is a disorder is determined in part by a social belief, the content of which is evaluative. That is, in order for a condition to count as a disorder, a relevant cultural group must believe that it is harmful to its bearer.

There are several problems with both the dysfunction and the harmfulness criteria, but let us begin by considering only the latter. The obvious problem with the harmfulness criterion, as it is stated in Definition 6, is that it has the absurd consequence that one could effectively cure one’s self of a disorder by leaving one cultural group and immersing one’s self in another. Consider, for example, dyschromia spirochetosis, which is also known as pinta. Pinta is a dermatological disease caused by bacterial infection and spread by non-sexual skin-to-skin contact. It is particularly prevalent in Central and South America. One of its characteristic symptoms is rose-coloured lesions on the skin. According to medical sociological reports, there is a South American tribe whose members find the pink-spotted skin of pinta sufferers attractive and believe it to be a sign of good health. Indeed, persons who don’t exhibit this symptom are considered sickly, and they are forbidden to marry. (Reznek 1987, 18) If one were to take Wakefield’s analysis, as stated in Definition 6, members of this tribe with pinta would not have a disorder, but all other people afflicted with pinta would have a skin disorder. However, this is the wrong conclusion to draw from this case. It is much more reasonable to claim that pinta is a skin disorder, but there are some cultures which mistakenly believe that it isn’t. If this is, indeed, the right way to respond to the pinta case, we should modify the harmfulness condition so that the analysis is stated thus:

**Definition 7** Organism O has disorder of structure or mechanism X iff:

1. there is a function F such that F is a natural function of X;
ii. $X$ fails to perform $F$; and

iii. (ii) causes some harm to $O$ or deprives $O$ of some benefit.

In Definition 7, the harmfulness criterion (iii) is such that a condition could be harmful to $O$ or deprive $O$ of some benefit regardless of anyone's true or false beliefs about the matter.

Even after modifying Wakefield's analysis, it is still inadequate because the dysfunction criterion is problematic. Since dysfunction is supposed to be a failure of an adaptation, some critics have pointed out that the failure of an exaptation to perform an important function would not count as a disorder because an exaptation does not have a natural function in Wakefield's preferred sense. Gould (1991) introduced the concepts of exaptation, co-opted adaptation, and spandrel. According to Gould, an exaptation is a feature that currently enhances fitness, but that did not emerge by natural selection for its current role. There are two kinds of exaptations: co-opted adaptations and spandrels. A co-opted adaptation is a feature that is 'now useful to an organism that did not arise as an adaptation for its present role, but was subsequently co-opted for its current function.' (Gould 1991, 43) For example, a bird's feathers are thought to have evolved for the purpose of thermo-regulation, but they were subsequently 'co-opted' to perform the function of flight. Gould describes spandrels as 'presently useful characteristics [that] did not arise as adaptations at all, but owe their origin to side consequences of other features.' (Gould 1991, 53) Gould borrowed the term 'spandrel' from the field of architecture in which this term refers to the space left over between the outer curve of an arch and the straight-sided wall and ceiling that bind the arch. Spandrels are often decorated, and the designs often fit the space cleverly. For example, the four spandrels in the dome of San Marco in Venice are decorated with depictions of the four evangelists. Although the spandrels are used to frame artistic representations of the evangelists, they were not constructed for this purpose. Similarly, Gould argues that the human brain may have evolved to its large size as a solution to some unspecified adaptive problem, but its current uses for reading, writing, religion, and other cultural tasks were not the roles for which our large brains were naturally selected. Since both co-opted adaptations and spandrels can perform important functions that aren't those for which they were selected, then any malfunctioning exaptation that is an obvious disorder would be a counterexample to Wakefield's account.
Lilienfeld & Marino (1995) point out that certain cognitive abilities such as our ability to do arithmetic are not plausibly products of natural selection, but rather, fortuitous by-products of the development of our brain to do other unspecified tasks to enhance our reproductive fitness. Yet, acalculia, which is an impairment of the ability to perform simple mathematical calculations like addition and multiplication, seems a perfectly legitimate neurological disorder that can result from lesions in the parietal and frontal lobes. Still another counterexample is amusia which is, roughly, the inability to recognize musical tones or reproduce them.

Wakefield’s reply to the exaptation objection is the following:

There is no requirement that the harm must be identical to the failed natural function. For example, in American society, reading is a socially valued and beneficial skill, so the inability to read is a harm. If that harm results from a dysfunction of one of the submechanisms of the many involved in reading, then there is a disorder. There is no requirement that the harm of not being able to read must be due to the failure of the mechanism that has the function ‘enabling one to learn to read’. It is only required that some mechanism must have some natural function that enables people to learn to read, and that the failure to learn to read is caused by the failure of such a function. (Wakefield 1999, 382)

Wakefield’s response is to modify his analysis thus:

**Definition 8** Organism $O$ has disorder of structure or mechanism $X$ iff:

i. there exists a structure or mechanism $Y$ such that $Y$ has a natural function $F$;

ii. there is a function $F'$ such that $X$ fails to perform $F'$ because $Y$ fails to perform $F$; and

iii. (ii) causes some harm to $O$ or deprives $O$ of some benefit.
Although such a modification gives Wakefield’s account the flexibility to handle amusia, acaclulcia, and reading disorders, the modified analysis in Definition 8 is too inclusive, as his critics have pointed out. (Lilienfeld & Marino 1999, Sadler 1999) For example, the inability to cook is arguably a condition that deprives one of some benefit, and that, in some environments where one cannot obtained food cooked by someone else, could be harmful. The ability to cook is a complex function that requires the use of one’s arms. Suppose that person P’s arms don’t function properly due to a neurodegenerative disorder. The analysis in Definition 8 leads to the absurd result that P would suffer from a cooking disorder over and above the neurodegenerative one. In an attempt to accommodate disorders that would otherwise be excluded by the harmful dysfunction analysis, Wakefield has settled on a modified account that brings with it just as many, and possibly more, problems that it can solve.

In addition to the difficulties with both the dysfunction and harmfulness criteria, Wakefield’s analysis of disorder does not appear to be the best account to fit available empirical evidence. It is commonly observed that mental disorders are characterized by clinical phenotypic heterogeneity. For example, some, but not all schizophrenics, have paranoid delusions. Some, but not all, schizophrenics have visual hallucinations. Some, but not all, schizophrenics are catatonic, and so on for any possible symptom of schizophrenia. In short, there is no one clinically observed property, or cluster of such properties that all and only schizophrenics share. There is also mounting evidence suggesting that there are no molecular, genetic, or other etiological properties that, alone or together, are shared by all and only schizophrenics. (Lewis & Lieberman 2000, Tsuang 2000, Thaker & Carpenter. 2001) Moreover, etiological, genotypic, and both gross and molecular phenotypic heterogeneity seem to be characteristic of many non-psychiatric medical disorders, ranging from Alzheimer’s disease, rheumatoid arthritis, and various cancers, to name but a few. (Tanzi & Bertram 2001, Schellenberg 1995, Weyland, McCarthy & Goronzy 1995, Cuny, Kramar, Courjal, Johannsdottir, Iacopetta, Fontaine, Grenier, Culin & Theillet 2000, Shah, Mehra, Chinnaiyan, Shen, Ghosh, Zhou, MacVicar, Varambally, Harwood, Bismar, Kim, Rubin & Pienta 2004)

It may be tempting for someone with Lockean sympathies to hypothesize that, given the data, schizophrenia, Alzheimer’s disease, rheumatoid arthritis, and breast and prostate cancer are not Lockean kinds, but that subtypes of these disorders are. While it is certainly possible that such a claim could be vindicated, committing ourselves to it would be an act of unwarranted optimism at this time. It is far more reasonable to take these data as motivation to explore some non-Lockean accounts of disorder.
that can accommodate and even explain some of the observed heterogeneity.

3.5 Disorders as Property Cluster Kinds

Since psychiatric and non-psychiatric disorders are remarkably heterogeneous, it seems strange to take, as a starting point, the postulation of a real essence that is supposed to explain the unity of disorders. To be sure, disorders seem to have enough in common to be distinguished from non-disorders, at least in paradigm cases, but they are also diverse enough that it is unlikely that they all share an essence. I do not take the project of providing an account of disorders to be a search for the essence of disorders. Rather, I suggest that we take, as our point of departure, two facts about disorders: (i) the fact that they are phenotypically, genetically, and etiologically heterogeneous; and (ii) the fact that many disorders have features that cluster together in enough cases that there is a genuine distinction between disorders and non-disorders, albeit one that isn’t always sharp. I seek, then, a model that can explain these facts about disorders.

A promising alternative to Lockean views of disorders is Boyd’s homeostatic property cluster (HPC) model which he first proposed as an account of biological species. (Boyd 1991, Boyd 1999) According to Boyd, an HPC kind is not determined by one or more properties that constitute the kind’s essence. Rather, an HPC kind is determined by the co-occurrence of a family $F$ of properties that cluster together in an important number of cases. Moreover, the co-occurrence of these properties is the result of underlying causal mechanisms such that either (i) the presence of some of the properties causally favours the presence of others, or (ii) the causal mechanisms maintain the presence of the properties of $F$. Moreover, imperfect homeostasis is nomologically possible or actual. That is, some members of an HPC kind display some but not all properties in $F$, or they exhibit some, but not all, underlying causal mechanisms responsible for the clustering of $F$.

An HPC kind concept applies to things in which homeostatic clustering of most properties in $F$ occurs. As such, the concept cannot be defined or analyzed into necessary and sufficient conditions for membership to the kind. An important consequence of the HPC model is that there will be many cases of extensional indeterminacy of an HPC concept. There will exist cases in which, given all relevant facts and true theories, no rational considerations dictate whether an HPC concept applies to the case or
not. Moreover, attempts to substitute a concept that is less extensionally indeterminate will be a distortion because it would require us to treat important distinctions as irrelevant to causal explanation or to ignore significant similarities that are important to causal explanation.

As mentioned earlier, biological species are supposed to be paradigm cases of HPC kinds so let us consider how *canis familiaris*, i.e. dogs, constitute an HPC kind. There are many properties that most dogs share such as furriness, four-legged-ness, two-eyed-ness, and so forth. The underlying causal mechanisms that are responsible for this clustering of properties includes, but are not limited to, interbreeding, molecular mechanisms for expression of genes, and mechanisms governing the ways in which epigenetic factors can influence genetic expression. Not all dogs have four legs, and not all dogs' physical dimensions need be determined by the exact same combination of diet, exercise, and genetic predisposition. Since evolution is a dynamic process, it is quite plausible that there once existed dog-like creatures that bore some resemblance to both dogs and their evolutionary predecessors, and that, given all the biological facts and true theories, it would still be indeterminate whether the concept of *dog* properly applies to these creatures.

Non-psychiatric disorders are also good candidates for HPC kinds. Consider, for example, basal cell carcinoma, a non-fatal skin cancer. The gross clinical phenotype is a skin lesion, but there is considerable variability in its properties. In some cases, it presents as an itchy, reddish patch. In other cases, it is a white, waxy scar-like area, and in still other cases, it is a mole-like bump. Like all other cancers, there are complex genetic and epigenetic factors that causally contribute to the gross phenotypic expression, including the expression of various oncogenes and suppression of certain tumor suppressor genes. Not all cases of basal cell carcinoma require activation of precisely the same oncogenes or deactivation of the exact same tumor suppressor genes. Moreover, activation of oncogenes or deactivation of tumor suppressor genes needn’t arise from precisely the same causal events. In some cases, exposure to radiation contribute to, say, the activation of a crucial oncogene, but in others, it could be exposure to a carcinogenic chemical substance.

The HPC model can also accommodate psychiatric disorders, like schizophrenia. Schizophrenia is a chronic disorder that is often, but not always, associated with psychotic symptoms such as delusions and hallucinations. Schizophrenic persons may also exhibit catatonia, alogia, disorganized speech, and affective flattening, amongst other possible symptoms. Although schizophrenic symptoms do tend to cluster to-
gether, there is no single property or cluster of properties shared by all and only schizophrenics. While many etiological, genetic, and pathophysiological factors have been identified as contributors to the onset of the condition, no one factor or group of factors is present in all cases of schizophrenia.

It is important to notice that, unlike biological species, non-psychiatric and psychiatric disorders include psychological properties, which may or may not be naturalized depending on one’s view of mental properties. In the case of non-psychiatric disorders, painfulness, soreness, and itchiness are often part of the family of properties that cluster together in instances of the disorder. In the case of psychiatric disorders, the psychological properties that may be present in relevant clusters are even more wide-ranging, and they may include emotions, moods, beliefs, desires, thoughts, and sensations. It is also noteworthy that neurological and psychiatric disorders may include behavioural and social properties in their homeostatic clusters. For example, pervasive developmental disorders such as autism and Asperger’s Syndrome include impairment in nonverbal behaviours such as eye contact, facial expression, and bodily postures that are important in ordinary social interactions. Behavioural symptoms such as violent outbursts or extreme social passivity can be present in schizophrenia or in Alzheimer’s disease.

The anatomical, physiological, genetic, psychological, behavioural, and social properties that are present in disorders don’t have one common unifying feature. Some, but not all, disorders may involve bacterial or viral infection. Some, but not all, disorders may involve genetic mutation. Some, but not all, disorders reduce an organism’s reproductive fitness. Some, but not all, threaten an organism’s life. Some, but not all, disorders are contagious. Some, but not all, disorders involve a statistical deviation from an anatomical, physiological, behavioural, or social norm. Some, but not all, disorders involve the failure of a naturally selected function. Statistical deviations, biological disadvantages, and dysfunctions are involved in many disorders, but notably absent in many others. None of these features unify disorders in the way required for latter to count as Lockean kinds. We can see, then, why the Lockean accounts presented earlier had some limited success in capturing our intuitions about disorders, and also why these accounts ultimately fail.

Thus far, I have only mentioned natural, psychological, behavioural, and social properties that may cluster together in a variety of disorders, but clustering of these properties alone is insufficient for a condition to count as a disorder. Consider the following two cases:
**Case 1** Jones loses his only child in a car accident. For months after the child’s death, Jones is in a depressed mood all day nearly every day. He loses interest in his work and hobbies. He doesn’t spend as much time as he usually would socializing with his friends and other family members. He is unable to sleep at night. He hasn’t much of an appetite, and he loses a significant amount of weight during these months. He is tired all of the time so he cannot find the energy to go for his daily jogs. Although he hasn’t the energy or motivation to engage in cognitive activities such as reading, on the few occasions when he does pick up a book, he is unable to concentrate. Jones is unable to overcome his grief and sadness. Jones’s symptoms worsen to the point where Jones lack of interest and his inability to concentrate cost him his job. Although friends and family offer their support and help, Jones is unreceptive, and he withdraws further from his social network. Eventually Jones becomes socially isolated. Jones can barely take care of himself, and he is no longer able to ask friends and family for help.

**Case 2** Smith also loses his only child in a car accident. For months after his child’s death, Smith experiences exactly the same symptoms as Jones. Smith is also in a depressed mood all day nearly every day, and he loses interest in his work and hobbies. He spends less time than usual socializing with friends and family. He is unable to sleep. He loses his appetite, and he also loses weight. He suffers from fatigue, and he is much less active than he used to be. He also has trouble concentrating. However, unlike Jones, Smith’s grief and sadness does not cause him to withdraw as severely from his friends, family, and colleagues. Although Smith does spend a few months spending markedly less time socializing, he is able to maintain some of his friendships, and at some crucial times, he is even able to seek out support from close friends and family members. After grieving for a few months, Smith is able to recover gradually. Slowly, he is able to find the energy and motivation to socialize, to engage in physical activities, and to work. He regains his appetite, and his sleep patterns eventually return to normal. Eventually, Smith’s mood lifts, and although he is sometimes saddened by the thought of his child’s death, he spends most of his days in a euthymic state.

In both Cases 1 and 2, Jones and Smith find themselves in a persistent depressed mood for a few months following the same triggering event-type, i.e. the death of a child. I have stipulated that they experience the same kinds of emotional, cognitive, physiological, and social symptoms. Although we don’t fully understand the causal
relationships between depressive symptoms, it is quite plausible that the clustering of depressive symptoms are supported by underlying homeostatic causal mechanisms. For example, loss of appetite contributes to metabolic disruptions that can result in weight loss, fatigue, lack of motivation, and even cognitive impairment. Thus, in this way, one of the symptoms helps to sustain the presence of some of the other symptoms. Although not all cases of the clustering of depressive symptoms need involve precisely the same symptoms or even the same underlying causal mechanisms, let us stipulate in Cases 1 and 2 that the symptoms are indeed supported by the same causal mechanisms.

Despite the obvious similarities between Case 1 and 2, there is an important distinction between them, viz. Jones suffers from an affective disorder, but Smith doesn’t. Jones and Smith experience the same kinds of symptoms so the crucial difference is not traceable to the absence or presence of one or more symptom. Furthermore, we have stipulated that the causal mechanisms responsible for the clustering of the symptoms are the same, and even the triggering event is the same event-type. What does seem to be different is that one of symptoms was less severe in Smith’s case than it was in Jones’s, viz. the degree to which Jones and Smith withdrew from their social network. Jones experienced that symptom to such a degree that it became debilitating. Despite the presence of caring people, Jones was unable to receive the emotional support that was available to him and to use this support to alleviate his grief and to begin a reversal of some of his other symptoms. Instead, he withdrew further making it increasingly difficult for him to get the emotional and social support that he needed and increasingly unlikely that he would be able to recover from what began as a normal emotional reaction to the loss of a child.

All of the symptoms experienced by Jones and Smith admit of degrees, but in Case 1, Jones exhibited at least one of his symptoms to a high enough degree that his human functioning is compromised. Here, I use the word ‘functioning’ differently than Wakefield does. When I speak of human functioning in this case, I am not talking about a natural mechanism or process selected through evolutionary processes. Rather, I use it in the Aristotelian sense according to which human beings have certain capabilities, and having the opportunity to develop these capabilities in varying degrees contributes to human flourishing. When one or more of these capabilities is absent, seriously impaired, or severely underdeveloped in a human being, human functioning is compromised. Some of the capabilities that have been suggested include practical reason, cognitive abilities, social affiliation, and the ability to laugh and play. (Nussbaum 1992) When a symptom, or cluster of symptoms, is present to
a certain degree, it can causally interfere with human functioning along one or more dimensions. Moreover, if human functioning along these dimensions falls below a given threshold, a disorder is present, as it was in Jones's case. In a case like Smith's, human functioning can be diminished, but still above a crucial threshold, and while it is unfortunate for the human being in question, it is not the case that he has a disorder.

At this point, a skeptic about the approach sketched here might object that an account of disorder in terms of human functioning looks circular as soon as we recognize that one of the important capabilities in human functioning is good health or the absence of disorder, illness, and disease. The circularity objection, as I shall call it, rests on the assumption that we are trying to define disorders in terms of human functioning which in turn must be defined, at least, partly in terms of the absence of disorder. Indeed, if we were offering a definition of the concept of disorder in terms of the real essence of disorders, we should worry about this objection because the real essence is supposed to reveal the nature of disorders and to explain what disorders are. Similarly, a definition of the concept of human functioning in terms of the essence of human functioning should reveal its nature and explain what it is. Surely, it is unsatisfactory to explain disorders in terms of something else, viz. human functioning, which is partially explained in terms of disorders or lack thereof. However, I am not in the business of providing such a definition of either disorder or human functioning. That is, I am not providing a Lockean account of disorder, or for that matter of human functioning, so the constraints on what counts as a satisfactory Lockean account do not apply here. Indeed, I suspect that human functioning is also a homeostatic property cluster kind, and I will sketch my reasons for thinking so shortly. Furthermore, disorders and human functioning are overlapping

---

2 Throughout this paper, I have only discussed Lockean essences, viz. real and nominal essences, and I write here as though one might try to understand human functioning by identifying its real essence. To do this, one would investigate human functioning, identify its nature, and articulate necessary and sufficient conditions for a certain human state to count as human functioning, or in plainer language, for a certain life to count as a good human life. An alternative to the Lockean essentialist approach is the Aristotelian essentialist approach. While concepts or properties have Lockean essences, individuals have Aristotelian essences. Roughly speaking, an Aristotelian essence is the property or properties without which an individual would cease to exist. For example, being human would be one of my essential properties in this sense. I would cease to exist if I were not human. On an Aristotelian approach to understanding human beings, human functioning includes all of the capabilities without which something would cease to be human. That is, the properties and capabilities that constitute human functioning are essential to any human individual. If one or more of these capabilities is absent, that individual ceases to exist as a human, or to exist at all.
HPC kinds. What matters for understanding disorders is not that we have a prior, independent account of one so that we can explain the other in terms of it. Rather, we need to have some, perhaps approximate, understanding of how each interacts causally with the other, and through this insight, we will come to learn about both disorders and human functioning.

As I suggested earlier, human functioning is best understood as a battery of capabilities which human beings may develop to varying degrees so that they may flourish to varying degrees. Although possession of all of these capabilities to a reasonable degree is plausibly sufficient for human flourishing, it is less obvious that any one of these capabilities is necessary for human functioning. Nussbaum (1992) thinks that the absence of any one of these capabilities in an individual is sufficient grounds for us to deny that this individual is human, I endorse a more moderate view. To be sure, the absence of practical reasoning in an individual who otherwise possesses most other distinctively human characteristics and capabilities is grounds for thinking that this individual will not flourish to the same extent as he would if he had the ability to reason practically, but it would be extreme to claim that this individual is not human or to claim that he cannot function as a human at all. I am more inclined to say that he is a human living with a disability, or even, a disorder, and that he can function as a human, albeit to a very limited extent. Indeed, severely mentally retarded individuals or people living with neurological disorders that compromise their practical reasoning faculties do not cease to function as humans, but they certainly do not flourish as they would if they had full command of their practical reasoning. Now, if we imagine a creature who failed to have quite a few distinctively human capabilities, it might seem more reasonable to deny that the creature is human or even a live animal. So, on my view, human functioning or flourishing is a cluster of human capabilities.

Moreover, the capabilities that cluster together tend to do so because enhancement of some causally promote the development of others. Take for example the abilities to use one’s perceptual and cognitive faculties, to shelter one’s self, to adequately nourish one’s self, and to avoid pain. Quite plausibly, the better developed one’s ability to use one’s perceptual and cognitive faculties, the better able one is to feed one’s self, to find shelter, and to navigate one’s physical and social environments in a way that minimizes painful experiences. Similarly, the better nourished one is, the greater one’s energetic resources for performing physical tasks such as finding food and shelter and for performing mental tasks such as plotting strategies for adequately meeting one’s needs and avoiding pain. The better one is able to avoid pain, the
less distracted one is from developing one's other capabilities. And, so forth. Of course, deficits in any of the capabilities will tend to lead to deficits in others. If one is unable to adequately nourish one's self, one will not have the energetic resources to think or concentrate, let alone to procure shelter, clothing, or ironically, food. If one's perceptual and cognitive abilities are underdeveloped, it is more difficult to take care of one's self and to ensure that one is adequately nourished and clothed. In short, development or underdevelopment of any human capabilities contributes to the development or underdevelopment of others. That is, human capabilities are sustained by homeostatic mechanisms. They cluster together because the presence and development of some tends to favour the presence and development of others.

Just as being in good health tends to promote the presence and development of other human capabilities such as the ability to use one's mental faculties, to take care of one's self, to seek the company of others, and to laugh and play, poor health can impair one's ability to develop these other human functions. Similarly, taking good care of one's physical, emotional, and social needs tends to promote good health, and living in unhygienic conditions or in social isolation or in a life devoid of laughter and play will tend to increase one's susceptibility to illness, disease, and disorder. The dimensions along which human flourishing may be compromised by disorders can vary widely. Some psychiatric and neurological disorders primarily limit the human capability for affiliation with other people. Some disorders are associated with chronic pain so that the human function of pain avoidance is compromised. Moreover, the degree to which human flourishing is limited by disorders can vary widely. Some disorders will have rather remarkable effects on human flourishing, such as cancer, AIDS, and schizophrenia. Others will exert limited influence on human capabilities, such as psoriasis. Thus, part of what it is for a condition to be a disorder is that it homeostatically interacts with human flourishing, and that it do so in degrees.

At this point, it is important to make some clarificatory remarks. In my discussion of Cases 1 and 2, I suggested that the crucial difference between Jones and Smith is that the former's condition interfered with his flourishing as a human being, and that this was a basis for distinguishing between the disorder and the non-disorder described in Cases 1 and 2, respectively. While it is tempting to take these cases to provide evidence that the crucial distinction between disorders and non-disorders is that the former always involves significant interference while the latter does not, this is not the proper conclusion to draw in light of the present discussion about the homeostatic relationship between disorder property clusters and human functioning property clusters. Given that the relationship is so characterized, it is better
to say that disorders often, but do not always, involve a significant impairment of human functioning, and that the demand for one or more properties that will, in all cases, clearly distinguish the disorders from the non-disorders is one that cannot, and needn’t, be met by a homeostatic property model. Indeed, this model predicts that there is no such property, or set of properties.

3.6 Conclusion

In this chapter, I have argued that Lockean accounts of disorder are inadequate, and that the homeostatic property cluster model is more promising. I have also argued that disorders are not purely natural kinds. The property clusters with which disorders are associated include psychological, behavioural, and social properties. Moreover, part of what it is to be a disorder is that relevant clusters of biological, psychological, behavioural, and social properties interact homeostatically with property clusters corresponding to human functioning.
Bibliography


Kleinman, A. & Good, B., eds (1985), *Culture and Depression*, University of California Press, Berkeley, CA.


