CHAPTER 3

THE PROBLEM OF THEORY-LADENNESS

I have argued that at least some types of mental disorder may well be natural kinds. Still, it may be difficult to produce a classification system that reflects their natural structure. This chapter examines the potential problems for the D.S.M. if observation is theory-laden. The next chapter examines the ways in which the D.S.M. has been shaped by pressures that arise from the ways in which it is used.

If observations are affected by the observer's theoretical beliefs this can be expected to reduce our chances of gaining empirical knowledge. Theory-ladenness implies that we can never have direct access to the world, and that appeals to "observations", "experimental results", and "facts" can never be definitive. Whenever someone reports seeing something we will always be left with the possibility that another observer, armed with different theoretical beliefs, might have reported something quite different. As such, we will have less reason to believe our hypotheses to have been reliably confirmed, or falsified, by empirical data. The ultimate consequence is that we will have less reason to believe that our theories are true.

Theory-ladenness would cause the most problems where our theories are worst. Theory-ladenness implies that when an observer has a false theory their observations will be distorted by their false beliefs. Unfortunately it is likely that much current psychiatric theory is mistaken. There are many competing frameworks for understanding mental illness – biomedical, psychoanalytic, behavioural, relational, and so on. At many points these approaches are inconsistent, so not all of them can be correct. Thus many practitioners will hold false theories. Even within the biomedical framework, the approach that's arguably most closely linked with the D.S.M., competing theories abound. With few exceptions, the causes of specific mental disorders are contested. Again, this competition implies that the odds are that the theories that a psychiatrist holds are mistaken. If observation in psychiatry is theory-laden, then psychiatrists' false beliefs can be expected to distort their observations of their patients and prevent them from seeing the true similarities and differences between types of mental disorder. There will thus be reason to doubt that the categories included in the D.S.M. will correspond to natural kinds of disorder.

In addition to these epistemic problems, if observation is theory-laden there will be ethical and political implications, at least within the human sciences. People's lives are often affected by the ways in which professionals perceive them. If Mrs Jones is seen as deeply unhappy and as possibly suffering from clinical depression, and Mr Smith is seen as angry and potentially dangerous, then these judgements are likely to have practical consequences. If psychiatrists' perceptions of patients are affected by their theoretical beliefs (which may be false) then there is a danger that patients will be perceived inaccurately and treated wrongly as a result. What's worse, evidence from the history and sociology of science suggests that in many cases errors in theories will be systematic as opposed to random. All too often, when we look back at scientific theories that in their day were believed by good and clever people, we can see that the theories were systematically biased against the disadvantaged.[1] The theories of the eugenicists constitute only the most glaring example.[2] We have no reason to think ourselves better or cleverer than our predecessors. A pessimistic induction thus suggests that the theories that we believe today will also frequently turn out to be not only false but systematically biased against the less powerful. If our observations are theory-laden, all too often we will see others falsely and unfairly.

In addition to these wide-ranging epistemic and political problems, if observation is essentially theory-laden there will be problems specific to the D.S.M. The D.S.M.-III committee set out with the intention of creating an atheoretical classification system.[3] Mental health professionals operate within numerous different theoretical frameworks, and there were fears that a D.S.M. based on any one particular theory would alienate many practitioners. The D.S.M.-III committee sought to avoid this problem by basing the D.S.M. on no theory at all.[4] Instead they set out to produce a purely descriptive classification system that makes no reference to hypotheses regarding the aetiology of disorders. However, if observation is necessarily theory-laden then it is impossible to construct an atheoretical classification system. The aim of the D.S.M.-III committee will simply be unattainable.

Within the mental health literature the D.S.M.-III's claim to be atheoretical has come in for much criticism. Much sport has been derived from pointing out the various ways in which the D.S.M.-III has manifestly failed in its aim. In their review of the D.S.M.-III, Cooper and Michels point out that many of the diagnostic criteria included in the D.S.M. require fairly complicated and prima facia theory-laden inferences to be made.[5] They cite "identity disturbance", which is a key symptom of Borderline Personality Disorder, as an example. Cooper and Michels also note that the way in which diagnostic categories are grouped into higher-level categories is informed by theoretical commitments. For example, the D.S.M.-III considers "dysthymic disorder" to belong with the affective disorders, whereas some psychodynamically inclined clinicians would consider it more akin to the personality disorders. To add to Cooper and Michels observations, many D.S.M. diagnoses contain exclusion clauses, which are also clearly informed by theory. Agoraphobia, for example, can only be diagnosed if the symptoms cannot be better explained by a major depressive episode, Obsessive Compulsive Disorder, Paranoid Personality Disorder, or Schizophrenia. These disorders trump agoraphobia because the pathological process underlying them are thought to be more "serious" or "deep-

[1] See, for example, Bleier 1984, Keller 1985, Harding (ed.) 1993.
[2] Gould 1983 discusses these debates.
[3] A.P.A. 1980 p.7
[4] A.P.A. 1980 p.7
[5] Cooper and Michels 1981

rooted" than that underlying agoraphobia. Such judgements quite clearly assume a theory of mental disorders.

In addition to noting that the D.S.M.-III is not actually theory-free, it has become commonplace for commentators to point out that the possibility of constructing an atheoretical classification system can be doubted on the grounds that observation is in some sense theory-laden.[6] Under this barrage of criticism the claim to be atheoretical was dropped from the introduction to the D.S.M.-IV. Still, many of the sets of diagnostic criteria included in the D.S.M.-IV remain the same as in the D.S.M.-III, and new criteria sets follow the same style as those that have been inherited. As such, if the aim to be atheoretical led to mistakes in the approach of the D.S.M.-III these mistakes are likely to also be present in the D.S.M.-IV.

1.THE PROBLEMS OF THEORY-LADENNESS CLARIFIED

1.1 What Is A Theory?

The question of what counts as "theory" casts its shadow over the whole of this chapter. Providing an account of what makes something a theory could easily take a book in itself, and so will not be attempted here. In any case, prototypical theories are easy enough to recognise. They are characteristically acquired through learning or invention, and are vulnerable to replacement by successor theories. Thus we have Newton's theory of gravitation, psychoanalytic theory, and so on.

Philosophers writing about theory-ladenness have sometimes worked with a notion of "theoretical knowledge" that is far broader than this, however. For example, it seems that our perception of the world depends on innate, hard-wired brain mechanisms that process the raw sensory inputs. One such mechanism "recognises" the kinds of 2D projections that 3D objects produce on the retina and enables us to perceive the 3D object. Paul Churchland considers the operations of such mechanisms to provide an example of one way in which our perceptions are theory-laden.[7] In a similar vein, Karl Popper writes, "there is no sense organ in which anticipatory theories are not genetically incorporated".[8]

I think that Churchland and Popper have made a mistake in thinking that such mechanisms provide evidence of theory-ladenness. The existence of a hard-wired brain mechanism that "recognises" the kinds of 2D projections that 3D objects produce on the retina no more implies that a perceiver possesses a theory of 2D projections than the fact that an organism respires proves that it possesses theoretical knowledge concerning its need for oxygen. Generally it is true that biological mechanisms are such that much theoretical knowledge would be required to design an artificial replica. However, this does not mean that the organism itself possesses

[6] Faust and Miner 1986, Carson 1991 p.306, Millon 1991, Morey 1991 p.291, Goodman 1994
[7] Churchland 1988 p.170
[8] Popper 1972 ch.2 §18

theoretical knowledge. That the mechanism that recognises 2D projections is located within the brain makes no difference and should not mislead us.

Other writers have taken "theoretical beliefs" to include beliefs such as "There is an external world filled with physical objects" and "Other people have minds".[9] I'm not sure whether such beliefs should be classed together with the theoretical beliefs that I wish to consider here. It seems to me possible that such common-sense beliefs might differ from prototypical theories in being hard-wired, whereas prototypical theories are learnt or invented, and are vulnerable to replacement by successor theories. In any case, the fact that such common-sense beliefs are universally held means that their possession will not make some scientists observe things that others do not. As all scientists hold such beliefs their effects on perception (if any) will be uniform across all observers. As such, the discussion here will be restricted to considering whether observation is affected by more prototypical theoretical beliefs, such as "All masses fall at the same rate" or "Chronic schizophrenics often have flattened or inappropriate emotional responses".

1.2 Three Kinds Of Theory-Ladenness

Three distinct, but inter-related, claims can be teased out from the general thought that observation is in some sense theory-laden. First, it may be claimed that perception itself is theory-laden, where "perception" here refers to what an organism actually sees, hears, or senses in some other way. Perceptions are neither wholly determined by stimuli nor wholly determine observers' judgements of what they have seen; when looking at the duck/ rabbit picture, for example, two observers are presented with the same stimulus, one may perceive a rabbit while the other perceives a duck, and both may judge what they have seen to be an ambiguous figure. Those who hold that perception is theory-laden claim that when scientists view the same stimulus their theoretical disagreements may cause them to perceive different things.

Second, it can be claimed that the language in which observation statements are couched is theory-laden. This would imply that even if scientists holding different theories had the same perceptions they would produce reports of their perceptions that differ in meaning.

Third, it can be claimed that where scientists choose to direct their attention will be influenced by their theoretical beliefs about what is important. A psychoanalyst will choose to investigate childhood experiences, sexual fantasies, and so on, while a biologically-orientated psychiatrist will choose to investigate brain scans and neurotransmitter levels. As the two scientists seek out different stimuli they will end up making different observations.

If observation is theory-laden in any of these three ways there will be implications for the D.S.M. Although, as we shall see in the next chapter, the D.S.M. can be shaped by economic, social, and political considerations, to a large extent it is

[9] For example Quine 1960 p.22 "The positing of those extraordinary things [molecules] is just a vivid analogue of the positing or acknowledging of ordinary things: vivid in that the physicist audibly posits them for recognised reasons, whereas the hypothesis of ordinary things is shrouded in prehistory."

based on empirical research. The committees that decide the categories to be included in the D.S.M., and write the sets of diagnostic criteria that define these categories, rely on empirical studies to inform their decisions. This is directly the case when the committees review the empirical literature. It is still indirectly the case when they rely on expert opinion (as expert opinion will itself be based on knowledge of the empirical literature). If observation is theory-laden in any of the three ways outlined above, then this empirical literature will be questionable. One will be left with the possibility that if investigators with different theoretical beliefs had performed the studies different results would have been obtained.

In what follows, I will examine each of the three ways in which observation might be theory-laden in turn. In each case, my aim is to determine whether it is true that observation is theory-laden in the specified way, and if so the extent to which this is likely to reduce the chances that the categories included in the D.S.M. will correspond to natural kinds of disorder.

2. IS PERCEPTION THEORY-LADEN?

Here I take "perception" to refer to what an organism sees, hears, or otherwise senses. Philosophers brought up on the work of Thomas Kuhn and Norwood Hanson tend to believe that perception is theory-laden.[10] They hold that scientists with different theories see the world in different ways. The claim that perception is theory-laden has become so entrenched within the philosophy of science that the evidence supporting it is seldom subjected to scrutiny. Here, however, I will re-examine the evidence that has traditionally been offered, and argue that it is insufficient to demonstrate that psychiatrists' perceptions of mentally ill people are theory-laden.

Whether perception is theory-laden is commonly held to be an empirical question. As such, philosophers who hold that perception is theory-laden back up their claims by citing experiments that are supposed to indicate that what a subject perceives is affected by their expectations under certain conditions. Jerome Bruner and Leo Postman's anomalous playing card experiment and George Stratton's inverting lenses experiment are the classic examples.[11] Bruner and Postman asked subjects to identify briefly presented playing cards. Some of the cards used were anomalous, for example, a black four of hearts. It was found that subjects took an average of 28 milliseconds to correctly identify normal cards and 114 milliseconds to identify anomalous ones. Anomalous cards presented for shorter periods of time tended to be mistaken for normal cards. For example, someone shown a black four of hearts might report seeing a black four of clubs. The conclusion often drawn is that people see what they expect to see,[12] and that perception is thus influenced by theory. George Stratton wore an inverting lens over one eye for 87 hours over eight days. His other eye was covered, and when not wearing the lens he was blindfolded.

[10] Kuhn 1970, Hanson 1969

[11] Bruner and Postman 1949, cited by Kuhn 1970 p.113, Goodman 1978 p.14. Stratton 1897 cited by Kuhn 1970 p.112.

[12] For example Kuhn 1970 p.113

At first Stratton found that everything looked to be upside-down. He had problems guiding his actions and bumped into things. However, after a few days he was able to control his actions, and by the eighth day Stratton's world at least occasionally appeared normal to him. Thomas Kuhn takes this experiment to show that the way we see the world is not fixed; we can learn to see the world differently.[13]

It is debatable whether Bruner and Postman's experiment actually shows that perceptions of playing cards are affected by beliefs about playing cards, or whether Stratton's experiment shows that perceptions of the environment can be affected by beliefs about orientation.[14] Here, however, for the sake of argument, I shall accept that these experiments demonstrate cases of theory-ladenness but argue that there are problems with extrapolating from these experiments to claim that perception in psychiatry is theory-laden.

I will argue that it is unjustifiable to extrapolate from the experiments discussed above to the claim that perception in psychiatry is theory-laden by arguing that perceiving people is importantly unlike perceiving inanimate objects, such as playing cards. There are various ways in which this argument could be made. Hopefully, the relationship between a psychiatrist and a patient (even between a research psychiatrist and an experimental subject) is quite different from that between a participant in an experiment and a playing card. In the first case two human beings interact. In the second case a disinterested observer views a passive stimulus. On many accounts of perception, such as that put forward by J.J.Gibson, and those linked with the extended or embodied/embedded cognition movement, interaction is important for perception and this difference will be key.[15] I have some sympathy with these accounts of perception. However, they are controversial. While I am sympathetic to these accounts, those who are tempted to claim that perception in psychiatry is theory-laden need not be, and so I will not rely on them for my argument here. Instead, I will take a different tack, and argue that even *if* one thinks of a psychiatrist as viewing "stimuli" there is reason to think that perceptions of patients need not be theory-laden. I will support this claim via a consideration of neuropsychological evidence that suggests that different types of perception are dependent on different areas of the brain, and are probably processed differently. This means that it is possible that perception is theory-laden in some domains but not in others. In particular, it might be the case that our perceptions of playing cards are theory-laden, but that our perceptions of people are not.

[13] Kuhn 1970 p.112, Churchland 1988 makes similar use of this experiment.

[14] Gilman 1992 p.294, ftnt 4 suggests that the implications of the Bruner and Postman experiment are limited. Gilman also provides a detailed analysis of two other "New Look" experiments cited in footnotes by Kuhn (Bruner, Postman and Rodrigues 1951 "Expectation and the perception of color", and Hastorf 1950 "The influence of suggestion on the relationship between stimulus size and perceived distance"). Some experimenters have repeated Stratton's experiment and do not interpret the results as showing that the perceptions of the lens-wearer change. G.Brown (1928) and I.Kohler (1964) think the lens-wearer's world always appears distorted to them, but that after a while they get used to this and learn techniques to enable them to control their bodily movements. C.Harris (1963) holds that adaptation is proprioceptive not visual. Jerry Fodor (1988 p.193) suggests that even if the perceptions of the lens-wearer change, this need not be taken to show that perceptions are in general flexible

[15] Gibson 1979, for an outline of the extended cognition approach see Clark 1996.

Some of the strongest support for the claim that different types of perception are processed in different areas of the brain comes from clinical case studies of people who have suffered brain lesions. Depending on the location of the lesions, certain types of perception may remain while other types are lost. One of the most famous such cases concerns a patient, D.F., who suffered brain damage as a result of carbon monoxide poisoning.[16] D.F. is unable to discriminate size, shape or orientation and is thus unable to recognise objects, places, or people. Despite this disability, D.F. is still able to perform actions that require perceptual information. For example, although when she is presented with a letter-box type slot D.F. is unable to say what orientation it is in, if she is asked to insert her hand through the slot D.F. can reach for it with her hand correctly positioned. It appears that D.F. can make use of perceptual information to guide her actions, although she is unable to recognise objects. Conversely, some other brain-damaged patients have problems grasping objects, but are able to recognise them.[17] This suggests that the type of perception involved in recognising objects ("perception-for-recognition") and the type of perception involved in guiding actions ("perception-for action") may be processed in different parts of the brain.

Not only is it plausible that different types of perception are processed in different areas of the brain, there is the possibility that some types of perception make more use of "top-down" processing than others. In "top-down" processing, higher cognitive information is used in processing lower-level data. In sentence recognition, for example, it seems that information about context enables us to "hear" appropriate words even when the incoming noise is somewhat degraded. Processing is "bottom-up" if no feedback from higher to lower levels is involved. Perception can only be theory-laden if it involves top-down processing. This is because theory-ladenness requires theoretical beliefs (which rely on higher cognitive processes) to affect perceptions. This being said, "involving top-down processing" and "theory-laden" are not synonymous. It might be the case that perception makes use of some higher cognitive information that yet falls short of counting as theoretical knowledge. Such an account has been put forward by Jerry Fodor in *The Modularity of Mind* (1983). Fodor holds that some top-down processing is required for perception. There are brain mechanisms that enable 2D retinal images to be interpreted as 3D scenes, for example. However, the perceptual "module" that contains such mechanisms is still isolated from theoretical knowledge and so perception is not theory-laden on Fodor's account. That a type of perception involves top-down processing is necessary but not sufficient for it to be theory-laden.

Certain illusions are generally thought to occur as a side-effect of top-down processing. These include the Müller-Lyer illusion, in which the apparent length of lines varies with the arrangement of arrows at the ends of the lines, and the Titchener circles illusion, in which the apparent size of a central circle is affected by the size of the circles surrounding it. To examine whether perception-for-action involves top-down processing, experimenters have examined what happens when

[16] Milner 1997

[17] Milner and Goodale 1995 ch.4

subjects are asked to grasp the central line of a Müller-Lyer figure, or the central circle in a set of Titchener circles. If perception-for-action is not fooled by the illusions this would suggest that perception-for-action does not employ top-down processing, and so could not be vulnerable to theory-ladenness. Unfortunately, the results of such experiments have been mixed. In a widely cited study Aglioti et al (1995) found that when subjects reach to grasp the central circle in a set of Titchener circles their grasp aperture corresponds to the actual rather than to the perceived size of the target. This would indicate that perception-for-action does not employ top-down processing. Similar results were obtained by Haffenden and Goodale (1998), and, using the Müller-Lyer illusion, by Post and Welch (1996), and Otto-de Haart, Carey and Milne (1999). More recently, however, it has been suggested that these results were merely artefacts of the experimental set-up, and other experimenters, using slightly different set-ups, have found that perception-for-action can be fooled by illusory effects.[18] These results, in their turn, have also been contested, and recent review articles argue that the issue can only be decided via further research.[19]

Whether or not it turns out that perception-for-action is vulnerable to illusory effects, for our purposes the central message is clear: It seems there are different types of perception that are processed differently. This opens up the possibility that some types of perception might be theory-laden while other types are not. As such, it is a mistake to simply ask whether perception as a whole is theory-laden, rather one must specify the types of perception one is interested in and then review the evidence that is relevant to that case.

Perception-for-action and perception-for-recognition appear to be distinct types of perception. There are also perceptual systems dedicated to other tasks. It is likely that face recognition (e.g. recognising Tony Blair) is dependent on one such system. Evidence for this hypothesis comes chiefly from the clinical condition of prosopagnosia, in which patients are unable to recognise faces but can recognise other stimuli. As is so often the case, the evidence is not clear-cut. Prosopagnosia is a rare condition, often the damage to the brain is diffuse, and most patients suffer from other disabilities besides an inability to recognise faces. There is also some debate whether prosopagnosia should be considered as primarily a perceptual or a memory deficiency; it might be the case that patients can’t recognise faces because they can’t perceive them, or it might be because they can’t remember what their friends and relatives look like. However, there is also other evidence that suggests that the perception of faces is significantly different from the perception of other stimuli. Behavioural studies have found differences between face and object recognition, for example, inverted faces are far harder to recognise than inverted objects.[20] Moreover, various brain-imaging studies show that face-perception and object-perception are processed in different areas of the brain in normal subjects.[21] When considered together the evidence strongly suggests that face recognition is processed by a dedicated system.

[18] Franz et al. 2001.
[19] Bruno 2001, Carey 2001.
[20] Yin 1969
[21] Biederman and Kalocsai 1997

There are also cases of brain-damaged people who can recognise faces but are unable to recognise facial expressions.[22] These people can recognise a photo as being of Tony Blair, say, but they are unable to tell whether he is happy, angry, or bored. This suggests that the perceptual-system that recognises facial expressions is different to that which recognises faces. The situation gets still more complicated, as the recognition of different emotions may depend on several distinct systems. Damage to the amygdala can result in a specific inability to recognise fear,[23] whereas people who suffer from Huntington's chorea can have specific difficulties recognising disgust.[24]

That different types of perception are processed differently in different areas of the brain is important for the discussion here because it opens up the possibility that some types of perception might be theory-laden while other types might not be theory-laden. Thus, rather than asking whether perception in general is theory-laden we must be more specific and ask whether a particular type of perception is theory-laden.

2.1 Is Perception In Psychiatry Theory-Laden?

Here I am interested in theory-ladenness in so far as it may affect the D.S.M. Given that different types of perception may be processed differently by the brain, any evidence suggesting that, say, perceptions of electron tracks are theory-laden will only be of peripheral interest. Instead it is necessary to consider specifically whether the types of perception involved in the collection of the data on which the D.S.M. is based are theory-laden. For the most part, the descriptions of conditions included in the D.S.M. are based on psychiatrists' observations of psychiatric patients. Thus we must ask whether perceptions of people are affected by theoretical beliefs. Psychological studies examining this question are scarce. However, there is a series of studies that examine whether perceptions of facial expressions are theory-laden, which shed some light on this question.

2.1.1 Studies Of Emotion Perception

Whether our perceptions of facial expressions are theory-laden is relevant to evaluating the importance of theory-ladenness within psychiatry. Psychiatrists often make judgements about patients' emotional states, and these judgements will in large part be dependent on the psychiatrists' perception of the patients' facial expressions. In addition, psychiatrists have theoretical beliefs concerning the types of emotions particular types of patients may be expected to exhibit. Chronic schizophrenics are expected to show flattened or inappropriate emotional responses, patients suffering from depression are expected to be miserable, those suffering from mania to be cheerful or irritable, and so on.

[22] Young et al 1993
[23] Adolphs et al 1994
[24] Spregelmeyer et al 1996

Unfortunately, although there are some studies that examine whether perceptions of facial expressions are affected by our expectations, they are scarcer and harder to interpret than one might have hoped. A number of experimenters have examined the effects of contextual information on subjects' perceptions of emotions. In such studies, subjects are shown a face, for example a woman smiling, and are given information regarding the context, for example they might be told that the woman's son has just died, and then they are asked to judge the emotion that the person is experiencing. Such experiments are relevant to determining the extent to which our perceptions of emotions are theory-laden. If the contextual information is found to influence the subjects' perceptions this will be because they have certain expectations regarding the emotions that people will feel in certain situations, for example, people are expected to be sad on the death of a relative, happy when they've been given a present, and so on. Thus, if the subjects' perceptions are affected by the contextual information this will indicate that the perception of emotions is theory-laden, and if they are not this will suggest that the perception of emotion is not theory-laden. Unfortunately the results of experiments have been inconsistent. Fernández-Dols and Carroll (1997) review eighteen studies. Of these, seven found that subjects' perceptions were unaffected by the contextual information, while the rest found that the contextual information had some effect. As the results of the experiments are mixed, no conclusion can be drawn.

A possible weakness of the studies reviewed by Fernández-Dols and Carroll is that they lump together scores for emotion recognition that are achieved when viewing various different emotions. As mentioned earlier, there is some evidence that the recognition of different emotions may depend on different neurological systems. This would mean that our perception of some emotions might be theory-laden, while our perception of other emotions might not be theory-laden. There is already evidence that the extent to which our ability to remember facial expressions is theory-laden depends on the facial expression in question. Woll and Martinez (1982) found that labelling pictures with inconsistent labels, for example labelling a picture of someone smiling as "angry", affected subjects' ability to recognise the picture later if the emotion depicted was positive or neutral, but not when the pictures were of negative emotions. Woll and Martinez's results may be due to memory for various emotions being differentially theory-laden, rather than to perception of various emotions being differentially theory-laden, but they are at least suggestive. An ability for observers to accurately perceive negative emotions regardless of their expectations might be expected on evolutionary grounds; plausibly those humans who failed to recognise their neighbour's fear because they expected the Sabre Tooth Tiger to be safely asleep got eaten straight after their neighbour.

2.1.2 Summary

In this section I have shown that the evidence traditionally cited by philosophers in support of the claim that perception is theory-laden is unsatisfactory. Determining whether perception is theory-laden is a task for empirical science and is far harder than many philosophers of science have thought. There is evidence that different types of perception are dependent on different areas of the brain and may be processed in different ways. This opens up the possibility that some types of perception might be theory-laden while other types might not be theory-laden. The most plausible examples of theory-laden perception come from highly technical domains - where experimenters look at microscope slides, electron tracks, or sun spots, for example. However, even if such perception is theory-laden, it is possible that perception in other domains is not. Research psychiatrists will spend their time looking at a variety of different types of stimuli. Importantly, much of their time will be spent looking at patients Unfortunately, evidence relevant to the question of whether psychiatrists' perceptions of their patients are theory-laden is insufficient to allow a judgement one way or the other. Still, it is worth noting that on evolutionary grounds one might expect even those who lack any relevant theory to be able to directly perceive facts that are of relevance to survival. Thus it should come as no surprise if humans need a theory to recognise cells on a microscope slide, but are innately disposed to recognise emotions, or indeed, symptoms of illness. In the absence of conclusive empirical studies, however, the final conclusion of this section can only be that the question of whether our perceptions are theory-laden is not closed, as is often assumed, but on the contrary should continue to be a live issue.

Even if scientists with different theories perceive the same thing, however, it may be that as soon as they formulate observation reports theory-ladenness creeps in. To see whether this would be the case, we must turn to consider whether language is necessarily theory-laden.

3. ARE OBSERVATION REPORTS NECESSARILY THEORY-LADEN?

Many philosophers have claimed that the language in which observation statements are couched is necessarily theory-laden.[25] As a consequence, even if scientists with different theoretical orientations perceive the same thing, their observation reports will have quite different meanings. Regardless of whether perception itself is theory-laden, as soon as scientists try to communicate their findings problems with theory-ladenness would emerge.

Those philosophers who claim that language is necessarily theory-laden claim that the meanings of the terms used in an observation statement are at least partially dependent on theory. To use Popper's example, suppose someone reports "Here is a glass of water".[26] This might seem like a straightforward observation statement. However, Popper claims that it is part of the meaning of terms such as "glass" and "water" that these are kinds of stuff that show law-like behaviour. If it turned out that the stuff in the glass could be ignited with a match, for example, then that would show that it was not water after all. Popper concludes that "Here is a glass of water" is not merely a report of what is seen, but assumes much theoretical knowledge.

[25] For example, Popper 1959, Kuhn 1970, Fleck 1979, Churchland 1988.

[26] Popper 1959 Ch.V §25. Fleck 1979 p.90 gives a similar discussion of the statement "Today one hundred large, yellowish, transparent and two smaller, lighter, more opaque colonies have appeared on the agar plate.".

3.1 Three Possible Ways Of Sidestepping The Problems Of Theory-Ladenness

3.1.1 Nagel's Suggestion

In his 1971 paper, "Theory and Observation", Ernst Nagel shows that the theory-ladenness of observation statements need cause no problems in practice, however. Nagel accepts that observation statements presuppose various theories and background information. Nevertheless he holds that theories can be tested by observations. This is because the theories assumed by the observation statements that report the results of some experiment will generally be different from the theory that the experiment is testing. For example, Newton conducted various experiments with a glass prism to test his theory that white light is made up of coloured light. The result of the experiment can be reported by an observation statement: "When light is shone into one side of a glass prism rays of red, green and purple light can be seen on the other side". This description is theory-laden. Calling something a "glass prism", for example, assumes that the prism is actually made from a particular substance. Nonetheless, the theories assumed by the description do not include Newton's theory of light. The observation statements are independent of this particular theory and so can test it.

Nagel's suggestion also implies that scientists with different theoretical orientations can often still mean the same thing by their observation statements. So long as the theories assumed by a description do not include those theories about which the scientists disagree, their differing theoretical beliefs will not lead to any difference in meaning. To illustrate, suppose a psychoanalyst and a biologically-orientated psychiatrist are discussing a patient. One says "Mrs Jones has been crying". It can be accepted that this is a theory-laden statement that assumes, for example, that Mrs Jones is a human being with mental states rather than a cunningly constructed robot. Still, following Nagel, even if "Mrs Jones has been crying" is theory-laden the psychoanalyst and biologically-orientated psychiatrist can mean the same thing by the statement. Admittedly lots of their theoretical beliefs and assumptions will be different, but others will be the same. So long as the beliefs assumed by "Mrs Jones has been crying" are amongst those they share, they should experience no problems in communicating.

3.1.2 Using Other Forms Of Communication

Philosophers who argue that observation statements are theory-laden tend to assume that this implies that scientists with different theories will necessarily have problems communicating with each other. This need not be the case, however, because we have other ways of communicating with people apart from linguistically. It may turn out that these other forms of communication are not theory-laden. If so, it might be possible to employ non-linguistic forms of communication to side step any problems that might result from using observation statements.

To take an example, suppose our psychoanalyst and biologically-orientated psychiatrist are discussing a patient's symptoms. The biologically-orientated psychiatrist is about to describe the patient as being anxious, but then he remembers that this description may have a different meaning for his colleague than it does for himself. So, instead of describing the patient's symptoms to his colleague and risking misunderstanding, he invites the psychoanalyst to have a look at the patient, so that the psychoanalyst can see the patient's symptoms for himself. No observation statements are used by this method of communicating information, and so whether they are theory-laden is irrelevant. The biologically-orientated psychiatrist hopes that by showing the patient to the psychoanalyst he has ensured that both of them have the same information regarding the patient's symptoms.

Will the biologically-orientated psychiatrist succeed in his aim? Amongst other things this depends on whether perception is theory-laden. If perception is theory-laden then, as the psychoanalyst and the biologically-orientated psychiatrist have different theoretical beliefs, they will probably perceive the patient differently. As discussed in the previous section, at present there is insufficient information for it to be possible to judge whether perceptions of people are theory-laden. Still, until it is shown that perceptions of people are theory-laden, the possibility that the biologically-orientated psychiatrist has succeeded cannot be ruled out on the basis of the theory-ladenness of perception.

In addition, whether the biologically-orientated psychiatrist can get the psychoanalyst to see the same symptoms as himself depends on whether the two clinicians look at the same aspects of the patient's behaviour. People are complicated stimuli, and it is plausible that an observer cannot take in all aspects of their appearance and behaviour. When the psychiatrist invites his colleague to "see for himself" the psychoanalyst still has to decide whether he is supposed to be looking at the way in which the patient is shaking, at their facial expression, at their freckles, or at the way they've tied their shoelaces - amongst other possibilities. Now it is plausible that as human beings we naturally find certain features of a person's behaviour or appearance salient. For example, people normally notice when others have facial twitches or dodgy looking rashes. However, there may well be other aspects of the patient's appearance or behaviour that observers will not naturally find salient but that they may be primed to be alert to if they possess the right theoretical beliefs. Whether the scars on someone's wrists go across or up, for example, will probably only be noticed by someone who knows that slashing upwards is a lot more dangerous. As observers with different theoretical beliefs may well be primed to notice different things, the different theoretical beliefs of the biologically-orientated psychiatrist and the psychoanalyst may well result in them noticing different aspects of the patient's behaviour and appearance.

Thus, in many cases, if a scientist is to get a colleague with different theoretical commitments to notice the same features of reality as himself he is going to have to use language in order to specify what his colleague should look at. Rather than just saying "See for yourself" and leading his colleague to Mrs Jones, he's going to have to add a more specific instruction, such as "Look at the way she's walking". In addition, the patient herself is likely to tell the observers about some of her symptoms. However, if we take Nagel's suggestion on board, so long as the terms used in what is said are not amongst those that are laden with the theories under debate no problems should result. The language may not be theory-free but it can be theory-neutral, in that it is neutral between the theories about which the observers disagree.

The possibility of using non-verbal communication to minimise the problems posed by the theory-ladenness of language is particularly relevant in the case of psychiatry. In psychiatry there is a long tradition of showing patients to colleagues and students so that they can see them for themselves. In the past, and to some extent today, students and colleagues are shown patients on ward rounds and at case conferences. Nowadays, video-clips showing students what symptoms look like are shown in lectures. Indeed, in the early 1970s there were some suggestions that the written D.S.M. could be accompanied by a "library of audio-visual definitions tapes which will be a visual definition of the terms used".[27] The idea was that for each symptom or disorder there would be a "piece of behavioural recording on 16mm film or videotape in the American Psychiatric Association 'Bureau of Standards' from which we may judge a patient's degree of 'anxiety', 'manic behaviour' or 'la belle indifference'." Such a library was never actually produced. I am tempted to think that producing such a library now would be unfeasible. The D.S.M. is currently in such wide circulation that videoing a real individual and treating them as a prototype for some psychiatric disorder would be ethically problematic. No-one would want to become known to most of the world's mental health professionals as "Mr Anti-Social Personality Disorder" or "Miss Bulimia". For this reason, one possible means of side-stepping problems caused by theory-ladenness in psychiatric classification must probably be ruled out.

3.1.3 Adopting A Causal Account Of Reference

Some writers have suggested that if a causal account of reference is adopted then there is no reason to believe that observation statements are theory-laden.[28] Causal accounts of reference claim that the meaning of a term depends on the causal history linking uses of the term with an initial baptism.[29] In the case of proper names, the causal theorist claims that the name refers to the person who was originally baptised with the name; so "Rachel Cooper" refers to me because that's the name I was christened with. In the case of natural kind terms ("Great Crested Grebe", "Dandelion", and so on) the term refers to things of the same type as an originally named specimen. A key feature of the causal theory is that a speaker's mistaken beliefs about a person or natural kind have no effect on the term's reference. Even if I think St Nicholas owns a sleigh pulled by flying reindeer, so long as some causal history links my utterances of "St Nicholas" to an original baptism of "Nicholas", my utterances refer to the man St Nicholas. Similarly, a natural kind term continues to refer to things of the type baptised no matter how our theories may change. "Whale" continues to refer to whales despite the discovery that whales aren't fish. On a causal account, the extensions of terms are not affected by changes in the theoretical beliefs of language users. It is this that has led to claims that terms for

[27] Froelich 1972
[28] For example Fodor 1984 pp.27-30
[29] Causal accounts of reference are most closely linked with Kripke 1980 and Putnam 1970.

which a causal account is appropriate will not be theory-laden. In the last chapter I argued that at least some types of mental disorder are plausibly natural kinds. Many philosophers think that some causal account of reference will be appropriate for natural kind terms.[30] If so, and if terms for which a causal account is appropriate are not theory-laden, this would suggest that a theory-free D.S.M. is a real possibility.

However, in his book *Theory and Meaning* (1979) David Papineau argues that terms will be theory-laden even on a causal account of reference. Imagine, he asks us, that a group of travellers have discovered a new country. Their leader points at a native and says "Let's call this one 'Hamlet'". Now, Papineau claims, this on its own is not sufficient to ensure that it's a person that is being named. It needs to be specified what kind of "one" is here being dubbed. Depending on the type of "one" being baptised, the leader may be naming a person, a race, a species, or a skin-colour. Similar problems arise when the causal theorist attempts to name a natural kind or property. A chemist can point at a sample and say "Let's call this 'Flash Silver'", but unless she specifies what kind of thing she is trying to name it remains unclear whether "Flash Silver" is to be the name of a kind of liquid, of a colour, or of an element. In both cases the namer must make their intentions clear before one can know what it is that's being named.

Let's suppose that the namers make it clear that "Flash Silver" is to be the name of a kind of element and that "Hamlet" is to be the name of the race. Still problems can arise. Different scientists can have different ideas about what it takes for two samples to be of the same element, or for two people to be of the same race. As Papineau points out, early 20th century scientists disagreed as to whether the term "lead" applied only to stable matter with atomic weight 206, 207, or 208, or whether it also applied to radioactive substances with atomic weights 210, 211, 212, or 214. Here the argument was over what it took for something to be the same kind of stuff as the samples originally baptised "lead".

Different scientists can have different ideas about what it is for one kind of stuff, or property, or thing, to be of the same kind as another. When different groups have different ideas about such identity principles they will extrapolate from the dubbed original samples in different ways. As a result, theoretical disagreements about the identity principles for types of stuff can lead different groups of scientists to use terms like "lead" to refer to different things.

The reason why it seemed that a causal account of reference might imply that terms need not be theory-laden was that it was hoped that by pointing at a sample and naming it the reference of terms such as "lead" could be fixed independently of any assumptions about lead. The thought was that the original sample plus the concept of "same stuff" might be sufficient to fix the reference. However, because scientists with different theories may have different concepts of "same stuff" this will not do. Even on a causal account, theories are required to specify the identity

[30] The causal account of reference for natural kind terms has traditionally been linked with essentialist accounts of natural kinds. In the last chapter I argue that essentialist accounts of natural kinds should be rejected. However, I think that the account of natural kinds proposed in the last chapter is also compatible with a causal account of natural kind terms.

conditions for the types of thing being named. As such, terms remain theory-laden even on a causal account of reference.

At this point, however, Nagel's suggestion can be employed once again. We can accept that theories are required in order to tell us what the general identity conditions are for various types of thing. Still, so long as scientists in any particular dispute do not disagree about *these* theories, if a causal account of reference is adopted it seems that scientists with different theoretical commitments can mean the same thing by their observation statements. If, for example, two scientists agree about the general identity conditions for biological species they can still disagree about whether badgers spread tuberculosis. When they talk of "badgers" this may assume theoretical beliefs concerning the identity conditions for species, but as these beliefs are not in dispute, the scientists can still communicate.

3.1.4 Summary

To conclude: All observation statements are theory-laden. But, it is still possible for scientists with different theoretical commitments to mean the same thing by an observation statement. As Nagel points out, although observation statements assume theories, often the theories assumed by the observation reports will not be the theories that are in dispute. As such, scientists who disagree about the correct theory of depression may both mean exactly the same thing by the report "Mr Smith has lost weight", so long as they agree about those theories that are assumed by talk of weight-loss.

In some cases the observation statement made by a scientist will assume a theory that is under dispute, for example a biologically-orientated psychiatrist might report that a patient is anxious to a psychoanalyst. Still, it may be possible for the problems caused by theory-ladenness to be sidestepped. The psychiatrist may invite his colleague to come and see the patient for himself. Plausibly, he will have to tell his colleague what to look at. However, the psychiatrist may be able to guide his colleague to direct his attention at the features to be considered without using terms that are under dispute, for example, he may say "Look at the way Mrs Jones keeps fidgeting".

Adopting a causal account of reference may also help to side step the problems caused by theory-ladenness. As we have seen, terms are theory-laden even on a causal account of reference. Theories concerning the general identity conditions for different types of stuff are required in order to get naming off the ground. In general, however, these theories will not be in dispute. Again, often the theories that are assumed by the observation statements are not those at issue, and so scientists with different theoretical commitments will be able to communicate with each other. In general the theory-ladenness of language does not prevent scientists with different theoretical commitments from communicating, as communication does not need to be theory free but just neutral between the theories under debate.

At this point it can be seen that the D.S.M. committee were wrong to describe their classification system as "atheoretical". Still, it may yet be possible for it to be neutral between the various competing accounts of mental disorder. If the D.S.M. can be neutral between different theories in psychiatry this will be a reassuring conclusion. The possibility that observation statements in psychiatry might be laden with psychiatric theory was especially worrying because there is reason to suspect that much psychiatric theory is wrong (if only because there are so many competing theories and they cannot all be right). On the other hand, if observation statements in psychiatry are laden with theories from some other and better established domain, for example biology, this is not so problematic, as we will have more reason to expect these theories to be correct.

Still, we have yet to consider the third way in which observation may be theory-laden, and the effect this could have on the D.S.M. In the next section I examine whether a theory is needed to determine which features of mental disorders are important enough to be worth studying, and consider whether theory-ladenness of this type might cause problems for the D.S.M.

4.PROBLEMS WITH DECIDING WHAT TO OBSERVE

Now we must turn to consider the third way in which observation might be theory-laden. Commonly it is claimed that phenomena are too numerous and too rich for a scientist to be able to set about observing everything. Rather, before being able to start collecting observations, the scientist must decide which features of the world are worth looking at. It is claimed this choice will invariably be informed by the theoretical beliefs of the scientist. For example, a psychoanalyst will choose to spend his time collecting data regarding the childhood experiences and fantasies of his patients, while his biologically-orientated colleague will spend time looking at brain scans and taking blood measurements. As scientists with different theoretical beliefs spend their time looking in different places it is only to be expected that they will end up seeing different things. Even team-based research can't avoid this problem. Team-work enables a group of scientists to examine more than would be possible for an individual researcher. Still, it is claimed, only a minute proportion of what could be looked at can ever be examined. Those who are attracted to this line of argument will claim that classification systems must draw on some theory or another, as a theory must be used to decide which features of the entities under study are of scientific interest.

Although philosophers generally accept that scientists require a theory to help them decide what to look at, there is a tradition in taxonomy that denies that this is necessarily the case. Proponents of numerical taxonomy sometimes deny that scientists must be selective with regard to the features of entities they consider in constructing classification systems.[31] Here I will examine the use of numerical taxonomy within psychiatry and consider whether the techniques employed allow classification systems to be constructed without a theory being needed to select the features of mental disorders that are to be considered important.

[31] For example Sneath and Sokal 1973 p.11

4.1 Cluster Analysis: The Techniques And Their Problems

Numerical taxonomists employ various statistical techniques to construct classification systems. Some of these techniques lend themselves to the construction of categorical classification systems (that is classification systems with discrete categories, such as the D.S.M.); other techniques are best used to construct dimensional classification systems (that is classification systems with dimensions, such as classifications of personality types that have dimensions of extroversion-introversion etc.).

"Cluster analytic" techniques are those most often used in the creation of categorical classification systems. These methods group entities into classes on the basis of their average similarity to each other. The intuitive idea of cluster analysis can be grasped by imagining data on many variables of the entities being analysed plotted in multi-dimensional space. Similar entities end up being close together on the plot, and the distance between any two entities is a measure of their average similarity. A classification can then be extracted by searching for "clusters", groups of highly similar entities.

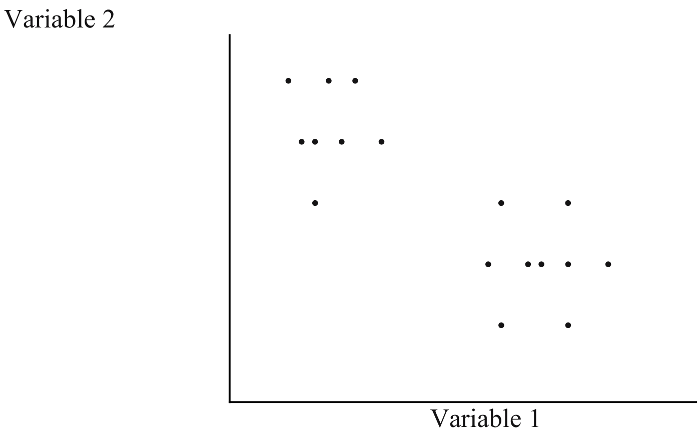

Figure 1. Two dimensions of a cluster solution showing two clusters

It is difficult for cluster analysts to achieve robust cluster solutions because the solution achieved is sensitive to decisions made at several stages in the clustering process. If different variables are measured, or a different sample of entities is studied, different cluster solutions may be obtained. Obviously, if the variables needed to characterise a cluster are not analysed that cluster cannot be found. For example, studies of depression employing purely cross-sectional variables are incapable of finding clusters of bipolar and unipolar depression, as these would be distinguished by differences in the history of the illness.[32] Similarly, excluding a sub-population from the sample will limit the clusters that can be obtained. For example, many studies of alcoholism have used male inpatients from U.S. Veterans Administration hospitals. As most of the subjects were middle-aged Vietnam veterans, there was no possibility of finding clusters of characteristically female or young alcoholics.[33]

Most clustering methods begin by calculating a matrix of similarities between entities. There are many different measures of similarity (Roger Blashfield writing in 1984 counts over fifty[34]), and the cluster solution eventually achieved may be sensitive to the measure used. The results of a cluster analysis also vary with the clustering method employed. There is as yet no consensus concerning the mathematical definition of “cluster” and many techniques (Blashfield counts 150[35]) have been developed. For example, some methods add an entity to a cluster on the basis of it being highly similar to a single member of the cluster, some methods require that an added entity be highly similar to all entities in the cluster, and some methods seek out regions of the multidimensional plot which have a high density of points. Especially when the data analysed lacks a well-defined structure the different methods may produce different solutions.

Different clustering solutions can also be produced using the same method. Regardless of the structure of the data, clustering methods either start with all the entities considered separately and finish with all the entities in one cluster, or vice versa. It is left to the analyst to decide which, if any, of the hierarchy of clustering solutions is acceptable. Various mathematical measures and clinical judgement are usually used in deciding which is the best solution produced.

The problems caused by the sensitivity of cluster solutions and ways of getting round these problems have long been recognised in the psychiatric literature.[36] It is generally accepted that the difficulties can be minimised if investigators reproduce their solutions using a second sample (or sub-sample of the original sample) and using a second clustering method. Analysts may also repeat the analysis using a second set of variables or a sub-set of the original variables. In addition, it is accepted good practice to attempt to “externally validate” clusters, that is to show that they are predictive of significant differences in variables that were not employed in the analysis. For example, clusters of types of depression constructed from an analysis of symptoms might be externally validated by showing correlations with responses to drug treatment.

4.1.1 Cluster Analysis Within Psychiatry

The rhetoric employed by the D.S.M.-III committee and by the early proponents of cluster analysis is strikingly similar. Biologists first developed modern cluster

[32] As noted by Blashfield and Morey 1979
[33] As noted by Skinner 1982
[34] Blashfield 1984 p.230
[35] Blashfield 1984 p.217
[36] For example Everitt 1972, Everitt 1975, Everitt et al. 1971, Strauss et al. 1973, Blashfield 1980

analytic techniques in the late 1950s.[37] The "Numerical Taxonomists" rejected attempts to construct classification systems based on evolutionary descendence on the grounds that in most cases it is impossible to determine the ancestral relationships between taxa. Instead they tried to develop means of classifying organisms based solely on the degree of similarity of their present characteristics. The numerical taxonomy movement in biology made much of the supposed "objectivity", "empiricism", and "naturalness" of the classes produced. Similarly, the D.S.M.-III committee called for a rejection of theory-based classification on the grounds of the paucity of theoretical knowledge. Like the Numerical Taxonomists, they also aimed at a classification system constructed on empirical, atheoretical grounds.

The apparent affinity between numerical taxonomy and the D.S.M. approach has been noted in the psychiatric literature. William Corning and Richard Steffy (1979) note that cluster analysis was developed by biologists in an attempt to overcome precisely the problems caused by a lack of theoretical knowledge that were perceived to be plaguing psychiatric classification.[38] Meehl (1986) considers the idea that a D.S.M. committee truly committed to the creation of a purely descriptive taxonomy should "have proceeded by applying an appropriate formal cluster algorithm to a huge batch of carefully gathered clinical data, 'letting the statistics do the whole job for them'".[39] The idea that cluster analysis might be suitable for constructing atheoretical classification systems would thus almost certainly be familiar to the committees developing the D.S.M.

Cluster analytic studies of mental illness only start appearing in any numbers in the late 1960s when computers became increasingly available. Thus when the D.S.M.-I was published in 1952 very few cluster analytic studies existed. However, cluster analytic studies did have some influence on the D.S.M.-II, published in 1968. In particular, the D.S.M.-II section for childhood disorders was heavily based on a 1966 cluster analytic study by Richard Jenkins. Categories for "hyperkinetic reaction", "withdrawing reaction", "overanxious reaction", "unsocialised aggressive reaction", and "group delinquent reaction" correlate with clusters in Jenkins' study. Only the category "runaway reaction", for children who run away from home, was added.[40] Methodologically, Jenkins' study is not impressive. Jenkins makes no attempt to check his cluster solution by using a second method, or a second sample, or by seeing whether the clusters correlate with variables not used in the analysis. Jenkins did have a seat on the D.S.M.-II committee, however - which doubtless goes some way to explaining the influence of his study.[41]

I have found no evidence that the D.S.M.-III committees were influenced by cluster analytic studies. Indeed by the early eighties some of the one-time proponents of cluster analysis had become disappointed that cluster analysis had not proved as influential as they had hoped. Writing in 1982, Harvey Skinner and Roger

[37] Sokal and Sneath 1963 is the key early text outlining the use of cluster analysis in biology.
[38] Corning and Steffy 1979 p.296
[39] Meehl 1986 p.224
[40] A.P.A. 1968 pp.50-1
[41] A.P.A. 1968 unnumbered first page

Blashfield attribute the failure of cluster analysis to methodological worries and a lack of "salesmanship" on the part of authors.[42]

Blashfield tries to explain the perceived failure of quantitative approaches in psychiatric nosology more fully in *The Classification of Psychopathology* (1984). Here he considers an understanding of inter-professional rivalry between psychiatrists and psychologists to be key to understanding the relations between the D.S.M. and cluster analysis. According to Blashfield the orientation of the D.S.M.-III was heavily influenced by a small "invisible college" of psychiatrists he terms the "Neo-Kraepelinians". Characteristically the Neo-Kraepelinians were keen to operationalise psychiatric diagnostic criteria, were deeply concerned with the reliability of psychiatric diagnosis, and strictly adhered to a medical model of mental disorder (that is they believed that mental diseases are fundamentally diseases like any other, and that psychiatry is a branch of medicine). When Robert Spitzer, a key member of this invisible college, became chair of the D.S.M.-III committee he was allowed to appoint its other members. For the most part he filled the positions with people who shared his theoretical outlook, and thus the Neo-Kraepelinians gained control of the D.S.M.

Medical models of mental illness, such as that adopted by the Neo-Kraepelinians, are sometimes considered a threat to the status of non-medical mental health workers. The fear is that if mental disorders are medical problems then this may imply that only medical doctors should treat them. According to Blashfield, and as previously discussed in Chapter One, the association of the D.S.M. with the medical model resulted in the D.S.M. being perceived as a threat by some psychologists.[43] Part of the American Psychological Association's response to this perceived threat was to try to develop its own, alternative classification scheme through the use of statistical techniques.[44] As a consequence, claims Blashfield, cluster analysis came to be thought of as a psychological method by psychiatrists, and thus came to be ignored by them. On Blashfield's account, the cluster analytic movement and the D.S.M. are opposed to each other. As the D.S.M. approach became ever more dominant, cluster analysis became less and less significant.

I have found some circumstantial evidence that suggests that Blashfield overstates the role of professional rivalries in determining the influence of cluster analytic studies on the D.S.M. In the archives of the A.P.A. and of the American Psychological Association I found correspondence relating to a conference on Classification and Nosology held in 1965.[45] Much of this conference was devoted to discussion of the use of numerical techniques in psychiatric nosology. The National Institute of Mental Health asked both associations if they would be interested in sponsoring the conference. The American Psychological Association was not interested, and so the conference was instead organised by the A.P.A. Thus in the 1960s the A.P.A. actively supported work in cluster analysis. I also went to speak to E.S. Paykel, who published some of the best-known cluster analytic studies on

[42] Skinner and Blashfield 1982
[43] Blashfield 1984 pp.73-4
[44] Blashfield 1984 p.74, p.139
[45] A.H.B. 1963, A.P.A. 1964

depression in the 1970s.[46] Paykel told me that his impression was that psychiatrists and psychologists often worked together on cluster analytic studies and that he thought it unlikely that psychiatrists would ignore studies because they used a "psychological" method. This implies that in the 1970s any tensions between psychiatrists and psychologists in the U.S. were not so great as to be evident to a British psychiatrist interested in cluster analysis. Finally, the D.S.M.-IV is accompanied by a *Sourcebook* that lists all the literature reviewed by the various Workgroups.[47] All cluster analytic studies cited are listed in Table 1.[48] Although the number of cluster analytic studies cited is quite small, several of them are by psychologists. Blashfield claimed that the A.P.A. rejected cluster analysis when the D.S.M.-III was being developed, that is from 1973 to 1980, because it was perceived as being a "psychological" method. Given that the A.P.A. actively supported cluster analysis in the 1960s, and that cluster analytic studies by psychologists were used in the development of the D.S.M.-IV, and that any tensions between American psychologists and psychiatrists during the 1970s were too slight to be noticed by a British cluster analyst, this seems unlikely.

This being said, it does seem to be the case that cluster analytic studies had little influence on the D.S.M.-III and, if Blashfield's explanation for this is to be rejected, it would be nice to be able to provide some alternative explanation. Those responsible for the D.S.M.-III have never publicly discussed cluster analysis. However, one plausible reason why more use was not made of cluster analytic studies is that the cluster analytic movement in psychiatry has constantly been dogged by allegations that the studies are methodologically poor. The clusters found in a cluster analysis need to be validated if the results are to be useful, but many published studies do not attempt to do this. In a review of studies on alcoholism, for example, only two of twenty-five studies were considered well validated by the reviewers.[49] In many cases, even those studies that are cited in the D.S.M.-IV *Sourcebook* are methodologically poor. Most fail to use a second clustering method or sample, and in some cases no attempt has been made to validate the cluster solution obtained (see Table 1 for details). I suggest that Blashfield is probably wrong in thinking that inter-professional rivalries explain the poor reputation of cluster analysis in psychiatry. Rather, a sufficient explanation is plausibly that many studies were methodologically poor, and that this was seen to be the case.

Whatever the true reason, it is clear that cluster analytic studies in psychiatry have had little influence on the D.S.M. Still, the question of whether it would be possible to use methodologically sound cluster analytic studies to construct an atheoretical classification of mental disorders remains. This is the question that must now be addressed.

[46] Paykel 1971, 1972
[47] Widiger et al 1994, 1996, 1997.
[48] The *Sourcebook* runs to thousands of pages. I have read it all once, but once only, thus it is possible that there are some cited cluster analytic studies that I have missed. Even if this is the case, however, there is no reason to think that the studies I found would be disproportionately likely to have been by psychologists.
[49] Morey and Blashfield 1981

Table 1. Cluster Analytic Studies Referred to in the D.S.M.-IV Sourcebooks (Explanatory notes follow)

Disorder and study (1st author and year)	Checks on cluster solution			Author's professional affiliation
	2nd method	2nd sample	External validation	
Insomnia				
Hauri 1983	No	No	No	Clinical psychologist
Dysthymia				
Paykel 1971	No	No	Strong	Psychiatrist
Melancholia				
Paykel 1971	No	No	Strong	Psychiatrist
Overall 1966	No	Yes	Strong	?
Post-Partum psychosis				
Hays 1978	Yes-4	Yes	Strong	Prof. of psychiatry
Specific (simple) phobia				
Curtis1990(unpublished)	No	No	Weak	Psychiatrist
Social Phobia				
Pilkonis 1977	No	No	Weak	Psychology PhD, now in psychiatry dept.
Fremouw et al 1982 Gross and Fremouw 1982	No	Yes & 2 sets variables	Weak	Psychology dept.
Turner and Beidel 1985	No	No	No	Psychiatric Institute
Mixed anxiety & dep.				
Blazer 1988(GOM)	N/A	Yes	Weak	Prof. of psychiatry
Davidson 1988 (GOM)	N/A	No	Strong	Psychiatrist and mathematician
Somato-form pain disorder				
Costello 1987	No	Yes	Strong	Psychologists in a dept. of psychiatry
Female Orgasmic Disorders				
Derogatis 1989	No	No	No	Psychologists and psychiatrists
PDDNOS				
Prior 1975	No	Yes	No	Psychologist and psychiatrist
Dahl 1986	No	No	No	Psychology dept.
Rescorla 1988	No	No	No	Psychologist
Siegel 1986	No	No	Weak	Psychologist
Attention-Deficit without hyper-activity				
Lahey 1988	No	No	No	Psychologists and psychiatrist
Hart (1991 unpublished)	No details given.			
Physical Abuse and Neglect of Children				
James and Boake 1988	No	No	No	Psychologists
Oldershaw et al 1989	Yes	No	Strong	Psychiatry dept.

Notes to table: External validation of cluster solutions is classed as "strong" if the clusters were found to correlate with variables unconnected with those used in the study, for example if clusters based on patients' symptoms are found to correlate with their response to drug treatment, and as "weak" when the clusters are found to correlate with variables connected to those used in the study, for example if clusters based on patients' ratings of symptoms are found to correlate with ratings of symptoms made by observers.

The studies by Dan Blazer and Jonathan Davidson use a technique called "Grade of Membership" analysis and are marked (G.O.M.) in the table. Grade of Membership analysis is a technique developed by Max Woodbury, a biomathematician at Duke University, in the early 1980s.[50] It differs from traditional cluster analysis in that instead of patients being exclusively assigned to one category they can be assigned a quantitative grade of membership for all categories. This means that the technique can cope with the possibility that one person may have more than one disorder. Despite this difference, Grade of Membership analysis is similar enough to cluster analysis for such studies to be considered here.

4.1.2 Could Cluster Analysis Be Used To Construct An Atheoretical Classification System?

As explained earlier, in cluster analysis data is collected on many variables of the entities being analysed. This data is then used to calculate how similar the various entities are to each other. The measurements of the average similarities between entities employed in cluster analysis can be thought of in two distinct ways:

1. As measurements of average similarity with respect to the variables actually utilised in the study.

2. As estimates of the overall true similarity between entities.

Cluster analysts who conceptualise "average similarity" in the first way commit themselves to nothing that is philosophically contentious. For them, cluster analysis is merely a tool that enables many variables to be taken into account when entities are classified. All cluster analysis allows them to do is to consider more variables simultaneously than they would be able to by other means. For these analysts there is no reason to expect that the classifications they obtain are unique. They will accept that someone conducting an analysis on the same entities but using different variables might well get different results. [51] On such a view, cluster analysis cannot be considered a means of producing atheoretical classification systems. Cluster analysis allows many variables to be considered, but these variables will still just be a sub-set of those that might have been analysed. As such, a theory is needed to select those variables that are scientifically interesting.

On the other hand, cluster analysts who conceptualise "average similarity" as being an estimate of the overall true similarity between entities commit themselves to particular metaphysical claims. "Overall true similarity" can be defined roughly as being the proportion of properties that two entities share (slight complications arise when two entities share no property but are somewhat similar because they

[50] See Woodbury and Manton 1982 for further details.

[51] View held by Jardine 1969 – p.211 asserts that there exists no optimal classification, p.216 condemns the idea "...that as increasing numbers of attributes are selected so the 'true' underlying dissimilarities between populations are estimated with increasing accuracy."

each possess a similar property, for example, one weighs 48.9kg while the other weighs 48.8kg). According to Robert Sokal and Peter Sneath (1963) the measures of average similarity provide an estimate of the true overall similarities because "we are taking a random sample from a very large number of characters which we could in theory sample and which would yield us a single, definite proportion of matches if we were able to sample all the characters."[52]

As has often been noted, "overall true similarity" is only a meaningful concept if the number of properties possessed by any one entity is finite.[53] Some writers have claimed that entities "obviously" possess infinitely many properties and have used this point to argue that the project of the analyst who is concerned with "overall true similarities" is incoherent.[54] I suggest that it is far from obvious that entities do possess an infinite number of properties. The plausibility of this claim depends on the metaphysical account of properties that one adopts, and there are many different accounts of properties currently on the philosophical market.[55] The claim that entities have a finite number of properties will seem plausible to those who adopt some realist account of properties – if properties are objective features of the world then their numbers will be fixed. The claim that entities possess an infinite number of properties is more likely to appeal to nominalists – if properties are in some sense generated by the human mind then their supply will plausibly be unlimited.

Debates over the nature of properties are too complex to enter into here, so for the sake of argument let's suppose that entities do have a finite number of properties. Could the cluster analyst who uses a large enough sample of variables then succeed in obtaining an atheoretical classification system? I suggest not. The measures of average similarity calculated by a cluster analyst are based on the *variables* analysed, but the true overall average similarity would depend on the proportion of *properties* that two entities share. The cluster analyst who aims to gain estimates of true overall similarities is committed to a realist account of properties. Only on such an account will there be some "overall similarity" between entities that can be measured. On a realist account, however, it is quite possible for us to miss the mark and choose variables that do not correspond to true properties. Thus, almost certainly "distance from my desk", or more seriously, "being a schizophrenogenic mother", are variables that fail to measure genuine properties.[56] Only if the cluster analyst's variables measure genuine properties will the measures of similarity be estimates of the overall average similarity.

If we want to pick variables that measure true properties our best bet is to rely on our best scientific theories. Those predicates found necessary in our best scientific theories provide our guide as to what properties actually exist. The variables employed in a cluster analysis should aim to measure these properties. As such, if the measures of average similarity are thought of as estimates of the true overall

[52] Sokal and Sneath 1963 p.114, Lorr 1982 takes a similar view and states that cluster analytic techniques "make possible the discovery of natural groupings" p. 461
[53] Sokal and Sneath 1963 pp.91-92, Ehrlich 1964 p.117, Johnson 1968 p.18
[54] Ehrlich 1964 p.117, Johnson 1968 p.18
[55] See Mellor and Oliver 1997 for an overview.
[56] Schizophrenogenic mothers were thought to induce schizophrenia in their children

average similarity, then a theory is needed to guide the choice of variables. As a consequence, these theorists are simply wrong if they claim that cluster analysis can be used to produce atheoretical classification systems.

This point can be made clearer by considering what happens if a variable that does not correspond to a property is included in an analysis. Let's call such a variable a "junk variable". Suppose we have a cluster analysis in which the variables are measures of various psychopathological symptoms - measures of feelings of worthlessness, hallucinations, speed of speech, and so on. Amongst other clusters such cluster analyses of general psychopathology usually produce a cluster that corresponds roughly to psychotic depression. The patients in this cluster feel worthless, they have thought about killing themselves, and their sleep-patterns are disturbed. Now let's consider what would happen if into this analysis we added a junk variable. The junk variable we will add codes for the day of the week on which the patient was born. Adding this variable will tend to split the depressive cluster into seven (whether it actually will split into seven as the result of this one variable depends on the clustering algorithm employed). These clusters will correspond to Monday-born depression, Tuesday-born depression, Wednesday-born depression, and so on. However, we don't believe that Monday-born depression is different in any interesting respect from Tuesday-born depression. Adding junk variables tends to produce cluster solutions that we do not believe corresponds to the natural structure of mental disorders.

Similar problems arise if variables are used that measure properties that are genuine properties but of no theoretical interest in the context. Suppose, for example, that the sex of the patient makes no difference to the nature of the mental disorder that afflicts them. Sex may be a genuine property, but if used in a cluster analysis it will tend to split groups into female and male schizophrenics. In such a case once again the effect of including the variable is to produce a cluster solution that we do not believe reflects the true structure of mental disorders.

As mentioned earlier, it is accepted good practice to repeat a cluster analysis using a second set of variables, or a sub-set of the original set. If this is done, then problems caused by the inclusion of the odd junk or irrelevant variable can be avoided. However, if a significant proportion of the variables are poorly chosen then even this practice won't protect against the generation of worthless cluster solutions. We can conclude that junk variables and irrelevant variables should not be included in cluster analyses, and the only way to try and avoid them is to use our best theories to guide the selection of variables. Thus, regardless of the way in which the analyst conceptualises the measures of average similarity a theory is needed to guide the selection of variables.[57]

Might it be possible to use Nagel's suggestion again at this point? Although the selection of variables requires some theory, might it be possible to use a theory that is not amongst those about which different mental health professionals disagree? I suggest not. Deciding whether a property is relevant requires a theory of the domain in question. For example, in a cluster analysis of psychopathology, biologically-

[57] Fleck 1979, p.92 also notes that the choice of characteristics to be analysed "depends upon the habits of thought of the given scientific discipline; that is, it already contains directional assumptions."

orientated psychiatrists will want to include biological variables but may well consider variables linked to "defence styles" to be irrelevant. Psychiatrists adhering to different theoretical frameworks will disagree.

We can conclude that, contrary to the claims of some of its proponents, cluster analysis is not a technique that can be used to construct atheoretical classification systems. Here we have only looked at cluster analysis because this is the numerical technique that can be used to produce categorical classification systems such as the D.S.M. However, it is worth noting that the discussion here is sufficient to show that no other numerical techniques are capable of producing atheoretical classification systems either. In all cases, a theory is going to be required to inform the selection of the variables that will be subjected to analysis. I conclude that classification systems must always draw on some theory or other, as a theory must be used to decide which features of the entities under study are of scientific interest. Classification cannot be theory-free. Furthermore, as the theory used must be a theory of mental disorders, classification cannot even be theory-neutral in Nagel's sense.

5.IMPLICATIONS FOR THE D.S.M.

All three of the ways in which observation might be theory-laden have now been examined. In the first section I concluded that there is insufficient evidence to determine whether or not perception in psychiatry is theory-laden (it will be remembered that by "perception" I mean what an organism sees, hears, or otherwise senses).

In the second section it was concluded that all observation statements are theory-laden but that the implications of this are limited. As Nagel points out, although observation statements assume theories, often the theories assumed by the observation reports will not be the theories that are in dispute. The scope of Nagel's observation can be maximised if a causal account of reference is appropriate, and through employing non-linguistic forms of communication. I conclude that while observation statements cannot be theory-free they can often be neutral between the theories that are in dispute. As theory-ladenness will only lead us astray when our theories are wrong this thought is comforting. Assuming that science has indeed made some progress, theories about which there is consensus are more likely to be right than those that are in dispute. Thus, if we can use language that assumes only widely accepted theories the epistemic problems caused by the theory-ladenness of language can be minimised.

In the final section I argued that a theory is always required for classification because a theory is required to guide a scientist in deciding which features of the world are relevant. Despite the claims of some numerical taxonomists this problem renders even so called "empirical" techniques, such as cluster analysis, theory-laden. To know whether or not some property is relevant one needs a theory of the domain in question, and so classification cannot even be theory-neutral in Nagel's sense.

If the D.S.M. cannot be theory-free what theory does it use? I suggest, as have many writers before me, that the D.S.M. tacitly assumes that some biological account of mental illness will prove to be correct. The *Sourcebook* published alongside the D.S.M.-IV reveals that the studies appealed to by the D.S.M. committees are mainly biological in orientation. These studies examine, for example, the biological correlates of disorder, they assess whether drug treatments differentially affect different groups of patients, they look at whether disorders run in families, and at whether particular disorders tend to affect people of a particular age and sex.

As mentioned at the beginning of this chapter if the D.S.M. is theory-laden this may give us reason to doubt that the categories it contains will reflect natural kinds of disorder. Biological accounts of mental illness are by no mean uncontroversial. As the D.S.M. tacitly assumes some biological explanation for mental disorder the D.S.M. categories stand, or quite possibly fall, with such an account.