

This is the Authors' Accepted Manuscript version of R.Cooper and R.K. Blashfield (forthcoming) 'The myth of Hempel and the DSM-III' *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences.*

The Myth of Hempel and the DSM-III

Rachel Cooper (corresponding author)

Philosophy, University of Lancaster, Lancaster, LA1 4YL, UK

Email: r.v.cooper@lancaster.ac.uk

Roger Blashfield

Psychology, Department of Psychology, 226 Thach Hall,

Auburn University, AL 36849-5214, U.S.

Email: blashrk@auburn.edu

Abstract

In 1959, the philosopher Carl Hempel presented a paper on psychiatric taxonomy at a conference of the American Psychopathological Association. In 1980, the American Psychiatric Association published DSM-III, the third edition of their hugely influential classification of mental disorders. The DSM-III sought to adopt an ‘atheoretical’ approach to classification, and introduced explicit diagnostic criteria setting out the number and combinations of symptoms required for diagnosis. Commentators now often claim that Hempel's paper was an important contributor to the DSM-III approach. This paper argues that this claim is mistaken and that the idea that Hempel influenced the DSM-III is a myth. This matters because the idea that Hempel influenced the DSM-III has played a key rhetorical role in discussions about the potential relevance and importance of the philosophy of psychiatry.

Key words

Hempel; operational definition; operationalism; diagnostic criteria; DSM-III

The Myth of Hempel and the DSM-III

The *Diagnostic and Statistical Manual of Mental Disorders* (DSM) is a classification of mental disorders published by the American Psychiatric Association (APA). This diagnostic manual now structures mental health research and influences clinical practice around the world. In 1959 Carl Hempel, the philosopher of science, presented a paper on psychiatric classification to the American Psychopathological Association. It is often claimed that this paper inspired the style of diagnostic criteria that has been employed in the DSM since its third edition (DSM-III, 1980).

The APA published the first edition of the DSM in 1952. The first and second editions of the DSM (1968) were slim volumes and were used mainly in the United States. In these editions, brief narrative paragraphs described typical manifestations of each disorder, and some descriptions made reference to psychodynamic concepts and theory. The third edition of the classification (DSM-III), published in 1980, broke with earlier tradition in multiple ways. Most obviously, the DSM-III was much bigger than the earlier editions, and it came to be used around the world.¹ In addition, among a number of other conceptual innovations, the DSM-III aimed to take a ‘descriptive’ or ‘atheoretical’ approach to classification; the idea was that symptoms should be described using as little

¹ We cannot fully discuss here the multiple ways in which the DSM-III differed from the predecessors and the reasons why it came to be so widely used. A detailed history of the DSM-III can be found in Decker 2013.

theory as possible (APA, 1980, pp.6-7).² The DSM-III also introduced explicit diagnostic criteria setting out the number and combinations of symptoms required for diagnosis.

While in the DSM-I and II a condition might be described as having a ‘prolonged course’; in the DSM-III there might be a requirement for ‘duration of at least three months’. In the DSM-I and II, symptoms are often described as ‘common’ or ‘typical’. In the DSM-III, explicit ‘polythetic’ criteria are employed; these make it clear exactly how many symptoms must be present (say, three from a list of five) before a diagnosis can be made.

It is regularly claimed that the DSM-III approach to diagnostic criteria was motivated by Hempel’s 1959 paper. For example, Allan Frances (who was chairman of the DSM-IV Task Force) has written (with co-authors who were also closely involved in the development of DSM-IV),

DSM-III introduced methodological innovations based on the earlier recommendations of Hempel (1961) and Stengel (1959) for explicit,

² It is widely accepted that the DSM-III failed to be ‘atheoretical’ (see Cooper and Michels (1981), Wakefield (1999)). Still the attempt to be atheoretical clearly affected the style of diagnostic criteria. The DSM-I and II referred explicitly to psychodynamic concepts. In the DSM-III explicitly theoretical concepts were avoided as much as possible.

descriptive criteria and a neutrality with respect to theoretical orientation.
(Frances et al, 1994, p.13)

Similarly, a recent book on the DSM-5 (2013) tells readers that,

DSM-III's (1980) revolutionary neo-Kraepelinians were dedicated to setting up a research program...Embracing Carl Hempel's logical empiricist agenda, they approached mental disorders in terms of operational definitions for the purpose of enhancing reliability in diagnosis (Mishara & Schwartz, 2013, p.125)

Similar claims have been made by many others (e.g. Schwartz & Wiggins 1986; Turner, 2003; Bentall, 2004; Sadler, 2005, p.74, Chung, 2007; Bolton, 2008, p.3; Bolton, 2012; Millon, 2012).

In this paper we re-evaluate the impact of Hempel's paper, and we show that the idea that Hempel influenced the DSM-III is a myth. The paper falls into four sections: In section 1, we discuss the context and contents of Hempel's paper. In section 2, we discuss the early reception of Hempel's paper. In section 3, the distinctive style of diagnostic criteria included in the DSM-III is traced to other origins. In section 4, we consider how it came to be widely believed that Hempel inspired the DSM-III.

The question of whether Hempel influenced the DSM-III is important chiefly for

understanding the potential role of philosophy for psychiatry. As we discuss in detail in section 4, the claim that Hempel influenced the DSM has often been used to demonstrate the potential significance of philosophy for psychiatry. If we are right, and Hempel had no impact on the DSM-III, then whether, and how, philosophy might contribute to psychiatry requires re-evaluation.

More broadly, assessing the role of Hempel's work in the evolution of the DSMs contributes to an understanding of the nature, history, and likely future development of the classification. Most obviously, analyzing Hempel's impact can inform discussions of the philosophical assumptions that underpin the DSM enterprise. A number of scholars have argued that the DSM is a 'logical empiricist' project, which might be better replaced by a classification based on other epistemological foundations (Schwartz and Wiggins 1986, 1987; Mishara and Schwartz 2013). Quantifying Hempel's influence is also relevant to debates concerning the relative input of different academic communities on the development of the DSM. On some accounts, the DSM-III was produced by a small 'invisible college' of like-minded psychiatrists (Klerman 1978, Blashfield 1984). If Hempel had a major impact on the creation of the DSM-III, this fact would support the view that the DSM-III was the result of broad scholarly input, rather than the work of a tight group of insiders (Kendell 1982b).

The role of Hempel's work in the development of the DSM series has previously been considered by other scholars. In a two-part paper, the psychiatrist and philosopher Massimiliano Aragona (2013a, b) has examined the historical and epistemological

relationship between neopositivist philosophy of science and the DSM-III. Aragona argued that it is unlikely that Hempel directly influenced the DSM-III, although in his view the DSM-III approach was consistent with Hempel's advice regarding classification. Aragona concluded that the DSM-III could fairly be seen as resting on neopositivist foundations, and he suggested that developing a classification on different epistemological presuppositions would be preferable. We agree with Aragona that it is unlikely that Hempel influenced the DSM-III, but in section 1.4 we argue that his interpretation of Hempel's paper is mistaken.

The influence of Hempel's paper has also been considered by Bill Fulford and Norman Sartorius (2009). Fulford is a psychiatrist and philosopher who has played a key role in establishing the subdiscipline of philosophy of psychiatry (Fulford, no date). Sartorius is a psychiatrist who for many years was head of the Division of Mental Health at the World Health Organization (Sartorius, no date). Fulford and Sartorius (2009) proposed that Hempel's paper indirectly led to DSM-III style classification systems, but only through influencing the thinking of the London-based psychiatrist, Aubrey Lewis. In their narrative, a philosopher and a psychiatrist both played indispensable roles in the conceptual development of psychiatric classification. Our reading of the argument of Hempel's paper is in broad agreement with Fulford and Sartorius, but in section 2.2 we dispute their suggestion that Hempel influenced the DSM-III through the work of Lewis.

The main contribution of this paper is that we bring together the evidence that makes it plausible that Hempel did not influence the DSM. We also trace how the myth of

Hempel's influence emerged, and how it has played a key role for authors seeking to demonstrate that philosophy is of potential relevance to psychiatry.

1. Hempel's paper.

1.1 The context of Hempel's paper: The 1959 conference

Hempel's paper was presented at a conference of the American Psychopathological Association in 1959. The conference was organized by Joseph Zubin, who was a prominent research psychologist at the Biometrics Unit at the New York State Psychiatric Institute (Zubin, 1961, p.v). The conference focused on methodological problems with 'field studies' of the mental disorders. Such studies were mainly epidemiological studies about variations in the incidence of mental disorders, and attempts to assess the efficacy of therapies in clinical settings. When Zubin organised the conference he had just completed a tour of European psychiatric centres. In his meetings with international colleagues, Zubin found a widespread interest in field studies. However, differences across countries in the use of psychiatric diagnostic terms made comparing the results of such studies difficult.

The 1959 conference was funded by a grant from the National Institute of Mental Health (NIMH) (Zubin, 1961, p.v). At the time, the NIMH had a large and rapidly increasing budget (increasing from \$50 million in 1959 to \$189 million in 1964), and funded much epidemiological research into the social determinants of mental health (Grob, 1991, pp.6,

64, 68). Many NIMH grants went to interdisciplinary groups of researchers, often headed by psychologists (who tended to have greater training in statistics and research methodology than did psychiatrists in this period) (Grob, 1991, p.66). In the late 1950s, other organisations were also investing heavily in epidemiological research into mental health. The 1959 conference can be seen as one in a series of international conferences on related themes that were sponsored by bodies including the NIMH, the Millbank Memorial Fund, the British Medical Research Council, the World Federation for Mental Health and World Health Organisation (World Health Organization, 1960, p.3).³

After the conference, a number of attendees (M.M. Frohlich, Ernest Gruenberg, Lothar Kalinowsky, Morton Kramer, and Benjamin Pasamanick) went on to serve on the committee that developed the DSM-II (Zubin, 1961, pp. 487-489; American Psychiatric Association 1968). Some of those present (Paul Hoch, Morton Kramer, Aubrey Lewis, Benjamin Pasamanick and Joseph Zubin) also went on to be involved in the US-UK Diagnostic Project (Cooper et al, 1972, pvii). The US-UK Diagnostic Project ran from

³ The introduction to the WHO report 'Epidemiology of Mental Disorders' (Expert Committee on Mental Health, 1960) presents the 1959 conference where Hempel spoke as part of series of conferences. The series included a meeting of the WHO Expert Committee on Mental Health in Geneva (held 1959), a meeting at the London School of Hygiene and Tropical Medicine, held 1958 and sponsored by the Milbank Memorial Fund, the British Medical Research Council, the World Federation for Mental Health, and the World Health Organization,

1966 to 1971 and was designed to investigate cross-national differences in the diagnosis of schizophrenia. Zubin headed the project, which was also funded by the NIMH. In his introduction to the US-UK project results, Zubin wrote that he organised the 1959 conference ‘to consider the possibility of introducing a more uniform vocabulary and a more comparable classification system’ but ‘it became clear that conferences would not settle the matter and that a field study was required [i.e. the US-UK project]’ (Cooper et al, 1972, p.ix). To sum up, the 1959 conference where Hempel spoke was an important international conference, which was itself part of a larger programme of activities that aimed to improve methods of classification to facilitate international research in mental health.

Zubin invited Hempel to speak at the conference about the general issues of classification in science (Hempel, 1961, p.3). Hempel would have been an obvious choice for this task. When Hempel presented his paper, he was a Professor of Philosophy at Princeton University and had already established himself as one of the most important philosophers of his generation. Hempel had also previously written on classification in psychology. His first monograph (written with P. Oppenheim) *Der Typusbegriff im Lichte der neuen Logik* (1936) was a work in applied logic that examined the concept of ‘psychological type’.

1.2 Hempel's paper

The original version of Hempel's paper was titled ‘Introduction to the Problems of Taxonomy’, and was first published in the conference proceedings (Zubin, 1961). It was

then reprinted as ‘Fundamentals of Taxonomy’ in Hempel’s collection of essays *Aspects of Scientific Explanation* (1965a). The two versions of the paper are identical (except for some very minor changes in wording as might be made by a copy-editor).

In his paper, Hempel stated his goal as intending to ‘provide a systematic background for a discussion of the taxonomy of mental disorders’ (Hempel, 1961, p. 3). Hempel divided his paper into six parts:

1. Object of this paper
2. Classes and concepts
3. Description and theoretical systematization as two basic functions of scientific concepts
4. Empirical import of scientific terms: Operational definition
5. Systematic import and ‘natural’ classifications
6. From classificatory to comparative and quantitative concepts

[These titles are quoted directly from Hempel (1961)]

Part 1 was a brief introduction. In Part 2, Hempel clarified some points regarding terminology. For example, he defined ‘classes’, ‘concepts’ and ‘extension’. In Part 3, Hempel presented the scientific enterprise as one which aims first to describe, and then to explain and predict phenomena. During the ‘natural history’ stage, scientists describe phenomena, while during the ‘theoretical stage’ theories are developed. Hempel explained that by ‘theoretical’ concepts he meant those that ‘refer to various theoretically postulated entities, their characteristics, and the processes in which they are involved; all

these are more or less removed from the level of directly observable things and events' (1961, p.6). Examples of theoretical terms in science included 'the electric and magnetic fields of physics... chemical valences; molecular and atomic structures...' (1961, pp.6-7). In Part 4, Hempel noted that science aims to be objective and that a widely recommended method for achieving this relative objectivity is via the use of operational definitions, as described by the physicist P.W. Bridgman. Hempel told his audience that 'An operational definition for a given term is conceived as providing objective criteria by means of which any scientific investigator can decide, for any particular case, whether the term does or does not apply to that case' (1961, p.8). For example, the concept of 'harder than' can be operationally defined in terms of the scratch test. Hempel recommended that 'operation' should be interpreted liberally, such that direct observation qualifies as an acceptable operation (1961, p.10). Hempel went on to say,

To apply the preceding considerations to the taxonomy of mental disorders:
If a classificatory system is to be used with a high degree of uniformity by different investigators, the concepts determining the various subclasses will have to possess clear criteria of application that can be stated in terms of publically ascertainable characteristics. (1961, p.11)

In Part 5, Hempel explained that, as science progresses, he expected classification to become more dependent on theory.

the concepts used in a given field of scientific inquiry will change with the systematic advances made in that field: the formation of concepts will go hand in hand with the formulation of laws and, eventually, of theories...In the course of such development, classifications defined by reference to manifest, observable characteristics will tend to give way to systems based on theoretical concepts. (1961, p.16)

Hempel's expectation was that psychiatric classification would become progressively more and more theory driven. He claimed that, with the DSM-I (which was the US classification in use when he spoke), the shift to a more theory-based classification was already occurring. He ended, in Part 6, by noting that categorical classifications often give way to dimensional systems in order to cope with borderline cases, and he commented that this might also happen with classifications for psychopathology (1961, pp. 20-21).

Those who think that Hempel inspired the approach of the DSM-III see his paper as a major influence on the changes in psychiatric classification during the last half of the 20th century. Hempel is claimed to have (i) proposed a new style of classification, through the introduction of the concept of 'operational definitions' (Kendell, 1975; Brockington et al., 1978; Farmer, 1997; Bentall, 2004; Cole et al., 2008), and to have ii) attacked the status quo in classification by criticizing the use of psychodynamic concepts in the DSM-I (Frances et al., 1994; Aragona, 2013a). We will argue that neither claim is true.

Hempel did not introduce operational definitions to psychiatry, and he did not attack the approach taken by the DSM-I.

1.3 Hempel did not introduce 'operational definitions' to psychiatric classification

While it is true that Hempel spoke approvingly of the use of 'operational definitions', we suggest it is unlikely that Hempel introduced this concept to his audience (Kendler et al., 2010 and Aragona, 2013b agree). The conference where Hempel spoke was about epidemiology and the measurement of psychopathology. Psychologists and psychiatrists interested in this field would likely have already been familiar with the idea that key scientific concepts should be 'operationally defined' when Hempel spoke.

The idea that scientists should make use of 'operational definitions' became widely discussed by psychologists shortly after the publication of P.W. Bridgman's (1927) *The logic of modern physics*. Bridgman, a physicist, explained how concepts might have an 'operational character',

The concept of length is therefore fixed when the operations by which length is measured are fixed: that is, the concept of length involves as much as and nothing more than the set of operations by which length is determined. In general, we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations. (1927, p.5)

By the 1930s, Bridgman's ideas were being discussed in psychology journals. Influential early advocates of operational definitions in psychology included B.F. Skinner (1931), S.S. Stevens (Stevens, 1935a, 1935b), and E.C. Tolman (1951 [1936]). Operationalist ideas became especially popular in the guise of behaviourism, and in debates about the interpretation of IQ tests (Rogers, 1989). In 1945, one of the most important psychology journals, *Psychological Review*, had a special issue devoted to discussion of operationalism (Vol 52 (5), Sept 1945).⁴ Psychologists would have been aware of operational definitions prior to Hempel's paper.

In psychiatry, too, there had been discussion of operational definitions prior to Hempel's paper. In his 1952 presidential address to the American Orthopsychiatric Association, James Cunningham (1952) discussed Bridgman's work and advocated for the use of operational definitions. A. Ellis (1956) had attempted to operationalise key psychoanalytic concepts. Thomas Szasz employed operationalist ideas to argue that there are important ambiguities in concepts of 'neurosis', 'psychosis' and 'normal' (Hollender and Szasz, 1957; Szasz, 1959). We conclude that in the late 1950s, psychologists and

⁴ Rogers (1989) and Green (1992) examine papers in this special issue and show that by this time 'operationalism' in psychology referred to a fairly diverse family of views. 'Operationalists' of all stripes would have been familiar with the idea that 'operational definitions' could improve reliability.

psychiatrists interested in research on classification would already have been familiar with the idea that key scientific terms should be ‘operationally defined’.

1.4 Hempel did not attack the DSM-I

According to Aragona (2013a), Hempel considered the DSM-I to be a psychodynamic classification,⁵ and also thought that psychodynamic concepts could not be given adequate operational definitions. Aragona claimed that Hempel’s paper therefore mounted a powerful argument against the DSM-I,

To sum up, Hempel asserts that: (a) the terms used in the DSM-I classification system (e.g. those defining a conversion reaction) ‘refer neither to directly observable phenomena, such as overt behavior, nor to responses that can be elicited by suitable stimuli’ (p. 140), and (b) there are no definite ways of putting psychodynamic theories ‘to a test by applying them to concrete cases’ (p. 141). On this basis he can conclude that the DSM-I mental disorders did not meet either the empirical interpretation criterion or the testability principle, and for this reason the DSM-I taxonomy should be rejected not as false but as meaningless. (Aragona, 2013a, p.174)

5 In later work, Aragona (2015) argues that the DSM-I was only partly psychoanalytic.

We disagree with Aragona's reading of Hempel's argument. We agree with Aragona that Hempel thought that the DSM-I had started psychiatric classification along a theoretical path, but we will argue that Hempel did not think that the use of psychodynamic theory made the DSM-I illegitimate (Wakefield, 1999 and Fulford & Sartorius 2009 concur).

It is true that Hempel had concerns about some specific uses of psychodynamic concepts. He mentioned that 'One of the main objections against various types of contemporary psychodynamic theories, for example, is that their central concepts lack clear and uniform criteria of application...' (Hempel, 1961, p.8) and he also stated that

the operational reformulation of psychoanalytic concepts proposed by Ellis, which relies on such 'operations' as thinking, remembering, emoting, and perceiving (in an enormously comprehensive sense) provides no clear criteria of application for the terms of psychoanalysis and no objective ways of testing psychoanalytic hypotheses (Hempel, 1961, p.11).

Clearly Hempel thought that some uses of psychodynamic concepts were problematic. However, we interpret these comments as attacks on specific problematic examples. We think that Hempel did not reject the use of psychodynamic concepts in general.

In support of our reading, consider this passage, where Hempel said that the DSM definition of conversion disorder made reference to theoretical notions, e.g., 'the impulse

causing the anxiety is “converted” into functional symptoms in organs’ (Hempel, 1961, p.7 quoting American Psychiatric Association, 1952, p.32). Hempel commented:

Clearly, several of the terms used in this passage refer neither to directly observable phenomena, such as overt behaviour, nor to responses that can be elicited by suitable stimuli. Those terms have a distinct meaning and function only in the context of a corresponding theory; just as the terms ‘gravitational field’, ‘gravitational potential’, and so on have a definite meaning and function only in the context of a corresponding theory of gravitation (Hempel, 1961, p.7)

Since ‘gravitational field’ is a theoretical, yet well accepted, concept in science, we infer that Hempel saw no inherent problem with concepts such as ‘conversion.’ A slightly earlier paper ‘A logical appraisal of operationism’ (1954) helps clarify Hempel’s views. In this paper, Hempel made it clear that he did not suppose that all properly scientific terms would be capable of simple operational definition. He supported Pierre Duhem’s claim that empirical testing is ‘properly applicable to entire theoretical systems rather than to individual hypotheses’ (Hempel, 1954, p.219). We conclude that Hempel saw no intrinsic problem with the use of theoretically-defined terms, and no intrinsic problem with psychodynamic theory. Indeed, Hempel concluded his paper by stating that psychiatric classification will continue to become increasingly theoretical. He said

It is not for me to speculate on the direction that theoretical developments in this field may take and especially on whether the major theories will be couched in biophysiological or biochemical terms or rather in psychodynamic terms lacking in over-all physiological or physiochemical interpretation. Theoretical systems of either kind can satisfy the basic requirements for scientific theories. (Hempel, 1961, p.18)

This quote is consistent with Hempel's comments elsewhere regarding psychodynamic and psychoanalytic theory. In writings from the late 1950s and early 60s, Hempel presented psychoanalytic explanations as somewhat sketchy, but potentially improvable, scientific explanations (1965b [1958], 1965c [1959], 1965d).

We conclude that, although Hempel thought that the DSM-I was partly based on psychodynamic theories, Hempel did not consider this to be a weakness in the DSM-I classification. We think that Hempel considered the DSM-I to be a reasonable step in a developing science, rather than a mistake that needed correcting.

2. Early reception of Hempel's paper

In our reading, Hempel did not introduce operational definitions to psychiatry, and he did not attack the DSM-I. The initial reception of Hempel's paper was consistent with our claims.

2.1 Early reception

Erwin Stengel, then a Professor of Psychiatry at Sheffield University, U.K. was the discussant of Hempel's paper at the conference. His discussion was published alongside Hempel's paper in the conference proceedings (Stengel, 1961). Much of Stengel's discussion was devoted to the question of whether psychiatric conditions could be considered mutually exclusive, as Hempel supposed, and with a consideration of the extent to which current psychiatric classifications might be considered 'descriptive' or 'theoretical'. Stengel also discussed, and approved of, Hempel's suggestion that diagnostic terms should be operationally defined. Stengel was clearly already familiar with the idea of operational definitions. Indeed, Stengel had himself previously argued that researchers talking about 'psychopathic personality' should provide an 'operational definition' for this concept (Stengel, 1958).

At the time of the 1959 conference, Stengel was working on a review of international psychiatric classifications for the WHO, and in addition to commenting on Hempel's paper, he presented an early version of his report at the conference. In the final version of his report, Stengel included a half-page summary of Hempel's paper (1959, pp.611-612). In Stengel's precis, Hempel's paper was characterised as providing a useful conceptual framework for thinking about classification. Stengel did not interpret Hempel as being critical of the DSM-I. Stengel's report recommended that diagnostic terms should be operationally defined but he did not cite Hempel as a source for this idea (1959, pp.612 and 619).

Early citations to Hempel by Michael Rutter, and by J.K. Wing, made somewhat similar use of Hempel's paper. Michael Rutter was a child psychiatrist, based at the Institute of Psychiatry, London. His much-cited 1965 discussion of classification in child psychiatry cited Hempel's paper and also advocated the adoption of 'operational terms', but Rutter did not cite Hempel in relation to operational definitions. Instead, the reader was directed to Hempel (1961) only as a source where 'the principles of classification have been clearly outlined' (Rutter, 1965, p.72). J.K. Wing, also based at the Institute of Psychiatry, developed explicit rating scales for psychiatric symptoms (most famously the Present State Examination (Wing et al. 1974)). In an early paper setting out a rating scale for schizophrenic symptoms, he cited a preprint of Hempel's paper – but only in relation to the claim that a classification should be mutually exclusive and jointly exhaustive (Wing, 1961, p.866).

Generally, in the first decades following Hempel's talk, when he was cited, Hempel's paper was interpreted as a handy overview of the role that classification might play in a science, and as outlining some of the logical requirements for an adequate classification system. Hempel was used as support for claims such as that classification plays an important role in science, that over time classifications tend to move from being descriptively to theoretically based, and that ideally the categories in a classification should lead to useful generalizations (Zubin, 1967; MacMahon & Pugh, 1970; Paykel, 1972; Strauss, 1973; Becker, 1974; Blashfield & Draguns, 1976; Adams et al., 1977; Wolman, 1978;).

Zubin wrote a critique of an early draft of the DSM-III in 1977. As the conference organiser who invited Hempel to speak, if Zubin thought that Hempel's paper had significantly shaped the DSM-III, Zubin might have been expected to mention Hempel in his review. Zubin did not cite Hempel's paper. Rather Zubin worried that the DSM-III had been shaped by economic and professional concerns, and that the DSM-III could not be considered a scientific classification. He went on to propose a way forward. Zubin suggested the development of an interdisciplinary 'commission for the classification of human behavior' – which would include input from not only psychologists, psychiatrists and epidemiologists, but also philosophers of science, historians, taxonomists, and linguists. A classification of abnormal behaviour would be created as a subset of this broader project (Zubin, 1977, p.7). Clearly, at this stage, Zubin considered constructing a philosophically-informed classification of psychopathology to be a future project as opposed to something that had been achieved or even attempted in the DSM-III. Zubin's failure to cite Hempel in the context of his proposal is notable.

2.2 Aubrey Lewis' Hempel

In their chapter, 'The secret history of ICD and the hidden future of DSM', Bill Fulford & Norman Sartorius (2009) argued that Hempel's paper motivated one audience member - Aubrey Lewis - to develop ideas that led to the development of symptom-based

diagnostic criteria in the DSM-III style. Fulford and Sartorius claimed that Hempel's paper influenced the DSM-III - but via Lewis.

Aubrey Lewis was Professor of Psychiatry at the Institute of Psychiatry, London, from 1948-1966, and was the most influential UK-based psychiatrist in the postwar period (Jones, 2003). Lewis had long been concerned about the methodological difficulties involved in assessing the prevalence of mental disorders, and in evaluating therapeutic efficacy. From as early as 1938, Lewis had been concerned that diagnostic terms were employed differently across various countries (Lewis 2003[1938], pp.89, 107, 120).

Transcripts of the discussion following the oral presentation of Hempel's paper were published in the Conference Proceedings (Zubin, 1961). Aubrey Lewis said this:

In psychiatry to make a classification based on theory is what we all would like, and what we believe we cannot at the moment attain – because, as Dr Hempel clearly stated, the requirements are not met by any of the theories prevailing in psychiatry at the present time.

Therefore I would suggest that for the purpose of public classification [i.e. those for international use] we should eschew categories based on theoretical concepts and restrict ourselves to the operational, descriptive classification... (Zubin, 1961, p.34)

In these comments, Lewis set out two key ideas: (1) current psychiatric theories are unsuitable for use as the basis of a public classification, and (2) public classifications should make use of operational definitions. Lewis's interpretation of Hempel's paper was peculiar. In his paper, Hempel did not claim, let alone 'clearly state', that there was a general problem with existing theories in psychiatry. Whatever the explanation for Lewis's misinterpretation, the idea of that a public classification should be descriptive took on a crucial significance. After the conference, Lewis came to head the committee that developed the UK-produced *A Glossary of Mental Disorders* (Sub-committee on Classification of Mental Disorders, 1968), and then the WHO committee that produced *Glossary of Mental Disorders and Guide to their Classification* (WHO, 1974). In Fulford and Sartorius' telling, the WHO *Glossary* made use of descriptive symptom-based definitions. The *Glossary* then fed into the drafting process of the ninth edition of the International Classification of Diseases (ICD, the classification published by the World Health Organization) and influenced the DSM-III. Fulford and Sartorius concluded 'It was from this glossary, as described earlier, that current descriptive international classifications of mental disorder, the ICD and DSM, are ultimately derived.' (Fulford & Sartorius, 2009, p.37).

We think that Fulford and Sartorius are mistaken and that the DSM-III approach to diagnosis cannot be traced to the WHO *Glossary*. The *Glossary* contains brief paragraphs that describe a characteristic patient. It is true that these descriptions are written in plain language, are generally symptom-based, and make frequent use of examples. However, the WHO *Glossary* lacks much of the specificity found in the DSM-III. Notably, the

Glossary is often vague about the time period during which symptoms must be present. It is also often unclear how many symptoms are required for diagnosis; some symptoms are described as being ‘characteristic’ or ‘common’ without it being clear whether or not they are essential. Take, for example, the description of 300.5 Neurasthenia

Includes a neurosis characterized by fatigue, irritability, headaches, depression, insomnia, difficulty in concentration, and lack of capacity for enjoyment (anhedonia). It may follow an infection or exhaustion, or arise from continued emotional stress.

Excludes depressive neurosis (300.4); anxiety neurosis (300.0); physical disorders of presumably psychogenic origin (305); mental disorders similar to that defined above associated with physical conditions (309) (WHO, 1974, p.40)

In contrast, the DSM-III criteria for Dysthymic Disorder (which the DSM-III index takes to be equivalent to neurasthenia) lists 13 possible symptoms of which at least three must be present, and specifies that symptoms must have been problematic for at least two years (APA, 1980, p.222).

We cannot rule out the possibility that the WHO *Glossary* had some influence on the DSM-III. Robert Spitzer, who became the chairperson for the DSM-III, was familiar with the WHO *Glossary*; he reviewed it for the *American Journal of Psychiatry* (Spitzer, 1975). The *Glossary* and the DSM-III are similar in that both sought to be descriptive,

symptom-based classifications. However, while the DSM-III characteristically specifies time periods and numbers of symptoms, the *Glossary* does not. In his review, Spitzer complained that the *Glossary* (and also the DSM-II) ‘...suffer from a lack of precision, of operational criteria for making the diagnoses...’ (1975, p.881). In developing the DSM-III, Spitzer sought to provide ‘operational criteria’; he did not think that the *Glossary* included such criteria. The key features of diagnostic criteria the DSM-III cannot be traced back via Lewis’s WHO *Glossary* to Hempel.

2.3 Robert Kendell’s Hempel

Robert Kendell was not at the 1959 conference, but was a key early citer of Hempel’s paper. Kendell was a British psychiatrist, working at the Institute of Psychiatry, London, and then at Edinburgh University (Kendell, 2000). He is best known for *The Role of Diagnosis in Psychiatry* (1975), an influential work that continues to be regularly discussed. In our analysis of citations to Hempel’s paper, Kendell is the first author to interpret Hempel as calling for ‘operational criteria’ in the DSM-III style, and the first to claim that Hempel’s paper influenced approaches to classification in psychiatry. Kendell knew Robert Spitzer, as they both worked on the US-UK Diagnostic project (Cooper, 2003), and it is worth considering whether Hempel could have influenced the DSM-III through Kendell’s work.

Kendell cited Hempel in various writings from the late 1970s and early 1980s. Kendell told his readers,

Some years ago, the philosopher Carl Hempel tactfully suggested to an audience of psychiatrists and clinical psychologists interested in problems of diagnosis and classification that they should tackle this situation [problems of diagnosis caused by the polythetic and inexact nature of syndromes] by developing ‘operational definitions’ for all the various categories of illness in their nomenclature’ (Hempel 1961). (quoted from Kendell, 1975, p. 25)

Hempel recommended the use of operational definitions to achieve uniformity. In response to this suggestion, psychiatrists in Europe and America have drawn up a large number of criteria for the diagnosis of schizophrenia, thus replacing inarticulate confusion by a Babel of precise but differing formulations of the same concept. (Brockington, Kendell & Leff, 1978, p. 387)

Kendell similarly cited Hempel in Kendell, Brockington & Leff (1979, p.25) and Kendell (1982a, p.1334). Kendell’s interpretation of ‘operational definition’ was close to that employed in DSM-III. Kendell made it clear that an operational definition of a mental disorder should set out rules detailing how many symptoms, in what combinations, are required for diagnosis. He wrote

The traditional polythetic criterion has to be converted to a monothetic one. This can be done quite easily. Instead of saying that the typical features of

disease S are A,B,C,D & E and most of these should be present before the diagnosis is made, A,B,C,D and E must be combined algebraically, so that which combinations satisfy the criterion O and which do not is specified unambiguously. It might be stipulated for instance that any three, or any four, of the five would satisfy O. (Kendell, 1975, p.26)

This idea was not explicitly explored by Hempel himself. Kendell plausibly interpreted Hempel's call for operational definitions in this way for two reasons. First, Kendell had previously worked using the numerical statistical techniques of factor analysis and cluster analysis to classify varieties of depression. Some of the depression symptom rating scales he used in this work included '3 symptoms from 5' type rules (Kendell, 1968). Second, during his studies in the philosophy of science, Kendell had been impressed by the works of Karl Popper (Kendell, 2000). To a Popperian, specifying the number of symptoms that is required for a diagnosis would seem natural (this specification is what a clinician needs in order to know exactly when someone does or does not qualify for a diagnosis).

Kendell interpreted Hempel as advocating for definitions of a form very like those included in DSM-III. Spitzer read, and liked, Kendell's 1975 book, *The Role of Diagnosis in Psychiatry*. In a review, Spitzer noted that

Kendell points out the central issue of definition – that is, indicating which features of the category are essential and always present, which features are

commonly present but may be absent, and, finally, which features must be absent for the diagnosis to be made. (1977, p.41)

However, we think that the use of polythetic definitions in the DSM-III cannot be traced back to Kendell's book. Kendell's book was published too late to have this influence. By the mid-1970s, when Kendell's book was published, work on the DSM-III was already well advanced (Spitzer et al. 1975a). We cannot rule out the possibility that Spitzer might have been influenced by informal conversations with Kendell during the US-UK Diagnostic Project (Cooper, 2003, p.280). However, whether or not Kendell influenced Spitzer in such a way, another, and better documented, predecessor for the DSM-III diagnostic approach can be found elsewhere, as we will describe in the next section.

3. Tracing the origins of the DSM-III diagnostic approach: The 'neo-Kraepelinians'

The diagnostic criteria included in the DSM-III were explicitly modelled on those of two earlier sets of diagnostic criteria that were developed for research purposes. The DSM-III was patterned on the Research Diagnostic Criteria (RDC) (Spitzer et al., 1975b), which in turn was based on the Feighner criteria (Feighner et al., 1972). The authors of the Feighner criteria were John Feighner, Eli Robins, Samuel Guze, Robert Woodruff, George Winokur and Rodrigo Muñoz. The authors of the Research Diagnostic Criteria were Robert Spitzer, Eli Robins and Jean Endicott. These eight authors and their

immediate colleagues at Washington University in St. Louis and the New York State Psychiatric Institute formed a tight-knit group of researchers (an ‘invisible college’) who have come to be known as the ‘neo-Kraepelinians’ (on the basis that their approach to classification had some commonalities with that of Emil Kraepelin, the great nineteenth century psychiatrist) (Klerman, 1978; Blashfield 1982, 1984). The Feighner criteria and the RDC employed the same style of polythetic definitions (i.e. ‘3 of 5’ rules) that appear in the DSM-III.

There is no evidence that any of the authors working on the Feighner criteria or the RDC were influenced by Hempel’s paper. None of them were in the audience at his talk, and none of them cited Hempel’s paper. When directly asked, both Spitzer and Muñoz have explicitly denied that Hempel had any influence on their thinking (Kendler et al., 2010, p. 140). Where, then, did the idea of constructing diagnostic criteria in this style originate?

In their paper on the history of the Feighner criteria, Kendler et al. (2010) note that polythetic definitions were employed for research on rheumatological disorders in the 1940s and 50s (Jones, 1944; Rutstein, et al., 1956; Ropes et al., 1957; Ropes et al., 1958). Kendler et al. suggest that Guze, who had trained in internal medicine, and then became part of the neo-Kraepelinian group, would have known of this work. However, the evidence that Guze knew of the rheumatism papers is somewhat weak; a Google scholar search suggests that Guze never cited these papers.

We think that Mandel Cohen, a neurologically-oriented psychiatrist at Harvard, provides a better documented link between the use of explicit diagnostic criteria for rheumatological disorders and the introduction of explicit diagnostic criteria in psychiatry. Cohen graduated in medicine from John Hopkins University in 1931, where he worked first as a cardiologist before switching to psychiatry (Healy, 2002). Early in his career, Cohen worked on ‘rheumatic heart disease’, which can occur as a complication of rheumatic fever (Healy, 2002). Thus, Cohen would plausibly have been familiar with the style of diagnostic criteria that had been employed in the 1940s work on rheumatic fever (Jones, 1944, p. 211).

Cohen went on to conduct a series of studies on anxiety neurosis for the Armed Forces. Cohen and his colleagues suspected that the same condition was being given a variety of different names – neuro-circulatory asthenia, Da Costa’s syndrome, soldier’s heart, effort syndrome, etc. In his work Cohen depended on precisely defining the symptoms required for diagnosis. He employed the characteristic ‘3 from 5’ style that would later be used in the Feighner criteria, the RDC, and the DSM-III. For example,

Symptoms of neurocirculatory asthenia as defined for this study included breathing trouble and symptoms from two of the following three groups: (1) palpitation or chest pain; (2) nervousness, dizziness, faintness, attacks or spells, and (3) feelings of fatigue or tiredness or limitation of activity’.
(emphasis added, Wheeler et al., 1950, p.878. For a similar approach see Cohen et al., 1951)

Cohen later went on to employ a similar definitional style when doing research on manic-depression (Cassidy et al., 1957).

Cohen's relationship with Eli Robins would have enabled his influence on the neo-Kraepelinian researchers. Robins worked with Cohen as resident in neurology and psychiatry. Robins was a co-author of a paper with Cohen on hysteria in men (Purtell et al., 1951; Healy, 2002). Cohen was also an author of a study that employed diagnostic criteria that became the basis for the definition of depression in the Feighner criteria (Cassidy et al., 1957; Feighner, et al., 1972; Kendler et al., 2010, p. 136).

Cohen could not have been influenced by Hempel's paper; he started employing the '3 from 5' diagnostic style years before Hempel's presentation. In her history of the DSM-III, Hannah Decker also identified Cohen as an early user of polythetic diagnostic criteria. She suggested Cohen might have been directly influenced by Bridgman, the physicist who first championed operational definitions, as they would have been contemporaries at Harvard (Decker, 2013). While this may have been the case, it is notable that Cohen never cites Bridgman, and his early papers did not use the language of 'operational criteria' (Wheeler et al., 1950; Cohen et al., 1951). We think it plausible that Cohen would have become familiar with the use of diagnostic criteria in this style through his work on rheumatic heart disease.

We think that the style of polythetic criteria employed in the Feighner criteria can be traced back to Cohen. The RDC was modelled on the Feighner criteria. Robert Spitzer

co-authored the RDC, and then later became chair for the DSM-III. Spitzer's interests prior to working on the RDC plausibly also primed his enthusiasm for explicit polythetic diagnostic criteria. During the 1960s, Spitzer created rating scales (such as the Mental Status Schedule) that aimed to improve the reliability with which researchers recorded symptoms (Spitzer, Fleiss, Burdock & Hardesty 1964; Spitzer, Endicott & Cohen 1964). He also developed a computer program, DIAGNO, which made DSM-II diagnoses when given input about a patient's symptoms (Spitzer & Endicott, 1968).

In the early 1960s, other researchers were also developing similar research programmes. Notably, J.K. Wing and his colleagues at the Institute of Psychiatry, London, were also seeking to develop reliable rating scales for psychiatric symptoms, and employed computer-based diagnosis (Brugha et al., 2011). Both Wing and Spitzer went on to work on the US-UK Diagnostic project, and both the Mental Status Schedule (developed by Spitzer and Endicott), and the Present State Examination (developed by Wing and colleagues) were used in the study (Cooper et al., 1972). The project also made some use of Spitzer's computer program, DIAGNO.

The computer programmes for diagnosis, developed by Spitzer and Wing and their colleagues, both employed explicit algorithms to turn symptom ratings into diagnoses. The explicit diagnostic criteria included in the DSM-III, were modelled on the RDC (and, the RDC in turn, on the Feighner criteria), but they can also be seen as paper-based versions of the sorts of rules employed by these computer programmes.

To summarize, we suggest that the approach to diagnostic criteria employed in the DSM-III had the following lineage. Explicit polythetic diagnostic criteria were first employed in work in rheumatism. Cohen, who would likely have known about this definitional approach through his work on rheumatic heart disease, employed a similar approach in studies of anxiety neurosis, manic-depression and hysteria. Robins had worked with Cohen when they both did research on hysteria. Robins then became a leader of the neo-Kraepelinians based in St Louis. The St. Louis group wrote the Feighner criteria. At around the same time, the computer programmes, CATEGO and DIAGNO, were developed by Spitzer and Wing and their colleagues, and both employed explicit algorithms to turn symptom ratings into diagnoses. The Feighner criteria were later modified and expanded, becoming the RDC. And, the RDC proved to be a pilot version for the revolutionary DSM-III. Hempel's paper plays no role in this history.

4. If the myth of Hempel is not true why do people believe it?

4.1 Overall citations to Hempel's paper

Changes in interest to Hempel's paper can be shown graphically in terms of citations. The data for the graph shown in Figure 1 were gathered from Google Scholar (as of 26 January 2018). Not all of these citations are papers connected to mental health – authors interested in classification in other areas also sometimes cite Hempel's paper. In addition, the figures are not fully reliable. Google Scholar misses some citations (particularly from

the earlier decades). Nonetheless the chart gives some insight into the shifts in interest in Hempel’s paper. The total volume of publications in mental health, the scope and accuracy of Google Scholar, and general interest in psychiatric classification, all vary over time. In order to gain some insight into shifts in interest in Hempel’s paper specifically we compare citations to his paper with those to two comparable papers. The comparison papers are Stengel’s review of international classification, discussed in section 2.1 (Stengel, 1959), and a paper by Ward, Beck et al. (1962) ‘The psychiatric nomenclature: Reasons for diagnostic disagreement’ which was published in the *Archives of General Psychiatry*. Ward, Beck et al. (1962) played a key role in convincing psychiatry that ambiguities in diagnostic criteria reduced the reliability of psychiatric diagnosis, and was published around the same time as the papers by Hempel and Stengel.

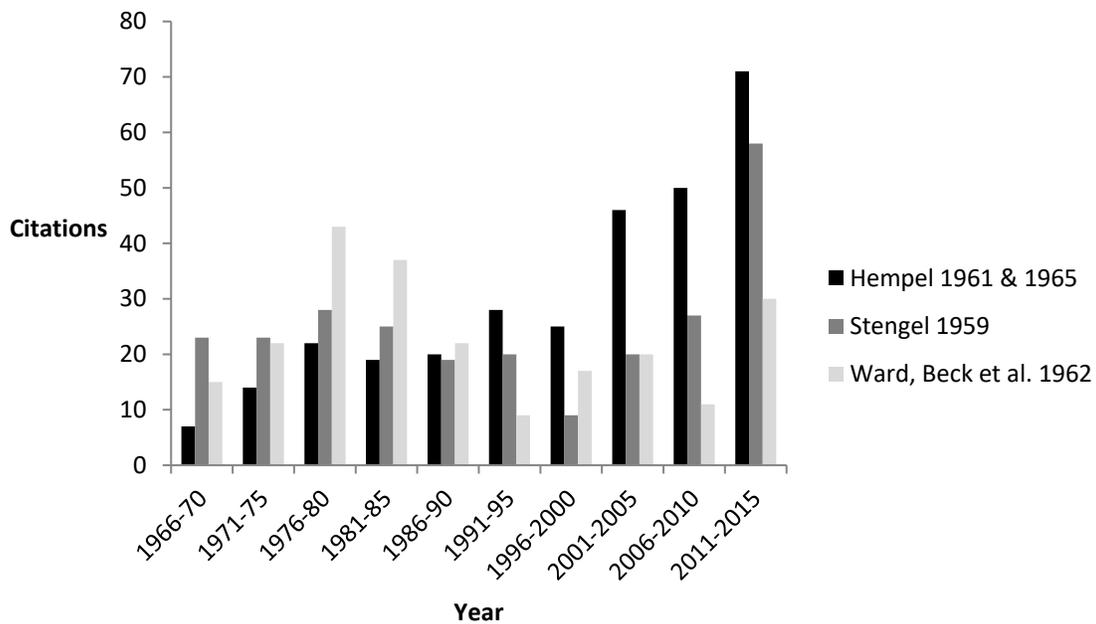


Figure 1. Citations to Hempel (1961 & 1965), Stengel (1959) and Ward, Beck et al. (1962)

In the fifteen years immediately after the publication of these papers and leading up to the publication of DSM-III (1966-1980), there were 80 citations to Ward, Beck et al. (1962), 74 to Stengel (1959) and 43 to Hempel (1961, 1965). The data shown in Figure 1 support our suggestion that Hempel's paper initially had a relatively small impact.

The citation data show a persistent secondary peak in the fifteen years leading to the DSM-5 (2001 to 2015). During this time, Hempel's paper collected nearly four times the number of citations that it did during 1966-1980 (i.e. the time period leading to the DSM-III). Citations to Stengel (1959) increased by a factor of 1.4 when comparing the same time periods. The interest in these papers suggests that the controversies and struggles surrounding the creation of the DSM-5 led a number of authors to think about the conceptual background for the DSM-III. Ward, Beck et al. (1962) was also cited in the lead up to DSM-5, but in this case the citations in the fifteen years leading up to DSM-5 were less than the original peak (25% drop).

The citation data in Figure 1 show that Hempel's paper has come to be considered far more important in recent decades than when it was initially published.⁶ While other papers concerned with fundamental issues connected to psychiatric diagnosis were also

⁶ A more detailed analysis of citations to Hempel's paper, which breaks down citations to Hempel 1961 and 1965 can be found in the supplementary materials.

revisited in the lead-up to DSM-5, interest in Hempel's paper increased more than did interest in comparable papers. Looking through the citations to Hempel's paper, it is clear that Hempel's paper is cited fairly regularly in both mental health journals and in works in the philosophy of psychiatry.⁷ In many cases, Hempel is cited only in passing. For most of those who cite Hempel's paper little hangs on the question of whether he influenced the DSM. However, the idea that a philosopher played a key role in shaping the DSM-III sometimes plays an important rhetorical role, as will be discussed below.

4.2 Hempel employed to argue for the importance of philosophy for psychiatry

Hempel's paper has been fairly regularly cited by writers who seek to demonstrate the importance of philosophy for psychiatry. Consider, for example, the role attributed to Hempel in a piece written for the *Psychiatric Bulletin* to advertise The Royal College of Psychiatrists Philosophy Group:

One final question remains when considering the future of the group.

Besides training and teaching, is there still significant research to be done at the interface of philosophy and psychiatry? In fact, the two disciplines continue to advance each other's knowledge, even if one discounts psychiatric medical ethics. The establishment of operational criteria for

⁷ More detailed discussion can be found in the supplementary materials

psychiatric disorder was an essential step in the development of modern psychiatry. This derived directly from the work of Carl Hempel (e.g. 1961) - a philosopher. (Foreman, 1993, p. 676)

As another example, a 1994 book titled *Philosophical perspectives on psychiatric diagnostic classification*, edited by John Sadler, Michael Schwartz and Osborne Wiggins, includes a reprint of Hempel's paper as an appendix. The introduction to the volume discusses Hempel's paper and its influence on the DSM.

Hempel's work and its place in psychiatric history are a classic example of the formative role of philosophical inquiry in psychiatry... We aspire to build on Hempel's example in this book – to reveal assumptions, to point to alternatives, to answer nosological questions outside the domain of empirical research while maintaining a commitment to an empirical psychiatry. As a monograph of the Association for the Advancement of Philosophy and Psychiatry, this book represents an outgrowth of that organization's values and interests (Sadler, Wiggins & Schwartz, 1994, pp. 2-3)

In these examples the idea that Hempel influenced the DSM is taken to show that philosophy of psychiatry is worth doing, and Hempel's paper illustrates the sort of project that philosophers of psychiatry might usefully undertake (for other papers where references to Hempel play a similar role see Stein 1991; Monti 1996, Fulford 2002;

Fulford, Sallah & Woodbridge, 2007). As a subdiscipline, the philosophy of psychiatry aims to be both academically respectable philosophy and also to be practically useful for psychiatry. The myth that Hempel influenced the DSM-III has played a key role in enabling the presentation of the philosophy of psychiatry as a subdiscipline that can demonstrably achieve both goals; when he presented his paper, Hempel was unquestionably a great philosopher and the DSM is clearly an important classification.

Kari McLeod (2000) examined the key role played in the construction of the discipline of medical geography by the idea that John Snow used mapping to discover the role of polluted water in spreading cholera. She argued that there is ‘a need for disciplines to have myths and for authors to memorialize disciplinary heroes in ways that reflect or help to create disciplinary identity’ (McLeod, 2000, p.932). We suggest that transforming Hempel into a disciplinary role model among those interested in the interface of the mental health field and philosophy has played a similar role in the development of the subdiscipline of philosophy of psychiatry. As in McLeod’s case, once the idea that Hempel influenced the DSM-III became current, references to Hempel’s influence snow-balled. The idea that Hempel influenced the DSM-III has been attractive to philosophers, and so has been regularly, and often somewhat uncritically, repeated.

We suggest that the acceptance of the myth of Hempel among philosophers of psychiatry has had at least one unfortunate consequence; it has made the question of whether and how philosophy might influence psychiatry seem unproblematic. In the myth, Zubin realized that classification might be a problem that could benefit from philosophical

input. Thus, he invited a great philosopher to speak. Hempel delivered a paper that addressed the concerns of psychiatrists, and the psychiatrists of the time then acted on his philosophical suggestions. We think it likely that such a vision of transdisciplinary influence may be naïve. Communicating across disciplines, so that philosophers can appreciate the problems that are of concern to mental health professionals, and so that these professionals will act on philosophical advice, may well be harder than the myth of Hempel suggests.

5. Conclusions

We have shown that Hempel did not argue that psychiatry should adopt an atheoretical approach to classification, and he did not introduce operational definitions in the context of psychiatric nosology. The diagnostic approach adopted by the DSM-III cannot be traced to the work of Hempel but has a different lineage.

We have shown that the idea that Hempel influenced the DSM is a myth. This has implications chiefly for the question of the potential significance of philosophy for psychiatry. The idea that Hempel's influenced the DSM-III has been used to argue that philosophy of psychiatry is worthwhile, and to provide an example of the sort of work that philosophers of psychiatry might seek to emulate. However, while the claim that Hempel influenced the DSM-III has been attractive to philosophers of psychiatry, it is a false claim, and the potential significance of philosophy for psychiatry must be defended on alternative grounds.

Acknowledgements

We are grateful to Massimiliano Aragona, Jared Keeley, Şerife Tekin and Peter Zachar, and to two anonymous reviewers for this journal, who read earlier versions of this paper and made many helpful comments.

References

- Adams, H.E., Doster, J.A. & Calhoun, K.S. (1977). A psychologically based system of response classification. In A. Ciminero, K. Calhoun & H. Adams (Eds.), *Handbook of behavioral assessment*. (pp.47-78). New York: Wiley.
- American Psychiatric Association (1952). *Diagnostic and statistical manual: Mental disorders*. Washington: American Psychiatric Association.
- American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders (Second Edition)*. Washington: American Psychiatric Association.
- American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders (Third Edition)*. Washington: American Psychiatric Association.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (Fourth Edition)*. Washington: American Psychiatric Association.
- Aragona, M. (2013a). Neopositivism and the DSM psychiatric classification: An epistemological history. Part 1: Theoretical comparison. *History of Psychiatry*, 24, 166-179.
- Aragona, M. (2013b). Neopositivism and the DSM psychiatric classification. An

- epistemological history. Part 2: Historical pathways, epistemological developments and present-day needs. *History of Psychiatry*, 24, 415-426.
- Aragona, M. (2015). Rethinking received views on the history of psychiatric nosology: minor shifts, major continuities. In P.Zachar, D. St.Stoyanov, M. Aragona & A. Jablensky (Eds.), *Alternative perspectives on psychiatric validation*. (pp.27-46). Oxford: Oxford University Press.
- Becker, J. (1974). *Depression: Theory and research*. Chichester: John Wiley.
- Bentall, R.P. (2004). *Madness explained: Psychosis and human nature*. London: Penguin.
- Blashfield, R. K. (1982). Feighner et al., invisible colleges, and the Matthew effect. *Schizophrenia Bulletin*, 8, 1-6.
- Blashfield, R. K. (1984). *Classification of psychopathology*. New York: Plenum Press.
- Blashfield, R. K., & Draguns, J. G. (1976). Toward a taxonomy of psychopathology: The purposes of psychiatric classification. *British Journal of Psychiatry*, 129, 574-583.
- Bolton, D. (2008). *What is mental disorder?: An essay in philosophy, science, and values*. Oxford: Oxford University Press.
- Bolton, D. (2012). Classification and causal mechanisms: a deflationary approach to the classification problem. In K.S. Kendler & J. Parnas (Eds.), *Philosophical issues in psychiatry II: Nosology*. (pp. 6-11). Baltimore: John Hopkins University Press.
- Bridgman P. W. (1927). *The logic of modern physics*. New York: Macmillan.
- Brockington, I. F., Kendell, R. E., & Leff, J. P. (1978). Definitions of schizophrenia: concordance and prediction of outcome. *Psychological Medicine*, 8, 387-398.
- Brugha, T., Wing, L., Cooper, J. & Sartorius N. (2011) Contribution and legacy of John Wing, 1923-2010. *British Journal of Psychiatry*, 198, 176-178.

- Cassidy, W.L., Flanagan, N.B., Spellman, M. & Cohen, M.E. (1957). Clinical observations in manic-depressive disease: A quantitative study of one hundred manic-depressive patients and fifty medically sick controls. *Journal of the American Medical Association*, *164*, 1535-1546.
- Chung, M.C. (2007). Conceptions of schizophrenia. In M. C. Chung (Ed.) *Reconceiving schizophrenia*. (pp. 29-62) Oxford: Oxford University Press.
- Cohen, M., Badal, D., Kilpatrick, A. Reed, E., & White, P. (1951). The high familial prevalence of neurocirculatory asthenia (anxiety neurosis, effort syndrome). *American Journal of Human Genetics*, *3*, 126-158.
- Cole, J., McGuffin, P. & Farmer, A.E. (2008). The classification of depression: Are we still confused? *The British Journal of Psychiatry*, *192*, 83-85.
- Cooper, A.M. & Michels, R. (1981). Diagnostic and statistical manual of mental disorders. *American Journal of Psychiatry*, *138*, 128-129.
- Cooper, J.E. (2003). Remembering Robert Kendell. *The British Journal of Psychiatry*, *182*, 279-280.
- Cooper, J.E., Kendell, R.E., Gurland, B.J., Sharpe, L., Copeland, J. & Simon, R. (1972). *Psychiatric diagnosis in New York and London: A comparative study of mental hospital admissions*. London: Oxford University Press.
- Cooper, R. (2005). *Classifying madness*. Dordrecht: Springer.
- Cunningham, J. (1952). Problems of communication in scientific and professional disciplines. *The American Journal of Orthopsychiatry*, *22*, 445-456.
- Decker, H.S. (2013). *The making of DSM-III: A diagnostic manual's conquest of American psychiatry*. New York: Oxford University Press.

- Ellis, A. (1956). An operational reformulation of some of the basic principles of psychoanalysis. In H. Feigl & M. Scriven (Eds.) *The foundations of science and the concepts of psychology and psychoanalysis*. (pp.131-150). Minneapolis: University of Minnesota.
- Farmer, A. (1997). Current approaches to classification. In R. Murray, P. Hill & P. McGuffin (Eds.) *Essentials of postgraduate psychiatry*. (pp.49-64).Cambridge: Cambridge University Press.
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26, 57-63.
- Foreman, D.M. (1993). The Royal College of Psychiatrists Philosophy Group—a brief history of ideas. *Psychiatric Bulletin*, 17, 675-676.
- Frances, A., Pincus, H., Widiger, T., First, M., Davis, W., Hall, W., McKinney, K. & Stayna, H. (1994). DSM-IV and international communication in psychiatric diagnosis. In J. Mezzich, Y. Honda, and M. Kastrup (Eds.), *Psychiatric diagnosis* (pp.11-22). New York: Springer
- Fulford, K.W. M. (no date) K.W.M. (Bill) Fulford. Available at <http://www.stcatz.ox.ac.uk/node/378> [accessed 21 Feb 2018]
- Fulford, K. W. M. (2002). Values in psychiatric diagnosis: Executive summary of a report to the chair of the ICD-12/DSM-VI coordination task force (dateline 2010). *Psychopathology* 35, 132-138.

- Fulford, K.W.M., Sallah, D. & Woodbridge, K. (2007). Philosophical tools for cultural psychiatry. In K. Bhui & D. Bhugra (Eds.) *Culture and mental health: A comprehensive textbook* (pp.37-46). CRC Press: Boca Raton, FL
- Fulford, K.W.M. & Sartorius, N. (2009). The secret history of ICD and the hidden future of DSM. In M. Broome & L. Bortolotti (Eds.), *Psychiatry as cognitive neuroscience: Philosophical perspectives*. (pp. 29-47). New York: Oxford University Press.
- Green, C. (1992). Of immortal mythological beasts: Operationism in psychology. *Theory & Psychology*, 2, 291-320.
- Grob, G.N. (1991). *From asylum to community: Mental health policy in modern America*. Princeton: Princeton University Press
- Healy, D. (2002). Mandel Cohen and the origins of the ‘Diagnostic and Statistical Manual of Mental Disorders: DSM-III.’ *History of Psychiatry*, 13, 209-230.
- Hempel, C.G., (1954). A logical appraisal of operationism. *The Scientific Monthly*, 79(4), 215-220.
- Hempel, C. G. (1961). Introduction to problems of taxonomy. In J. Zubin (Ed.) *Field studies in the mental disorders*. (pp. 5-32). New York: Grune & Stratton.
- Hempel, C.G. (1965a). Fundamentals of taxonomy. In C.G. Hempel (Ed.) *Aspects of scientific explanation and other essays in the philosophy of science* (pp. 137-154). New York: Free Press.
- Hempel, C.G. (1965b). The theoretician’s dilemma: A study in the logic of theory construction. In C.G. Hempel (Ed.) *Aspects of scientific explanation and other essays in the philosophy of science* (pp. 173-226). New York: Free Press.

- Originally published 1958.
- Hempel, C.G. (1965c). The logic of functional analysis. In C.G. Hempel (Ed.) *Aspects of scientific explanation and other essays in the philosophy of science* (pp.297-330). New York: Free Press. Originally published 1959.
- Hempel, C.G. (1965d). Aspects of scientific explanation. In C.G. Hempel (Ed.) *Aspects of scientific explanation and other essays in the philosophy of science* (pp. 333-496). New York: Free Press.
- Hempel, C.G. & Oppenheim, P. (1936). *Der typusbegriff im lichte der neuen logik*. Leiden: Sitjhoff.
- Hollender, M. H., & Szasz, T. S. (1957). Normality, neurosis and psychosis: Some observations on the concepts of mental health and mental illness. *The Journal of Nervous and Mental Disease*, 125, 599-607.
- Jones, E. (2003) Aubrey Lewis, Edward Mapother and the Maudsley. In K. Angel, E. Jones & M. Neve (Eds). *European psychiatry on the eve of war: Aubrey Lewis, the Maudsley Hospital and the Rockefeller Foundation in the 1930s*. (pp.3-38). Medical History, Supplement 22. London: The Wellcome Trust Centre for the History of Medicine.
- Jones, T.D. (1944). The diagnosis of rheumatic fever. *Journal of the American Medical Association*, 126, 481-484.
- Kendell, R.E. (1968). *The classification of depressive illnesses*. London: Oxford University Press.
- Kendell, R.E. (1975). *The role of diagnosis in psychiatry*. London: Blackwell Scientific Publications.

- Kendell, R. E. (1982a). The choice of diagnostic criteria for biological research. *Archives of General Psychiatry*, 39, 1334-1339.
- Kendell, R.E. (1982b). Comments on Blashfield's article. *Schizophrenia Bulletin*, 8, 11-12.
- Kendell, R.E. (2000). Ten books. *The British Journal of Psychiatry*, 177, 185-188.
- Kendell, R. E., Brockington, I. F., & Leff, J. P. (1979). Prognostic implications of six alternative definitions of schizophrenia. *Archives of General Psychiatry*, 36, 25-31.
- Kendler, K.S., Muñoz, R.A., & Murphy, G. (2010). The development of the Feighner criteria: a historical perspective. *American Journal of Psychiatry*, 167, 134-142.
- Klerman, G.L. (1978). The evolution of a scientific nosology. In J.C. Shershow (Ed). *Schizophrenia: Science and practice*. (pp.99-116). Cambridge MA: Harvard University Press.
- Lewis, A. (2003) Aubrey Lewis's report of his visits to psychiatric centres in Europe in 1937. In K. Angel, E. Jones & M. Neve (Eds.). *European psychiatry on the eve of war: Aubrey Lewis, the Maudsley Hospital and the Rockefeller Foundation in the 1930s*. (Pp.64-147) Medical History, Supplement 22. London: The Wellcome Trust Centre for the History of Medicine. Report originally written 1938.
- MacMahon, B. & Pugh, T. (1970). *Epidemiology: Principles and methods*. Boston: Little, Brown and Company.
- McLeod, K. S. (2000). Our sense of Snow: the myth of John Snow in medical geography. *Social Science & Medicine*, 50, 923-935.
- Millon, T. (2012). On the history and future study of personality and its disorders. *Annual*

- Review of Clinical Psychology*, 8, 1-19.
- Mishara, A.L. & Schwartz, M.A., (2013). What does phenomenology contribute to the debate about DSM-5? In J. Paris & J. Phillips (Eds.), *Making the DSM-5* (pp. 125-142). Springer: New York.
- Monti, M. (1996) Review of *Philosophical Perspectives on Psychiatric Diagnostic Classification* by John Z. Sadler, Osborne P. Wiggins, Michael A. Schwartz. *History and Philosophy of the Life Sciences*, 18, 247-250
- Paykel, E.S. (1972). Correlates of a depressive typology. *Archives of General Psychiatry*, 27, 203-210.
- Purtell, J.J., Robins, E. & Cohen, M.E. (1951). Observations on clinical aspects of hysteria: A quantitative study of 50 hysteria patients and 156 control subjects. *Journal of the American Medical Association*, 146, 902-909.
- Rogers, T. B. (1989). Operationism in psychology: a discussion of contextual antecedents and an historical interpretation of its longevity. *Journal of the History of the Behavioral Sciences*, 25, 139-153.
- Ropes, M.W., Bennett, G.A., Cobb, S., Jacox, R. & Jessar, R.A. (1957). Proposed diagnostic criteria for rheumatoid arthritis: report of a study conducted by a committee of the American Rheumatism Association. *Annals of the Rheumatic Diseases*, 16, 118-125.
- Ropes, M.W., Bennett, G.A., Cobb, S., Jacox, R. & Jessar, R.A. (1958). 1958 Revision of diagnostic criteria for rheumatoid arthritis. *Bulletin on the Rheumatic Diseases*, 9, 175-176.
- Rutstein, D.D., Bauer, W., Dorfman, A., Gross, R.E., Lichty, J.A., Taussig, H.B.,

- Whittemore, R., Hagberg, K. & Parker, M.E. (1956). Jones' criteria (modified) for guidance in the diagnosis of rheumatic fever report of the Committee on Standards and Criteria for Programs of Care. *Circulation*, 13, 617-620.
- Rutter, M. (1965). Classification and categorization in child psychiatry. *Journal of Child Psychology and Psychiatry*, 6, 71-83.
- Sadler, J.Z., (2005). *Values and psychiatric diagnosis*. Oxford: Oxford University Press.
- Sadler, J.Z., Wiggins, O.P. & Schwartz, M.A. (Eds.) (1994). *Philosophical perspectives on psychiatric diagnostic classification*. Baltimore: Johns Hopkins University Press.
- Sartorius, N. (no date) Norman Sartorius. Available at <https://www.kcl.ac.uk/sspp/departments/sshm/people/Visiting-Professors/normansartorius.aspx> [accessed 21 Feb 2018]
- Schwartz, M.A. & Wiggins, O.P. (1986). Logical empiricism and psychiatric classification. *Comprehensive Psychiatry*, 27, 101-114.
- Schwartz, M.A. and Wiggins, O.P., (1987). Diagnosis and ideal types: a contribution to psychiatric classification. *Comprehensive Psychiatry*, 28, 277-291.
- Skinner, B.F. (1931). The concept of the reflex in the description of behaviour. *The Journal of General Psychology*, 5, 427-458.
- Spitzer, R.L. (1975) Review of *Glossary of Mental Disorders and Guide to Their Classification for Use in Conjunction with the International Classification of Diseases, 8th Revision*. *American Journal of Psychiatry*, 132, 881.
- Spitzer, R. (1977). Review of *The Role of Diagnosis in Psychiatry*. *Psychiatric Annals*, 7, 40-41.

- Spitzer, R.L. & Endicott, J. (1968). DIAGNO: A computer program for psychiatric diagnosis utilizing the differential diagnostic procedure. *Archives of General Psychiatry*, 18, 746-756.
- Spitzer, R.L., Endicott, J. & Cohen, G.M. (1964). *Manual of instruction for the Mental Status Schedule*. New York: Biometrics Research, New York State Department of Mental Hygiene.
- Spitzer, R.L., Endicott, J. & Robins, E. (1975a). Clinical criteria for psychiatric diagnosis and DSM-III. *American Journal of Psychiatry*, 132, 1187-1192.
- Spitzer, R.L., Endicott, J. & Robins, E. (1975b). *Research diagnostic criteria: Instrument no. 58*. New York: New York State Psychiatric Institute.
- Spitzer, R.L., Fleiss, J.L., Burdock, E.I. & Hardesty, A.S. (1964). The mental status schedule: rationale, reliability and validity. *Comprehensive Psychiatry*, 5, 384-395.
- Stein, D. J. (1991). Philosophy and the DSM-III. *Comprehensive psychiatry*, 32, 404-415.
- Stengel, R. (1958). Correspondence: Psychopathic personality. *British Medical Journal*. Feb 15 1958. p.402
- Stengel, E. (1959). Classification of mental disorders. *Bulletin of the World Health Organization*, 21, 601-663.
- Stengel, E. (1961). Discussion following Introduction to problems of taxonomy. In J. Zubin (Ed.) *Field studies in the mental disorders*. (pp. 23-26). New York: Grune & Stratton.
- Stevens, S. S. (1935a). The operational basis of psychology. *American Journal of Psychology*, 42, 323-330.

- Stevens, S. S. (1935b). The operational definition of psychological concepts. *Psychological Review*, 42, 517-527.
- Strauss, J. S. (1973). Diagnostic models and the nature of psychiatric disorder. *Archives of General Psychiatry*, 29, 445-449.
- Subcommittee on Classification of Mental Disorders (1968). *A glossary of mental disorders*. London: Her Majesty's Stationary Office.
- Szasz, T. S. (1959). The classification of 'mental illness'. *Psychiatric Quarterly*, 33, 77-101.
- Tolman, E.C. (1951). Operational behaviorism and current trends in psychology. Reprinted in E.C. Tolman (1951), *Behavior and psychological man* (pp.115-129). Berkeley: University of California Press [Original work published in 1936].
- Turner, M. A. (2003). Psychiatry and the human sciences. *The British Journal of Psychiatry*, 182, 472-474.
- Wakefield, J. C. (1999). Philosophy of science and the progressiveness of the DSM's theory-neutral nosology: response to Follette and Houts, part 1. *Behavior Research and Therapy*, 37, 963-999.
- Ward, C.H., Beck, A.T., Mendelson, M., Mock, J.E. and Erbaugh, J.K., 1962. The psychiatric nomenclature: Reasons for diagnostic disagreement. *Archives of General Psychiatry*, 7(3), pp.198-205.
- Wheeler, E.O., White, P.D., Reed, E.W. & Cohen, M.E., (1950). Neurocirculatory asthenia (anxiety neurosis, effort syndrome, neurasthenia): a twenty year follow-up study of one hundred and seventy-three patients. *Journal of the American Medical Association*, 142, 878-889.

- Wing, J.K. (1961). A simple and reliable subclassification of chronic schizophrenia. *The British Journal of Psychiatry*, 107(450), 862-875.
- Wing, J.K., Cooper, J.E. and Sartorius, N. (1974). *Present state examination*. London: Cambridge University Press.
- Wolman, B.B. (1978). Classification and diagnosis of mental disorders. In B.B. Wolman (Ed.), *Clinical diagnosis of mental disorders* (pp.15-45). New York: Plenum Press.
- World Health Organization (1960). *Epidemiology of mental disorders. Eighth report of the Expert Committee on Mental Health*. Geneva: World Health Organization.
- World Health Organization (1974). *Glossary of mental disorders and guide to their classification for use in conjunction with the International Classification of Diseases, 8th revision*. Geneva: World Health Organization.
- Zubin, J. (1961). (Ed.) *Field studies in the mental disorders*. New York: Grune & Stratton.
- Zubin, J. (1967). Classification of the behavior disorders. *Annual Review of Psychology*, 18, 373-406.
- Zubin, J. (1977). But is it good for science? *Clinical Psychologist*, 31, 5-7.

Supplementary materials

Further discussion of the quantitative analysis of citations to Hempel's paper.

Breakdown of citations to the two versions of Hempel's paper

Figure 1 (in the main paper) compared citations to both versions of Hempel's paper combined with citations to papers by Stengel and by Ward et al. In the main paper we combine citations to both editions of Hempel's paper, rather than considering citations to each edition separately, as we are interested in tracing the influence of his argument (which is identical in the two editions). For many purposes, one might have concerns about the fairness of comparing two editions of a paper published once in conference proceedings and once in a collection of philosophy essays, with papers published in medical journals. However, we think that such concerns do not threaten our primary argument. Our main aim is to show that Hempel's paper was originally not very influential, and is now cited much more than in the years leading up to the publication of DSM-III, and our methodology is sufficient to establish this claim.

We have also examined the breakdown of citations to the two versions of Hempel's paper (Fig.2).

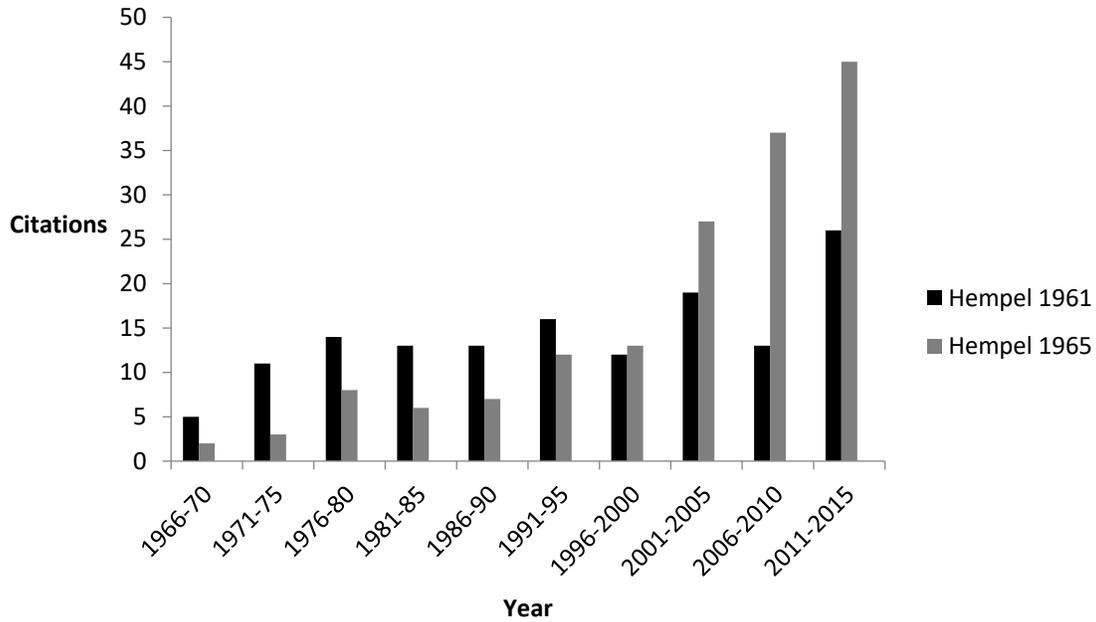


Fig.2 Citations to Hempel 1961 and 1965

The earlier (1961) edition, published in the conference proceedings, is the most frequently cited version up to 1996-2000, when the 1965 edition, published in a volume alongside other philosophical essays by Hempel, starts to become cited more frequently. We think it likely that this shift can be largely explained by changes in ease of access to the different editions, rather than, for example, suggesting that Hempel's paper came to seem more relevant to philosophers than to mental health researchers. Academic libraries tend to retain philosophical works much longer than they do medical works. For this reason, Hempel's 1965 collection of philosophy essays, *Aspects of Scientific Explanation*, is still readily available in most academic libraries, while the 1961 conference proceedings, *Field Studies in the Mental Disorders*, is now much harder to access.

On the possibility of quantitatively analysing citations to Hempel by type of literature.

One might wonder whether it would be possible to reveal more about the sorts of use that have been made of Hempel's paper through further quantitative analysis of citations.

Looking through citations to Hempel's paper in Google scholar, it is clear that many citations are in mental health journal papers, that many are in work in the philosophy of psychiatry, and that Hempel's paper is also cited in work across a range of other disciplines. We have decided not to attempt a quantitative analysis of the extent to which Hempel's paper is cited in different literatures. A number of methodological problems would plague any such attempt. First, it is not clear whether there is equal coverage across disciplinary areas (for example, of philosophy of psychiatry and mental health publications) in Google scholar. Second, citation practices differ between disciplines (papers in psychiatry and psychology tend to contain rather more passing references than do papers in philosophy), and it is unclear how this might be weighted. Third, in many cases it is unclear how a work might be categorised, for example, sometimes philosophically-oriented work is written by psychiatrists and published in psychiatry journals.

