

Full Title: **Psychiatric diagnosis, tacit knowledge and criteria**

Short title: **Psychiatric diagnosis, tacit knowledge and criteria**

Authors (i.e. name as it should appear on the paper):

**Tim Thornton**

Position (i.e. job description as it should appear on the paper):

**Tim Thornton**  
**Professor of Philosophy and Mental Health**

Address for correspondence:

**School of Health**  
**University of Central Lancashire**  
**Preston**  
**PR1 2HE**

Other Contact Details:

phone **01772 894646**

fax

email:

**TThornton1@uclan.ac.uk**

Key-words (not appearing in title):

**Explicit knowledge, judgement, Gestalt judgement, Polanyi, McDowell.**

Word count (excluding title page but including references):

**~8,700**

Character count (abstract):

Software used: Microsoft Word

Filename(s):

## Psychiatric diagnosis, tacit knowledge and criteria

### Abstract

The two main psychiatric taxonomies set out codifications of psychiatric diagnoses via lists of symptoms with the aim of maximising the reliability of diagnostic judgements. This approach has been criticised, however, for failing to capture the precise connection between diagnostic judgements and symptoms as detected by skilled clinicians. Assuming that this criticism is correct, this chapter offers related two accounts of why this might be so. First, skilled diagnostic judgement may be an exercise of tacit knowledge: a practical skill the exercise of which requires the presence of the patient or client. Second, the conception of criteria implicit in DSM and ICD is based on a mistaken view of how what people say and do connects to their mental states. On an alternative account, in an overall gestalt diagnostic judgement the various criteria are abstractions from a whole which directly expresses the underlying psychopathological state of patients or clients.

### Introduction

For the last 50 years, both of the major psychiatric diagnostic systems – DSM and ICD – have aimed at reliability at the potential cost of validity. They have done this by codifying diagnosis in the form of criteria, influenced by operationalism from the philosophy of physics and down playing aetiological theory. It is an empirical question whether DSM-III, -IV and now -5 and the parallel ICD classifications have achieved this aim overall.

There have been criticisms, however, that the explicit criteria under-determine the diagnoses made by skilled clinicians. That is, the criteria themselves have a vagueness or indeterminacy for which experienced psychiatrists have to compensate in diagnostic judgements in response to particular patients expressing particular signs and symptoms. The overall top-down or gestalt judgement is more precise than the component criteria on which it is supposed to be based.

The aim of this chapter is not to address whether this is so but rather how it could be so. In doing so, I will make two suggestions. First, diagnosis may involve an important tacit element. As a recognitional judgement, it may share characteristics of an uncodifiable form of know-how. Second, the postulation of criteriological intermediaries between the skilled clinician and their patients' or clients' actual conditions may distort the recognitional process. Judgement of the underlying mental states of patients and clients may be more secure than the operationalised criteria.

The first section outlines the reasons for the emphasis, since the second half of the twentieth century, on operationalism in both the main psychiatric taxonomies. The second section sets out three similar clinically based criticisms of the resulting criteriological model of diagnosis.

The final two sections set out to shed light on how this might possible first by suggesting that diagnostic judgements are an instance of a broader class of tacit knowledge and then by suggesting that the criteriological view distorts the way that clinical signs and symptoms bespeak, to a skilled clinical, underlying pathologies directly.

### Background: the rise of criteriological diagnosis

Over the last half century, there has been a concerted effort to improve the reliability of psychiatric diagnosis by pruning the two main diagnostic systems of possibly over hasty aetiological theory and stressing instead more directly observational features of presenting subjects. Two main factors explain this. (For a fuller account, see [Fulford et al 2005].)

First, on its foundation in 1945, the World Health Organisation set about establishing an International Classification of Diseases (ICD). The chapters of the classification dealing with physical illnesses were well received but the psychiatric section was not widely adopted. The British psychiatrist Erwin Stengel was asked to propose a basis for a more acceptable classification. Stengel chaired a session at an American Psychological Association conference of 1959 at which the philosopher Carl Hempel spoke. As a result of Hempel's paper (and an intervention by the

psychiatrist Sir Aubrey Lewis) Stengel proposed that attempts at a classification based on theories of the causes of mental disorder should be given up (because such theories were premature), and suggested that it should instead rely on what could be directly observed, that is, symptoms.

In fact, Hempel's paper provided only *partial* support for the moral that was actually drawn for psychiatry. He argued that:

Broadly speaking, the vocabulary of science has two basic functions: first, to permit an adequate description of the things and events that are the objects of scientific investigation; second, to permit the establishment of general laws or theories by means of which particular events may be explained and predicted and thus scientifically understood; for to understand a phenomenon scientifically is to show that it occurs in accordance with general laws or theoretical principles. [Hempel 1994: 317]

These two requirements – that terms employed in classifications should have clear, public criteria of application and should lend themselves to the formulation of general laws – correspond to the aims of *reliability* and *validity* respectively. Clear public criteria promote both test-retest and inter-rater reliability whilst general laws are a step, at least, towards construct validity. But it was the former that was adopted by psychiatry as the key aim at the time. With respect to it, Hempel claims that

Science aims at knowledge that is *objective* in the sense of being intersubjectively certifiable, independently of individual opinion or preference, on the basis of data obtainable by suitable experiments or observations. This requires that the terms used in formulating scientific statements have clearly specified meanings and be understood in the same sense by all those who use them. [ibid: 318]

He commends the use of operational definitions (following Bridgman's book *The Logic of Modern Physics* [Bridgman 1927]), although he emphasises that in psychiatry the kind of measurement operations in terms of which concepts would be defined would have to be construed loosely. This view has been influential up to the present WHO psychiatric taxonomy in ICD-10.

The second reason for the emphasis on reliability and hence operationalism was a parallel influence within American psychiatry on drafting DSM-III. Whilst DSM-I and DSM-II had drawn heavily on psychoanalytic theoretical terms, the committee charged with drawing up DSM-III drew on the work of a group of psychiatrists from Washington University of St Louis. Responding in part to research that had revealed significant differences in diagnostic practices between different psychiatrists, the 'St Louis group', led by John Feighner, published operationalised criteria for psychiatric diagnosis. The DSM-III task force replaced reference to Freudian aetiological theory with more observational criteria.

This stress on operationalism has had an effect on the way that criteriological diagnosis is codified in DSM and ICD manuals. Syndromes are described and characterised in terms of disjunctions and conjunctions of symptoms. The symptoms are described in ways influenced by operationalism and with as little aetiological theory as possible. (That they are neither strictly operationally defined nor strictly aetiologically theory free is not relevant here.) Thus one can think of such a manual as providing guidance for, or a justification of, a diagnosis of a specific syndrome. Presented with an individual, the diagnosis of a specific syndrome is justified because he or she has enough of the relevant symptoms, which can be, as closely as possible, 'read off' from their presentation. The underlying syndrome is connected to more accessible, epistemologically basic signs and symptoms.

### **An objection to criteriological approaches**

Although the rationale for a criteriological, or bottom up, approach to diagnosis seems clear, it has not escaped criticism. The charge outlined in this section is that combining individual signs and symptoms understood initially in isolation from context and only assembled in the conjunctions that add to diagnosis makes the signs and symptoms imprecise.

In a paper called 'Phenomenological and criteriological diagnosis: different or complementary?' Alfred Kraus, professor of psychiatry at Heidelberg, argues that diagnostic systems such as DSM and ICD miss out an important element of psychiatric diagnosis [Kraus 1994]. Because they assume that diagnoses are built up from a number of individual and conceptually independent symptoms they cannot capture top-down and holistic elements of diagnosis.

One key criticism that Kraus makes of what he calls this criteriological approach to diagnosis, is that rather than providing a reliable foundation, the connection between individual symptoms and conditions lacks *specificity*.

[S]ymptomatological/criteriological diagnosis not only makes the reality of the patient accessible in a very reduced way but also portrays the pathological phenomena in a very imprecise and broad manner... The reduction of phenomena to symptoms and criteria has as its consequence a loss of specificity. [ibid: 153-4]

Taking delusions as an example of a symptom, on the criteriological model, for schizophrenia, Kraus argues that there is not a reliable connection between delusions in general and schizophrenia. 'Delusion' is a vague term picking out a variety of psychological states. The reliable connection is between particular kinds of delusional structure and schizophrenia. But the identification of delusions with a specific schizophrenic colouring presupposes, Kraus argues, a top down holistic model rather than a bottom up description. The assumption on the criteriological approach that symptoms can be recognised and described independently of the psychopathological diagnostic categories of which they are a part introduces vagueness to their descriptions and hence undermines the specificity of their connection to diagnostic judgements.

Kraus also argues that in the bottom-up model, symptoms can only be added together through conjunction. But no mere conjunction of individual symptoms—a 'Chinese restaurant menu' approach—can capture the psychological integrity up to which the individual parts add. For that, one again needs a top down holistic approach. This is not to say, however, that particular elements cannot be identified in a holistic diagnosis. It is just that the individual elements have a different logic.

One way of marking this distinction (although not Kraus' own) is to contrast parts that are independent pieces and parts that are essential aspects. The pieces of a jigsaw add up to a whole, but each piece can exist independently of the others. By contrast a musical note has both a tone and a pitch, but neither aspect can exist independently of the other. Thus, according to a holistic approach, psychological symptoms are not independent building blocks towards diagnostic judgements but are interdependent aspects of a psychological unity.

Kraus combines with these two comments on the limits of a criteriological model of diagnosis with a further philosophical explanation of the difference in approach. This is why he contrasts the criteriological with a phenomenological rather than merely a holistic model. This concentrates not on psychiatric diseases but on the mode of being of whole persons, the 'whole of the being in the world of schizophrenics or manics'. Thus the phenomenologically based diagnosis of schizophrenia turns on an overall assessment of the patient—a 'praecox feeling'—as having a very different form of 'being-in-the-world'. Whether or not that more general view is correct, Kraus' criticism suggests that the operational structure of psychiatric manuals introduces a vagueness into the description of symptoms and hence undermines the specificity of the link between symptom, when properly understood, and diagnostic judgement.

Mario Maj makes a similar criticism. Again taking the example of schizophrenia, he argues that:

One could argue that we have come to a critical point in which it is difficult to discern whether the operational approach is disclosing the intrinsic weakness of the concept of schizophrenia (showing that the schizophrenic syndrome does not have a character and can be defined only by exclusion) or whether the case of schizophrenia is bringing to light the

intrinsic limitations of the operational approach (showing that this approach is unable to convey the clinical flavour of such a complex syndrome). In other terms, there may be, beyond the individual phenomena, a 'psychological whole' (Jaspers, 1963) in schizophrenia, that the operational approach fails to grasp, or such a psychological whole may simply be an illusion, that the operational approach unveils. [Maj 1998: 459-60]

In fact, Maj argues that this shows the weakness of the operational approach. He argues that the DSM criteria fail to account for aspects of a proper grasp of schizophrenia: for example, the intuitive ranking of symptoms (which have equal footing in the DSM account). He suggests that there is, nevertheless, no particular danger in the use of DSM criteria by skilled, expert clinicians for whom it serves merely as a reminder of a more complex prior understanding. But there is a problem in its use to encode the diagnosis for those without such an additional underlying understanding:

If the few words composing the DSM-IV definition will probably evoke, in the mind of expert clinicians, the complex picture that they have learnt to recognise along the years, the same cannot be expected for students and residents. [ibid: 460]

Maj's criticism that the DSM criteria do not capture a proper, expert understanding of the diagnosis of schizophrenia raises the question of how or why that could be the case. If the criticism is right, is it that the wrong criteria have been used: either the wrong symptoms and / or the wrong rules of combination? Or is there something more fundamentally wrong with the criteriological approach as applied to psychiatry?

Josef Parnas suggests the latter. In a paper describing pre-operational approaches to taxonomy and diagnosis as a 'disappearing heritage' he comments on an underlying difference in attitude towards signs and symptoms of schizophrenia.

When the pre-DSM-III psychopathologists emphasized this or that feature as being very characteristic of schizophrenia, they did not use the concept of a symptom/sign as it is being used today in the operational approach. This latter approach envisages the symptoms and signs as being (ideally) third person data, namely as reified (thing-like), mutually independent (atomic) entities, devoid of meaning and therefore appropriate for context-independent definitions and unproblematic assessments. It is as if the symptom/sign and its causal substrate were assumed to exhibit the same descriptive nature: both are spatio-temporally delimited objects, ie, things. In this paradigm, the symptoms and signs have no intrinsic sense or meaning. They are almost entirely referring, ie, pointing to the underlying abnormalities of anatomo-physiological substrate. This scheme of 'symptoms = causal referents' is automatically activated in the mind of a physician confronting a medical somatic illness. Yet the psychiatrist, who confronts his 'psychiatric object', finds himself in a situation without analogue in the somatic medicine. The psychiatrist does not confront a leg, an abdomen, not a thing, but a person, ie, broadly speaking, another embodied consciousness. What the patient manifests is not isolated symptoms/ signs with referring functions but rather certain wholes of mutually implicative, interpenetrating experiences, feelings, beliefs, expressions, and actions, all permeated by biographical detail. [Parnas 2011: 1126]

The claim here is that the criteriological approach has the wrong model of psychiatric symptoms and signs in two respects. Just as smoke can indicate fire or tree rings the age of a tree, the criteriological approach takes signs and symptoms to be free standing items which merely causally indicate underlying states. Furthermore, these relations are independent of one another: they are atomic. By contrast, Parnas suggests, psychiatric signs and symptoms are both essentially meaning-laden rather than brutally causal and also mutually interdependent wholes. It is the latter claim, which plays the more important role in his criticism.

One argument for their interdependence is that it is only in particular contexts that symptoms are reliable. Thus, for example, mumbling speech is comparatively widespread (Parnas estimates 5% of

the population) but in – and only in – the context of other features such as ‘mannerist allure, inappropriate affect, and vagueness of thought, it acquires a psychopathological significance’ [ibid: 1126]. So the effectiveness of the sign is context-dependent. In some contexts it is indicative and in others not. Excluded from context – as it is in the criteriological context – it is vague. But it is precise in context. Parnas goes further by suggesting a more than merely additive view. Grasp of psychiatric symptoms is likened to seeing the figure of the duck-rabbit first as a rabbit and then suddenly as a duck: seeing the signs and symptoms under an overall aspect or gestalt.

A Gestalt is a salient unity or organization of phenomenal aspects. This unity emerges from the relations between component features (part-whole relations) but cannot be reduced to their simple aggregate (whole is more than the sum of its parts)... A Gestalt instantiates a certain generality of type (eg, this patient is typical of a category X), but this typicality is always modified, because it is necessarily embodied in a particular, concrete individual, thus deforming the ideal clarity of type (universal and particular). [ibid: 1126]

So the model of diagnosis is one in which the skilled clinician grasps the right diagnosis as an integrated whole in which different aspects can be seen as abstractions from that whole rather than as its basic building blocks. Such a view would accommodate Kraus’ rejection of a ‘Chinese restaurant menu’ approach and Maj’s suggestion that criteriological elements serve as reminders for already skilled clinicians. They do – on this view – in the sense that, after the fact, such articulations of the overall picture are possible, as a musical note may be divided into its pitch, tone and duration whilst it cannot be built up from those as independent building blocks. But that does not imply that the expert judgement of the whole could be built up from the individual criteria understood in isolation.

There is a further possibility hinted at in the criticism of Kraus, Maj and Parnas. On a criteriological view symptoms are not merely independent of each other (as Kraus points out) they are conceptually independent of the underlying psychopathological state that they indicate. But in the case of Kraus and Parnas, at least, there is a suggestion that the connection between symptoms (when correctly understood) and psychopathological state is more direct: the state is expressed directly in the signs and symptoms to those, at least, with the skill to see it.

### **Diagnosis and tacit knowledge**

The criticisms of the criteriological approach set out in the preceding section prompt two further questions. The bottom up codification of diagnosis through simpler, more basic signs and symptoms suggests an explanation of how complex diagnostic judgement is possible. It is possible because it is based on simpler more epistemically accessible building blocks. The first question concerns the nature of an overall ‘gestalt’ judgement if that explanation is rejected. On what is top-down judgement based and what is its relationship to the criteriological approach? In this section, I will suggest an analogy with context-dependent tacit knowledge to try to make the rejection of the above explanation seem a less puzzling possibility [for a more detailed discussion see Thornton 2013]. But it will also help highlight how the move from context-dependent recognition to explicit criteria introduces vagueness into the description of psychiatric symptoms.

Second, if diagnostic judgement is not based on more observational features of a clinical encounter, how can it yield knowledge of underlying mental states? In the final section, I will suggest an analogy with the more general ‘problem of other minds’ and outline what may initially seem a counter-intuitive view outlined by the philosopher John McDowell which inverts the epistemic priority of judgements about behavioural signs and symptoms and judgements of underlying mental states. Again it will suggest that reliance on basic criteria comes at the cost of introducing vagueness into description of psychiatric symptoms which undermines the potential directness of psychiatric diagnosis as described by Kraus, Maj and Parnas.

I suggested at the start that the development of the theoretically minimal criteriological approach to diagnosis in psychiatry was partly influenced by operationalism in the philosophy of science in the first part of the twentieth century. The aim was to minimise uncodified elements in psychiatric diagnosis so as to maximise reliability. But there was, in the second half of the century, a contrasting view about the nature of scientific knowledge: the chemist turned philosopher Michael Polanyi's arguments for the importance of tacit knowledge. (Polanyi himself talks of tacit *knowing* rather than knowledge. I will, nevertheless, use 'knowledge' whilst talking about his views but will return to emphasise the practical dimension to what is tacit.) Top-down or gestalt judgement in psychiatry can be thought of as an instance of tacit knowledge. I will use Polanyi to introduce this notion but will deviate from his account shortly.

Polanyi gives the following example:

We know a person's face, and can recognize it among a thousand, indeed among a million. Yet we usually cannot tell how we recognize a face we know. So most of this knowledge cannot be put into words. [Polanyi 1967b: 4]

This is an instance of what he takes to be a general phenomenon. Indeed, he begins his book *The Tacit Dimension* with the following bold claim:

I shall reconsider human knowledge by starting from the fact that *we can know more than we can tell*. [Polanyi 1967b: 4]

The broad suggestion is that knowledge can be tacit when it is, on some understanding, 'untellable'. 'Tellable' knowledge is a subset of all knowledge and excludes tacit knowledge. But the slogan is gnomic. Does it carry, for example, a sotto voce qualification 'at any one particular time'? Or does it mean: ever?

The very idea of tacit knowledge presents a challenge: it has to be tacit and it has to be knowledge. But it is not easy to meet both conditions. Emphasising the tacit status, threatens the idea that there is something known. Articulating a knowable content, that which is known by the possessor of tacit knowledge, risks making it explicit. There is a second strand through Polanyi's work which helps address this problem. At the start of his book *Personal Knowledge* in which he says:

I regard knowing as an active comprehension of things known, an action that requires skill. [Polanyi 1958: vii]

These two features suggest a way to understand tacit knowledge: it is not, or perhaps cannot be made, explicit and it is connected to action, the practical knowledge of a skilled agent. The latter connection suggests a way in which tacit knowledge can have a content: as practical knowledge of how to do something. Taking tacit knowledge to be practical suggests one way in which it is untellable. It cannot be made explicit except in context-dependent practical demonstrations. It is not that it is mysteriously ineffable but that it cannot be put into words alone.

Psychiatric diagnostic judgement can be thought of as an example of such a skill: the ability to recognise, in a particular context, the manifestation of psychiatric illness. Polanyi also compares recognition to a practical skill, likening it to bicycle riding:

I may ride a bicycle and say nothing, or pick out my macintosh among twenty others and say nothing. Though I cannot say clearly how I ride a bicycle nor how I recognise my macintosh (for I don't know it clearly), yet this will not prevent me from saying that I know how to ride a bicycle and how to recognise my macintosh. For I know that I know how to do such things, though I know the particulars of what I know only in an instrumental manner and am focally quite ignorant of them. [ibid: 88]

In both cases, the 'knowledge-how' depends on something which is not explicit: the details of the act of bike riding or raincoat recognition. Whilst one can recognise one's own macintosh one is, according to Polanyi, ignorant, in some sense, of how. Thus how one recognises it is tacit. Polanyi

suggests here that explicit recognition of something as an instance of a type is based on the implicit recognition of subsidiary properties of which one is focally ignorant. He explains the distinction of focal and subsidiary awareness using the example of focusing attention on what a pointing finger points to. In looking from the finger to the object, the object is the focus of attention whilst the finger, though seen, is not attended to. It is not invisible, however, and could itself become the object of focal attention.

Polanyi seems to assume that the question of how one recognises something always has an informative answer and then to cover cases where it is not obvious what this is he suggests it can be tacit. But, firstly, whilst it sometimes may have an informative answer, there is no reason to think that it always has (cf recognising that a wall is red). Secondly, even in cases where one recognises a particular as an instance of a general kind in virtue of some further properties *and cannot give an independent account of those properties*, it is not clear that one need be focally ignorant of them. It may be, instead, that the awareness one has of the 'subsidiary' properties is simply manifested in the act of recognition. I might say, I recognise that this is a, or perhaps my, macintosh because of how it looks *here* with the interplay of sleeve, shoulder and colour even if I could not recognise a separated sleeve, shoulder or paint colour sample as of the same type. Whilst it seems plausible that one might not be able to say in context-independent terms just what it is about the sleeve that distinguishes a or my macintosh from any other kind of raincoat (one may, for example, lack the vocabulary of fashion or tailoring) that need not imply that one is focally ignorant of, or not attending to, just those features that make a difference. Recognition may depend on context-dependent or demonstrative elements, such as recognising shapes or colours for which one has no prior name. But if anything, that suggests one has to be focally aware, not focally ignorant, of them.

Thus Polanyi's own account of the tacit nature of recognition faces objections. But such criticism suggests the possibility of a more minimal account of tacit knowledge. Recognition is tacit because it is a skill – for example, developed through repetition and critical practice and demonstrated in applications – and because it can thus be articulated only in context-dependent terms such as 'like this!'. It cannot be explicated in words alone independently of additional practical demonstrations in context.

If the skilled diagnostic judgement described in the previous section by Kraus, Maj and Parnas is thought of as tacit knowledge as just explicated then it can be contrasted with criteriological diagnosis in the following way. The criteria set out in ICD and DSM are an attempt to make psychiatric diagnosis *explicit*, to put it into words alone. They attempt to set out context-independent descriptions of psychiatric syndromes.

Such an attempt is akin to attempting to model an ability to recognise colours and shades on general knowledge of the names for colours that ordinary people have. For most people, the ability to recognise, think about and recall (at least for some period) particular shades of colour goes beyond what they can make explicit linguistically. The ability can instead be manifested by pointing to particular instances of colour themselves. By contrast with the fine discriminations that can be made in the presence of actual colours and shades, colour *vocabulary* is generally vague.

Similarly, by contrast with the context depending discriminations of skilled clinicians made in the presence of their patients and clients, the criteria set out in diagnostic manuals are vague. Because they are fully linguistic, the criteria in DSM and ICD are portable. There is an advantage in communication of a linguistic codification of diagnosis that floats free of particular inter-personal relations. But it is bought at the cost of precision. By contrast, the features that play a role in the top-down diagnoses of skilled clinicians are identified in the presence of a particular patient's or client's psychological whole. Such recognition cannot be captured in words alone.

The analogy suggested in this section has been between clinical judgement made possible by the presence of a patient or client and recognition of a macintosh, either as an instance of a kind or as a particular one, or recognition of a colour or shade in its presence. The analogy suggests that the

patient herself is passive and plays no active role. Since clinical judgement depends a great deal on what patients say and do, the general picture of tacit knowledge needs augmenting with a specific account of the recognition of mental states. That is the subject of the next section.

### Criteria and other minds

In the previous section, I suggested that tacit knowledge can be used to shed light on the idea that an overall top down or gestalt diagnostic judgement could be more specific than a diagnosis based on general but vague criteria. A skilled clinician has a recognitional skill which can only be exemplified in context-dependent judgements in the presence of patients or clients. That is to approach the problem from an epistemological perspective: what it is to have knowledge in this way. In this section, I will complement that by taking an ontological view. What could the relation be between the underlying mental states and conditions amounting to mental illness or disease syndromes and the more apparently epistemically accessible criteria set out in DSM and ICD? Addressing this question will also address the active role of patients and clients raised just now.

To sketch an answer to this question I will consider a debate from the philosophy of mind about whether our knowledge of other minds in general is based on behavioural *criteria*. Although the argument against that view that I will outline does not directly carry over to the case of psychiatric diagnosis, it does suggest why criteriological diagnosis is vague compared to top-down or gestalt judgement.

The concept of a criterion was introduced into the philosophy of mind as a solution to the problem of other minds by followers of the philosopher Ludwig Wittgenstein. The influential Wittgenstein exegete PMS Hacker, writing in the *Oxford Companion to Philosophy*, defines a criterion thus:

A standard by which to judge something; a feature of a thing by which it can be judged to be thus and so. In the writings of the later Wittgenstein it is used as a quasi-technical term. Typically, something counts as a criterion for another thing if it is necessarily good evidence for it. Unlike inductive evidence, criterial support is determined by convention and is partly constitutive of the meaning of the expression for whose application it is a criterion. Unlike entailment, criterial support is characteristically defeasible. Wittgenstein argued that behavioural expressions of the 'inner', e.g. groaning or crying out in pain, are neither inductive evidence for the mental (Cartesianism), nor do they entail the instantiation of the relevant mental term (behaviourism), but are defeasible criteria for its application. [Honderich 1995]

Key features of this definition are that the criteria of, for example, an 'inner' state like pain are fixed by convention and are partly constitutive of what we mean by the word 'pain'. Thus groaning and crying out are not mere symptoms but rather part of what we understand by 'pain', connected by definition not induction. At the same time, however, the criteria of pain are *defeasible*.

The reason for this qualification is the following intuition. Whilst, in general, pain behaviour is the expression of underlying pain, on occasion behaviour which resembles pain behaviour in every detail is not the expression of pain. It may be the result of acting or pretence. (And equally, genuine underlying pain may sometimes be stoically kept from expression.) As a result, the criterial support that apparent pain behaviour gives for a judgement that someone is in pain is taken to be defeasible. It can, on occasion, be overturned.

The idea that criteria give only defeasible support for a claim is combined with a further assumption which the philosopher John McDowell, in his criticism of this very notion, describes thus: 'if a condition is ever a criterion for a claim, then any condition of that type constitutes a criterion for that claim, or one suitably related to it' [McDowell 1982: 462-3]. In other words, criteria are types. Whilst on most occasions, when instances of some general type of criterion are satisfied the underlying fact for which those instances are criteria also obtains, on some occasions the type of

criterion is satisfied (by some particular circumstances) but the fact does not obtain. In such cases, the criterion is satisfied but is nevertheless also defeated.

This suggests that there is an essential underdetermination in the support that criteria, so understood, provide for judgements about mental states. In any particular case, on this picture, some expression, some sign or symptom of pain for example, may or may not actually *mean* that the person expressing it is actually in pain. Hence the behavioural expression is vague. Its meaning is imprecise.

This worry provides the basis for McDowell's criticism of the use of criteria, understood in this way, to explain how knowledge of other minds is possible. On the assumption that it is sometimes, at least, possible to know someone else's mental state, McDowell asks how such knowledge is supposed to be based 'on an experiential intake that falls short of the fact known... in the sense [of]... being compatible with there being no such fact' [McDowell 1982: 459].

The worry is this. If one knows something, then it cannot be the case that - 'for all one knows' - things may be otherwise. That possibility is ruled out precisely because one *knows* what is the case. But if criteria fall short of implying the fact that they are supposed to enable one to know, then they cannot themselves rule out the possibility that the fact does not obtain. So if our everyday concept of knowledge *does* rule this out then such knowledge cannot be based on perception that the criteria for some mental state are satisfied. A possible alternative view in which the perceived the criteria is supposed merely to be enough to satisfy linguistic *conventions* for the *ascription* of knowledge would also not address this objection, either.

If experiencing the satisfaction of 'criteria' does legitimise ('criterially') a claim to know that things are thus and so, it cannot also be legitimate to admit that the position is one in which, for all one knows, things may be otherwise. But the difficulty is to see how the fact that "criteria" are defeasible can be prevented from compelling that admission; in which case we can conclude, by contraposition, that experiencing the satisfaction of 'criteria' cannot legitimize a claim of knowledge. How can appeal to "convention" somehow drive a wedge between accepting that everything that one has is compatible with things not being so, on the one hand, and admitting that one does not know that things are so, on the other? [McDowell 1982: 458]

Imagine that there are two observers who both see that the behavioural criteria, so construed, for two other people being in pain are satisfied but that only one of them really is in pain: the other is pretending. If the observers' experiences are the only grounds for them knowing the mental state of their respective subject and if their perceptions are the same in both cases (seeing that the criteria for pain are met) then how can one observer know their subject's mental state and the other observer not? Surely, neither has *knowledge* even if one has, by chance, a true belief. It seems merely a matter of luck that one observers' experience is of undefeated criteria whilst the other's is of defeated criteria, that in one case the observed subject really is in pain and in the other merely pretending. The luckier observer has done nothing extra to earn the right to knowledge. Construing criteria as defeasible to try to accommodate the fact that we are fallible at knowing other people's minds cannot work because it rules out that we ever have knowledge.

There is, however, an alternative view of criteria and of knowledge of other minds based on them. Rather than assuming that, in the case of pretence, the criteria for mental states are satisfied but are also defeated - by the fact that it is a case of pretence - one can instead construe it as a case of the criteria only *appearing* to be satisfied. This is a rejection of the idea that criteria are *defeasible types* of situation. Instead, McDowell presses the idea that, when criteria are satisfied, one's experience *does not fall short of the facts*. So there cannot be cases where the criteria are satisfied without the fact for which they give criterial support also holding.

McDowell supports this interpretative possibility by considering a passage in which Wittgenstein discusses criteria in a non-mental context.

The fluctuation in grammar between criteria and symptoms makes it look as if there were nothing at all but symptoms. We say, for example: "Experience teaches that there is rain when the barometer falls, but it also teaches that there is rain when we have certain sensations of wet and cold, or such-and-such visual impressions." In defence of this one says that these sense-impressions can deceive us. But here one fails to reflect that the fact that the false appearance is precisely one of rain is founded on a definition. [Wittgenstein 1953 §354]

Wittgenstein rejects the temptation to say that both the fall of a barometer and also sensations of wet and cold (or visual impressions) are mere *symptoms* of rain. Instead, and by contrast with the barometer fall, the connection between the sensations (or the visual impressions) and rain is definitional or criterial. They are used in an explanation of what 'rain' means. This thought can, however, be interpreted in two ways.

Commentators often take this to imply that when our senses deceive us, criteria for rain are satisfied, although no rain is falling. But what the passage says is surely just this: for things, say, to look a certain way to us is, as a matter of 'definition' (or 'convention'...), for it to look to us as though it is raining; it would be a mistake to suppose that the 'sense-impressions' yield the judgement that it is raining merely symptomatically - that arriving at the judgement is mediated by an empirical theory. That is quite compatible with this thought... when our "sense-impressions" deceive us, the fact is not that criteria for rain are satisfied but that they *appear* to be satisfied. [McDowell 1982:466]

Someone who steps outside their house when the lawn sprinklers are switched on may think that by having experiences of wet and cold they have experienced the criteria for rain, albeit on this occasion defeated. After all, when being taught about rain they may have been taught it through practical definitions involving experiences that felt similar. But the experiences used in the practical definition were not just any experiences of wet and cold but wet and cold experiences of rain falling. Similarly in the case of criteria for mental states, pretence can make it *seem* that the criteria for pain, for example, are satisfied when, in fact, they are not.

Taking the criteria to be merely *any* experience of wet and cold (for rain) or *any* experience of high pitched cries (for pain) makes them too vague to sustain knowledge. Correcting this requires rethinking the generality and the descriptive nature of criteria. If the criteria for pain are given in general and behavioural terms, they are too vague to underpin knowledge. Such 'criteria' do not only *mean* pain. So one might think of them as particular though still behavioural. If so, only particular instances of behavioural criteria (particular instances of crying out and rubbing knees etc) are valid guides to underlying pain. Such a suggestion maintains the behavioural character of criteria for mental states but denies their generality. But this threatens the idea that one can learn how to recognise pain. The alternative is to maintain (something of) their generality but deny the restriction to merely behavioural signs and symptoms. On such an account, the criteria for pain do not have in common anything that could be given in mind-free behavioural terms. Rather they share the essentially mind-involving generality of being *expressions of pain*.

McDowell offers a philosophical diagnosis of why such a view of criteria seems to go unnoticed which goes back to the influence of Cartesian dualism. If one starts from that basic picture then it invites a contrast between the behavioural states of other people to which one can have direct perceptual access, and mental states, which are, in some sense, hidden behind them. According to Descartes, they even exist in different kinds of space (*res cogitans* and *res extensa*). Cartesian dualism suggests an alienated picture of human behaviour in which all that anyone else can ever see is bodily movement which is only contingently associated with minds. Because perception of, and judgements about, such 'behaviour' is taken to be unproblematic whilst access to other people's

*mental* states is taken to be problematic, a route is needed from one to the other. Thus it seems plausible to think that judgements about mental states have to be grounded in independent judgements about behaviour. The alienated picture of human behaviour survives in approaches to the philosophy of mind which have long since rejected Descartes' conception of the mind as *res cogitans* (or thinking stuff) existing in a different dimension to matter (*res extensa*).

This picture of the relation of mind and body is neither obligatory nor natural, however. One can instead think of mind and body as more closely linked. What one says and does *expresses* what one thinks and feels. Whilst one person's mental states do not themselves fall within the direct experience of another their expression of their mental state does. Such expression is *not* one that is consistent with the *absence* of the inner state. So McDowell replaces an account in which all that is visible to an observer is another person's intrinsically brute or meaningless behaviour, standing in need of further interpretation and hypothesis, with one in which that behaviour is charged with expression.

This claim addresses the worry raised at the end of the previous section that an analogy with a tacit recognitional judgement of a macintosh or a shade of colour suggests that patients and clients are passive in the face of a clinical gaze. If the analogy held closely then one person's mental state would have to fall directly within the experience of another just as a colour can. The nuanced view is that this is not so. Patients and clients have to reveal their mental states through speech and action. But, to continue to describe the nuanced view, what they say and do makes their mental lives available to others in a way that requires no inference. This accounts adds to the more general picture of tacit knowledge in the previous section the further idea that recognitional judgement of others' mental states requires that the other people actively express them.

By denying that our 'access' to the minds of others must proceed through a neutrally described behavioural intermediary (their behaviour), McDowell can offer a much less technically charged account of criteria which he summarises thus:

I think we should understand criteria to be, in the first instance, ways of telling how things are, of the sort specified by "On the basis of what he says and does" or "By how things look"; and we should take it that knowledge that a criterion for a claim is actually satisfied - if we allow ourselves to speak in those terms as well - would be an exercise of the very capacity we speak of when we say that one can tell, on the basis of such-and-such criteria, whether things are as the claim would represent them as being. [McDowell 1982: 470-1]

Knowledge of other minds depends on what people say and do. It does not require a kind of direct mind reading. The judgement is based on, emerges from, what they say and do. But the conceptualisation of what they say and do need not be couched in mind-independent neutral terms. As Dowell comments:

This flouts an idea we are prone to find natural, that a basis for a judgement must be something on which we have firmer cognitive purchase than we do on the judgement itself; but although the idea can seem natural, it is an illusion to suppose it is compulsory. [McDowell 1982: 471]

It may be easier to see patterns and generalities in behaviour construed as essentially expressive of minds than in neutrally described bodily movement. So even though judgements about others' minds may be based on their behaviour, the description of the behaviour may be less secure than the description of what it expresses.

I have set out two contrasting accounts of criteria from the philosophical discussion of the problem of other minds to shed light on the more specific issue of mental illness diagnosis. There are, however, two related important differences between the two cases which need mention.

First, the application of the idea of criteria in the more general problem of other minds and in the case of psychiatric diagnosis differ in one clear respect. It is merely a theoretical idea in the former case but set out in practical detail in recent editions of the DSM and ICD in the latter case. Second, and related to this, is an important difference in the dialectical context of criticism of behavioural criteria in the two cases. The argument above assumes that it is possible to have knowledge of other minds. Since the standard model of criteria (as defeasible behavioural types) makes knowledge impossible, it cannot be the basis of our knowledge of other minds.

But one might object that psychiatry does not aspire to *knowledge* when it comes to diagnosis but some weaker state such as a belief with a particular degree of probability. And hence an argument which shows that knowledge cannot be based on criteria, so understood, need not undermine that project. Such an objection carries risk, however. Since psychiatry is a practical discipline, diagnoses form the basis for action (concerning treatment and management). Thus clinicians need more than merely having beliefs with a particular (suitably high) probability of being true, they need to know that they do.

Nevertheless, even if psychiatric diagnosis need not aspire to knowledge itself but merely to some known probability of being correct, it could be based on criteria understood as behavioural types (ie the target of the criticism of this section). Providing that there are other methods of arriving at diagnoses, such as the considered judgement of skilled clinicians or longitudinal studies, it would be possible to make an assessment of the sensitivity and specificity – in probabilistic terms – of types of behavioural criteria. The dialectical context differs for defenders of defeasible criteria for knowledge of other minds because they assume that there is no more fundamental way of having such knowledge and hence no independent test of the construct validity of the criteria.

Despite these differences, McDowell's discussion of the two accounts of criteria and the role, in the account he defends, of the idea that behaviour can be more than mere behaviour but rather expressive mental states sheds light on the possibility, at least, of the relative vagueness of criteriological diagnosis compared to the specificity of gestalt judgement. Both the DSM and ICD stress operationalised descriptions as opposed to more essentially psychiatric descriptions couched in aetiological terms. They do this in an attempt to provide secure foundations for diagnosis. But that very strategy makes the criteria mere approximations of the underlying psychopathological states they aim to capture. As Kraus, Maj and Parnas suggest, precision requires thinking of psychiatric symptoms as abstractions from a diagnostic whole rather than built up from neutral – or more neutral – criteria whose obtaining does not strictly imply the presence of the psychiatric syndrome for which they are supposed to be signs.

An alternative view of diagnostic criteria, drawing on McDowell's account and influenced by the empirical claims of Kraus, Maj and Parnas would stress the specific schizophrenic colouring of particular delusions, for example. It may seem that this carries the risk that identifying that a patient or client is experiencing such a delusion is riskier than the vaguer claim that they are experiencing some sort of delusion or other. But this may not be so in context. In particular cases, the justification for thinking that the delusion carries a specific schizophrenic colouring may be what warrants the more general claim that they are thus experiencing some more general category of delusion.

This view also helps address a suggestion in Kraus' and Parnas' description mentioned earlier: that the connection between symptoms and psychopathological state is more direct than a mere evidential or causally indicating relation. The state is expressed directly in signs and symptoms to those, at least, with the skill to see it. On the view developed above, skilled clinicians do not merely infer the diagnostic state of their patients and clients from signs and symptoms that are independent of or distinct from them. Rather, they see (or hear) in what their patients say and do the expression of a diagnostic condition.

It is natural to object to such a view (as the editors of this book did) that clinicians are fallible beings, too and so the shortcoming of the criteriological approach cannot be that criteria do not strictly

imply the presence of what they are criteria for. But on the view sketched, this objection presupposes the wrong account of the fallibility of such judgements. If criteria for mental illnesses were both general and defeasible, that would explain how knowledge claims could fail but it would also fail to explain how knowledge is ever possible. On the alternative view sketched above, when all goes well a skilled clinician is able to respond to the expressions of, say, schizophrenia which do indeed necessitate that the patient has schizophrenia. Fallibilism is explained by the fact that some apparent criteria for schizophrenia are not in fact such criteria. But it is a mistake to assume that the best that even a skilled clinician can rely on is a description of the signs and symptoms that merely indicates that it is likely that someone has that syndrome. Skilled judgement is more precise than the vague descriptions of symptoms found in the DSM.

### **Conclusions**

I have considered the charge made against criteriological models of diagnosis that, compared with the gestalt judgement of a skilled clinician, criteriological descriptions of symptoms are essentially vague. I have argued that two independently plausible considerations help explain how this could be so. Epistemologically, diagnosis based on gestalt judgement could be akin to the kind of context-dependent practical skill that underpins one model of tacit knowledge. Such skill resists codification in general context-independent terms akin to the DSM and ICD's diagnostic criteria but is nevertheless a form of conceptually structured knowledge. Ontologically, the diagnostic criteria of the DSM and ICD may be merely more or less behavioural abstractions from underlying psychological reality. Skilled clinicians need not rely on neutral criteria but on the direct expression of complex psychological wholes.

### **Acknowledgement**

This chapter was written whilst a fellow of the Institute for Advanced Study, University of Durham. My thanks both to the IAS, Durham and the University of Central Lancashire for granting me research leave.

### **References**

- Bayer, R. and Spitzer, R.L. (1985) 'Neurosis, psychodynamics and DSM-III' *Archives of General Psychiatry* 42: 187-96
- Brigman, P.W. (1927) *The Logic of Modern Physics*, New York: Macmillan
- Fulford, K.W.M., Thornton, T. and Graham, G. (2006) *The Oxford Textbook of Philosophy and Psychiatry* Oxford: Oxford University Press
- Hacker, P.M.S. (1995) 'Criterion' in Honderich, T. (ed) *Oxford companion to philosophy*, Oxford: Oxford University Press: 171
- Hempel, C.G. (1994) 'Fundamentals of taxonomy' in Sadler, J.S. Wiggins, O.P. and Schwartz, M.A. (eds) *Philosophical Perspectives on Psychiatric Diagnostic Classification*, Baltimore: Johns Hopkins: 315-331
- Kraus, A. (1994) 'Phenomenological and criteriological diagnosis: different or complementary?' In *Philosophical Perspectives on Psychiatric Diagnostic Classification* (ed. J.S. Sadler, O.P. Wiggins, and M.A. Schwartz) Baltimore, MD: Johns Hopkins University Press, pp. 148-162.
- McDowell, J. (1982) 'Criteria, defeasibility and knowledge' *Proceedings of the British Academy* 68: 455-79
- Maj, M. (1998) 'Critique of the DSM-IV operational diagnostic criteria for schizophrenia' *The British Journal of Psychiatry* 172: 458-460
- Parnas, J. (2011) 'A Disappearing Heritage: The Clinical Core of Schizophrenia' *Schizophrenia Bulletin* 37: 1121-1130

- Polanyi, M. (1962) *Personal Knowledge*, Chicago: University of Chicago Press
- Polanyi, M. (1967a) 'Sense-giving and sense-reading' *Philosophy* 42: 301-325
- Polanyi, M. (1967b) *The Tacit Dimension*, Chicago: University of Chicago Press
- Ryle, G. (1949) *The Concept of Mind*, London: Hutchinson.
- Ryle, G. (1945) 'Knowing How and Knowing That' *Proceedings of the Aristotelian Society*, 46: 1-16
- Shorter, E. (1997) *A History of Psychiatry*, New York: John Wiley and Sons
- Thornton, T. (2013) 'Clinical judgement and tacit knowledge' for Fulford, KWM (Bill) et al (ed) *Oxford Handbook of Philosophy and Psychiatry*, Oxford: Oxford University Press
- Wittgenstein, L. (1953) *Philosophical Investigations*, Oxford: Blackwell.