In Defence of Psychiatric Diagnoses

Sam Fellowes

OPEN ACCESS

palgrave
macmillan

In Defence of Psychiatric Diagnoses

"*In Defence of Psychiatric Diagnoses* has the great virtues of being really clear and really interesting. Sam Fellowes lucidly sets out some fundamental issues in the philosophy of science and uses them to illuminate contested issues in psychiatric nosology. Philosophers, clinicians and mental health researchers can all learn a lot from this, and find new ways to talk to each other."

—Dominic Murphy, *Professor of Philosophy, University of Sydney, author of Psychiatry in the Scientific Image (MIT)*

"Sam Fellowes is one of the most innovative and sophisticated thinkers in the philosophy of psychiatry today. In *Defence of Psychiatric Diagnoses*, he develops a novel neo-Kantian approach to understanding psychiatric diagnosis and classification. Through focussing on the importance of idealisation in psychiatric science, Fellowes manages to explain how mental health diagnoses can be scientifically legitimate without merely mirroring reality."

—Rachel Cooper, *Professor of History and Philosophy of Science, Lancaster University. Author of Classifying Madness (Springer, 2005), Psychiatry and the Philosophy of Science (Acumen, 2007), Diagnosing the DSM (Karnac, 2014)*

"In this thought-provoking book, Sam Fellowes attempts to undermine many influential criticisms of contemporary psychiatric diagnoses by undermining the assumptions about the nature of science on which they rest. He defends, instead, a neo-Kantian view of science as a whole, which highlights the role of construction and idealisation rather than a simple mirroring of the natural world. The result is both a detailed scholarly defence of neo-Kantian philosophy of science and a timely intervention in discussions of psychiatric classification."

—Tim Thornton, *Emeritus Professor of Philosophy and Mental Health, University of Central Lancashire, author of Essential Philosophy of Psychiatry (OUP)*

Sam Fellowes

In Defence of Psychiatric Diagnoses

palgrave
macmillan

Sam Fellowes
Politics, Philosophy & Religion
University of Lancaster
Lancaster, UK

This work was supported by Lancaster University.

ISBN 978-3-031-74477-8 ISBN 978-3-031-74478-5 (eBook)
https://doi.org/10.1007/978-3-031-74478-5

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

For my brother Simon Fellowes, a schizophrenic individual who died tragically early aged forty-seven

Preface

The origin of this book can be traced back to me being diagnosed as autistic half way through my part-time Masters in History and Philosophy of Science. After finishing the Masters I spent a lot of time reflecting upon who I am, my past and my future, in light of being diagnosed as autistic. Additionally, I spent a lot of time around autistic people whom I met at support groups, on internet forums for autistic people and sometimes in person outside of those. I was struck by two seemingly opposing notions. On one hand, the diagnosis of autism seemed an immensely powerful tool for understanding myself and other autistic people. I interpreted behaviour in a new way and made links between behaviour which I previously had not. On the other hand, the differences between autistic people often seemed very high. I tried to work out how notions of autism could increase understanding despite those differences.

My thinking about this was influenced by three philosophies of science that were heavily on my mind during and following my Masters. Firstly, that all observations are theory-laden whereby what we observe can be heavily influenced by theory. I wondered if autism could act as a theory which influences what we observe. Secondly, that science often involves models that only imperfectly describe what actually occurs in the world. I wondered if autism might be an idealised and abstract model which only imperfectly describes actual people. Thirdly, structural realism whereby science relates to describing mathematical structures rather than

entities. I wondered if autism could be understood as a mathematical structure rather than an entity. Eventually, part of the way through my PhD I encountered neo-Kantianism and all these different notions merged together into one cohesive framework. Rather than seeing science as being about describing mind-independent entities, science is about applying something that comes from us to produce idealised and abstract models that describe probabilistic and statistical relationships. I hope that through this framework I can help understand how the diagnosis of autism can provide genuine knowledge despite the heterogeneity of autistic people.

This book has been influenced in at least two ways by me being an autistic person. Firstly, much has been written recently about how people with psychiatric diagnoses have unique insight into their diagnosis through lived experience. Whilst I never directly appeal to my lived experience, my thoughts about how autism does or does not relate to the world will have been influenced by the way in which I think parts of myself are or are not instances of autism. However, I think the move from lived experience to knowledge is quite epistemologically complicated whereby lived experience is theory-laden and needs interpreting through concepts. It is certainly possible that my lived experience has led to insights that otherwise would be difficult to reach but I urge readers to retain in mind that I have not shown any interpretations influenced by my lived experience are epistemologically secure. There is a possibility that my lived experience has not improved or has even reduced the quality of this book. Secondly, as an autistic individual I have a preference to focus on detail but this can sometimes come at the cost of seeing the bigger picture. I have actively tried to keep that wider perspective in mind and remember that more detail is not always preferable, but I suspect I have had limited success in this endeavour. My preference for detail may have many positives but I suspect I will not have appreciated any limitations imposed by my difficulties with seeing wider contexts.

I would like to thank multiple people. I would like to thank Rachel Cooper, Peter Zachar, Awais Aftab, Lucy James, Justin Garson, George Turner, Annika Dhalin, Jane Fellowes, Hane Maung, Alison Stone, Gareth Williams, Jacob Barlow and Brian Garvey who either read parts of this book or had helpful conversations with about this book. This book

was largely written whilst I was a Wellcome Trust funded Research Fellow [209868/Z/17/Z]. I would like to thank the Wellcome Trust for their support. Also, I would like to thank those who supported me despite knowledge of my past mental health problems and legitimate concerns over future mental health problems. You hopefully know who you are.

Competing Interests I declare no competing interests.

Lancaster, UK Sam Fellowes

Contents

1

Introduction

There is a long history of scepticism over psychiatric diagnoses and these concerns are still very much present today. These concerns have come from many different sources, such as psychiatrists, philosophers, psychotherapists, sociologists, historians and diagnosed individuals. Psychiatric diagnoses have been described as "arbitrarily constructed" (Burrows 2010, p. 252), "constructed" (Gains 1992, p. 4; Summerfield 2001, p. 95), "arbitrary" (Cushing 2013, p. 38; Horwitz 2002, p. 5), "invention" (Read 2004, p. 21; Summerfield 2001, p. 95) and "made-up" (Watson 2019, p. 2). In this book I will argue that such concerns are largely misplaced. They seem to be driven by an explicit or implicit view of how science relates to reality which I will argue is untenable. I will draw upon an alternative way to think of science. That alternative understanding of science will leave psychiatric diagnoses looking much more respectable and legitimate compared to how their critics portray them. In this chapter I will outline how critics view psychiatric diagnoses and suggest that alternative approaches to understanding reality and science can be held. In later chapters I will argue for those alternative approaches and show how they strengthen the status of psychiatric diagnoses.

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_1

The type of psychiatric diagnosis I primarily focus upon is categorical polythetic psychiatric diagnoses. These are the main types of psychiatric diagnoses in the main psychiatric diagnostic manuals, the DSM-5-TR and ICD-11. The DSM [Diagnostics and Statistical Manual] is a diagnostic manual published by the APA [American Psychiatric Association]. It provides the diagnostic criteria for 298 psychiatric diagnoses. The ICD [International Classification of Disease] is a diagnostic manual published by the WHO [World Health Organisation]. It largely lists the same psychiatric diagnoses (though, unlike the DSM, it also lists non-psychiatric diagnoses). Later in this book I will discuss alternative types of psychiatric diagnoses but for now it is sufficient to know that I am discussing stereotypical DSM and ICD style psychiatric diagnoses. The vast majority of psychiatric diagnoses in these manuals are categorical polythetic psychiatric diagnoses. Examples would be autism, depression and schizophrenia.

The main reason why psychiatric diagnoses are seen as scientifically illegitimate is that they are considered not to be real. I now outline three motives for thinking this.

Firstly, reality is thought of in terms of causation and entities. Real causes are taken to produce real entities. What makes two things instances of the *same* type of entity (say, two apples) is that they both have the same underlying causes (the biological structure of the apple). What makes two things instances of *different* types of entities (say, an apple and a banana) is that they both have different causes. If scientists posit a type of thing (say, autism) but instances of that entity (people diagnosed with autism) have much causal diversity then autism is not a real entity. In this situation, what groups people together would not be real causes but rather mere opinion. As it happens, the vast majority of DSM and ICD psychiatric diagnoses appear to cover individuals who are causally heterogeneous (Cuthbert and Insel 2013, p. 3; Kozak and Cuthbert 2016, p. 287; Sanislow et al. 2010, p. 632). Some critics take this to mean psychiatric diagnoses are the product of ideology rather than science. For example, the psychiatric diagnosis of autism is highly causally heterogeneous (Devlin and Scherer 2012, p. 233; Weiskopf 2017, p. 179). This causal diversity (Timimi et al. 2011, p. 139) means "the field of autism rests on ideological assumptions, *not* scientific evidence" (Timimi et al. 2011, p. 4, emphasis original; see also Cushing 2013, p. 38). Instead, "the

autism spectrum has become a catch-all metaphor for focusing on a disparate range of behaviours that suggest a lack of the type of social and emotional competences thought to be necessary for the functioning of society dominated by neo-liberal economic and political foundations" (Timimi et al. 2011, p. 7). Similar claims are made against psychiatric diagnoses more generally. Kirk et al. write that "[s]ince the DSM is unable to document scientifically that these diverse behaviors represent any medical or mental disease, its [the DSM's] endurance may be reasonably attributed to other powerful social purposes" (2015, p. 67). Other critics would not see causally heterogeneous psychiatric diagnoses as the product of ideology. Rather, they claim that such psychiatric diagnoses are scientifically inadequate and need replacing with superior psychiatric diagnoses which are not causally heterogeneous (Cooper 2005, p. 150; Murphy 2006, p. 11; Cuthbert and Insel 2013, p. 4; Poland 2014, p. 34).

Secondly, science is taken to describe actual entities. The world is taken to consist of actual entities and any scientific theory which inadequately describes an actual entity is not scientifically legitimate. The world is there to be discovered and either we discover it or we do not. What we should not do is *decide* what is in the world. Scientists try to *discover* what is in the world, they do not get to decide that, say, atoms exist. However, there is a well-documented process of psychiatrists deciding which psychiatric diagnoses should make up the diagnostic manuals (Kendlar et al. 2011, p. 1149; Kincaid 2014, p. 151; Jablensky 2008, p. 90). Psychiatrists debate which psychiatric diagnoses should be included in the DSM and ICD, what their diagnostic criteria should be, if they have subtypes or are on a spectrum. Psychiatric diagnoses have been described as "unscientific, made-up constructs voted into existence by the 'opinions' of psychiatrists, populating the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* committees" (Watson 2019, p. 2, emphasis original; see also Raskin and Lewandowski 2000, p. 16; Cushing 2013, p. 38; Horwitz 2002, p. 5; Read 2004, p. 21). If psychiatric diagnoses cannot be discovered but must instead result from decisions then they are not real or legitimate science.

Thirdly, science is taken as actually describing the world rather than missing out important aspects. Any scientific theory which attempts to describe part of the world but misses out many significant details would

fail to describe reality and would not be scientifically legitimate. However, psychiatric diagnoses typically only provide a very limited description of particular people. Knowing that someone has a psychiatric diagnosis means that you know they have enough symptoms to meet the diagnostic criteria but this does not reveal which specific symptoms of that psychiatric diagnosis they have and does not reveal other aspects of the individual such as their personality, their life history, their life goals or their social and economic situation (Johnstone 2018, p. 33; Kinderman et al. 2013, p. 2; Pietikainen 2015, p. 323). Some critics take this to mean that psychiatric diagnoses are illegitimate. Kinderman et al. writes that

> two people with a diagnosis of 'schizophrenia' or 'personality disorder' may possess no two symptoms in common, [so] it is difficult to see what communicative benefit is served by using these diagnoses. Surely a description of a person's real problems would suffice? A description of an individual's actual problems would provide more information and be of greater communicative value than a diagnostic label. (2013, p. 3; see also Johnstone 2018, p. 39; Kinderman et al. 2013, p. 3; Runswick-Cole 2016, p. 27; Hassall 2016, p. 51; Timimi et al. 2011, p. 1; Vanheule 2017, p. 85)

Psychiatric diagnoses should not be considered legitimate science if they fail to accurately describe actual people.

In this book I shall respond to these concerns. I broadly accept some of the premises behind them. The vast majority of psychiatric diagnoses do indeed have a heterogeneous causal basis, decisions do need to be made over which psychiatric diagnoses are employed, and they do not describe significant aspects of particular individuals. Despite this, I shall argue that psychiatric diagnoses can be seen as being scientifically legitimate. Whether this will be the case with any particular psychiatric diagnosis will be case specific. I will argue it is possible for psychiatric diagnoses to be scientifically legitimate and it is likely that some currently employed psychiatric diagnoses are scientifically legitimate.

I will respond to these concerns by outlining an alternative account of how we should think of reality and how we should think of scientific legitimacy. I suggest these critics are employing an untenable account of science. I suggest that critics of psychiatric diagnoses often take a view of

science as being about describing reality *as it is in itself*. In contrast, I will outline a view of science as being about describing reality *as we can experience it*. I outline this alternative account of science by employing neo-Kantianism. We cannot see reality as it is but only instead as it appears to individuals like us. On this understanding we need to reconsider how we think about reality and legitimate science. Science should not be judged upon whether it describes some set of mind-independent entities in the external world. What counts as science depends upon concepts and principles we apply to the world rather than our ideas reflecting the mind-independent external world. I will argue that this alternative view of science means that psychiatric diagnoses can be seen in a much more positive light.

I will use this neo-Kantian position to address the claim that psychiatric diagnoses do not describe real people or biomedical entities. On this neo-Kantian account, the aim of science is not to simply describe or categorise what we see in the world around us. Rather, science "continually go[es] beyond the given" (Cassirer 1953 [1910], p. 127; see also Natorp 2004 [1921], p. 202; Rickert 1986 [1902], p. 217). Science takes what we observe and transforms it into something that does not resemble what we observe. As such, that psychiatric diagnoses do not conform to anything that we observe around us, such as actual people, does not compromise their scientific legitimacy. Additionally, on this neo-Kantian account, the aim of science is not to describe mind-independent entities. It takes such entities to be in some sense inaccessible to us. This is because we cannot know about mind-independent entities except through applying something to them which they do not themselves have. Our knowledge of mind-independent entities is mediated by the concepts and principles we apply to them. As such, we cannot see mind-independent objects as they are but only as they appear to us. Consequently, we should not judge psychiatric diagnoses upon whether they conform to mind-independent entities, be those mind-independent entities understood as actual people or biomedical entities.

Following this, we need an alternative standard for judging the scientific legitimacy of psychiatric diagnoses which is not based upon describing real people or biomedical entities. The philosophy of science I outline is one where we apply concepts and principles to sensation to build

idealised and abstract models that do not conform to what we see around us or to mind-independent entities. As such, we should see psychiatric diagnoses as idealised and abstract models. Our standard for judging the scientific legitimacy of psychiatric diagnoses needs instead to focus upon what concepts and principles we employ when building those idealised and abstract models. Neo-Kantianism provides standards for making this judgement. Neo-Kantianism aims to avoid relativism by providing an alternative basis for thinking about objectivity in science. Different neo-Kantians have supplied different standards for assessing science. I shall largely draw upon the notion that we can derive those concepts and principles from reason. This then gives those concepts and principles a level of objectivity. As such, we can still potentially think of psychiatric diagnoses as anchored in objectivity even if they do not describe real people or mind-independent entities. Not all currently employed psychiatric diagnoses will meet those criteria and so not all currently employed psychiatric diagnoses should be considered scientifically legitimate. Additionally, I will outline how alternative, not currently employed psychiatric diagnoses, could be formulated from those concepts and principles. They do not simply entail one set of psychiatric diagnoses. This gives us a level of choice over which psychiatric diagnoses to employ although not all choices will result in scientifically legitimate psychiatric diagnoses. A neo-Kantian approach will thus provide an important defence of psychiatric diagnoses in principle and of some currently employed psychiatric diagnoses in particular.

There are two alternative approaches to categorical psychiatric diagnoses which might intuitively seem better at describing the reality of people. The first is dimensional approaches. This is where the presence of traits is measured throughout the population rather than just within people who meet the DSM diagnostic criteria. Given that the symptoms and causes associated with a particular DSM diagnosis can occur in people who do not meet diagnostic criteria, it might seem that dimensional approaches are better able to detect how symptoms and causes manifest in reality. The second is person centred approaches. The aim with these is to understand an actual person who is engaging in therapy, to understand them as a unique individual. This contrasts with categorical and dimensional diagnoses where the aim is to group people together despite those people

having unique differences. It might seem that person centred approaches are better able to describe the reality of actual people, describing the symptoms and causes and other attributes present in specific people. In this book I will argue that neither dimensional nor person centred approaches can better access reality than categorical approaches can. Rather, I argue that categorical, dimensional and person centred approaches are ways of making knowledge possible. They are effectively measurement systems which give data form rather than being descriptions of how people actually are. In this sense I incorporate all three under the neo-Kantian philosophy I endorse. My goal in this book is to argue that categorical psychiatric diagnoses can be scientifically legitimate and can make possible useful data but my goal is not to argue against the use of dimensional and person centred approaches. Rather, all three are scientifically legitimate, all three can make possible different knowledge and all three should be used.

There are four aims to this book. Firstly, in response to the above concerns, I show why DSM style psychiatric diagnoses can be scientifically legitimate. This is not to suggest they are good science, merely that they are the sort of things that scientists can legitimately posit. Secondly, I give an indication of what a good DSM style psychiatric diagnosis would look like, giving the example of autism as a relatively good diagnosis though I suggest it could be improved by adding subtypes. Thirdly, I respond to calls to reform rather than abandon the DSM. Typically, these are calls to make the DSM dimensional. I consider whether such a reformulated DSM should be entirely dimensional or should also still employ categories, and I suggest categorical diagnoses should still be employed. Finally, I consider scientific projects like RDoC and HiTOP that wish to completely dispense with the DSM. Whilst it is difficult to know what psychiatric diagnoses they will eventually suggest or be used to construct, I suggest their aim of finding relatively homogeneous dimensional groupings can plausibly be supplemented by adding categories on to those dimensional groupings.

References

Burrows, Kathryn. 2010. What Epidemic? The Social Construction of Bipolar Epidemics. *Advances in Medical Sociology* 11: 243–261. https://doi.org/10.1108/S1057-6290(2010)0000011017.

Cassirer, Ernst. 1953. *Substance and Function and Einstein's Theory of Relativity*. Dover Publications.

Cooper, Rachel. 2005. *Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders*. Dordrecht: Springer.

Cushing, Simon. 2013. Autism: The Very Idea. In *The Philosophy of Autism*, ed. Jami Anderson and Simon Cushing, 17–46. Plymouth: Rowman & Littlefield.

Cuthbert, Bruce N., and Thomas R. Insel. 2013. Towards the Future of Psychiatric Diagnosis: The Seven Pillars of RDoC. *BMC Medicine* 11 (126): 1–8. https://doi.org/10.1186/1741-7015-11-126.

Devlin, Bernie, and Stephen W. Scherer. 2012. Genetic Architecture in Autism Spectrum Disorder. *Current Opinion in Genetics & Development* 22: 229–237. https://doi.org/10.1016/j.gde.2012.03.002.

Gains, Atwood D. 1992. From DSM-I to III-R; Voices of Self, Mastery and the Other: A Cultural Constructivist Reading of U.S. Psychiatric Classifications. *Social Sciences and Medicine* 35 (1): 3–24. https://doi.org/10.1016/0277-9536(92)90115-7.

Hassall, Richard. 2016. Understanding This *Thing* Called Autism. In *Re-Thinking Autism*, ed. Katherine Runswick-Cole, Rebecca Mallett, and Sami Timimi, 49–66. London: Jessica Kingsley Publishing.

Horwitz, Allan V. 2002. *Creating Mental Illness*. University of Chicago Press.

Jablensky, Assen. 2008. The Nosological Entity in Psychiatry: A Historical Illusion Or Moving Target. In *Philosophical Issues in Psychiatry II: Nosology*, ed. Kenneth S. Kendler and Josef Parnas, 77–94. Oxford: Oxford University Press.

Johnstone, Lucy. 2018. Psychological Formulation as an Alternative to Psychiatric Diagnosis. *Journal of Humanistic Psychology 58* (1): 30–46. https://doi.org/10.1177/0022167817722230.

Kendlar, Kenneth S., Peter Zachar, and Carl Craver. 2011. What Kinds of Things Are Psychiatric Disorders? *Psychological Medicine* 41: 1143–1150. https://doi.org/10.1017/S0033291710001844.

Kincaid, Harold. 2014. Defensible Natural Kinds in the Study of Psychopathology. In *Classifying Psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 145–174. Cambridge, MA: MIT Press.

Kinderman, Peter, John Reed, Joanna Moncrieff, and Richard P. Bentall. 2013. Drop the Language of Disorder. *Evidence Based Mental Health* 16 (1): 1–3. https://doi.org/10.1136/eb-2012-100987.

Kirk, Stuart A., David Cohen, and Tomi Gomory. 2015. DSM-5: The Delayed Demise of Descriptive Diagnosis. In *The DSM-5 in Perspective: Philosophical Reflections on the Psychiatric Babel*, 63–81. Dordrecht, The Netherlands: Springer.

Kozak, Michael J., and Bruce N. Cuthbert. 2016. The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. *Psychophysiology* 53: 286–297. https://doi.org/10.1111/psyp.12518.

Murphy, Dominic. 2006. *Psychiatry in the Scientific Image*. Cambridge, MA: Massachusetts Institute of Technology.

Natorp, Paul. 2004. *Plato's Theory of Ideas*. An Introduction to Idealism: Academia Verlag Sankt Augustin.

Pietikainen, Petteri. 2015. *Madness: A History*. London: Routledge.

Poland, Jeffrey. 2014. Deeply Rooted Sources of Error and Bias in Psychiatric Classifications. In *Classifying Psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 29–64. Cambridge, MA: MIT Press.

Raskin, Jonathan D., and Adam M. Lewandowski. 2000. The Construction of Disorder as Human Enterprise. In *Constructions of Disorder: Meaning Making Framework for Psychotherapy*, ed. Robert A. Neimeyer and Jonathan D. Raskin, 15–40. American Psychological Association.

Read, John. 2004. The Invention of Schizophrenia. In *Models of Madness, Psychological, Social and Biological Approaches to Schizophrenia*, ed. John Read, Loren Mosher, and Richard P. Bentall, 21–34. Hove: Brunner-Routledge.

Rickert, Heinrich. 1986. *The Limits of Concept Formation in Natural Science: A Logical Introduction to the Historical Sciences*. Cambridge University Press.

Runswick-Cole, Katherine. 2016. Understanding This *Thing* Called Autism. In *Re-Thinking Autism*, ed. Katherine Runswick-Cole, Rebecca Mallett, and Sami Timimi, 19–29. London: Jessica Kingsley Publishing.

Sanislow, Charles A., Daniel S. Pine, Kevin J. Quinn, Kozak J. Michael, Majorie A. Garvey, Robert K. Heinssen, Philip Sung-EN Wang, and Bruce N. Cuthbert. 2010. Developing Constructs for Psychopathology Research: Research Domain Criteria. *Journal of Abnormal Psychology* 199 (4): 631–639. https://doi.org/10.1037/a0020909.

Summerfield, Derek. 2001. The Invention of Post-traumatic Stress Disorder and Social Usefulness of a Psychiatric Category. *British Journal of Medicine* 332: 95–98. https://doi.org/10.1136/bmj.322.7278.95.

Timimi, Sami, Neil Gardner, and Brian McCabe. 2011. *The Myth of Autism*. Palgrave Macmillan.

Vanheule, Stijn. 2017. *Psychiatric Diagnosis Revisited: From DSM to Clinical Case Formulation*. Cham: Palgrave Macmillan.

Watson, Jo. 2019. Introduction. In *Drop the Disorder: Challenging the Culture of Psychiatric Diagnosis*, ed. Jo Watson, 1–7. Monmouth: PCCS Books.

Weiskopf, Daniel A. 2017. An Ideal Disorder? Autism as a Psychiatric Kind. *Philosophical Explorations* 20 (2): 175–190.

2

Psychiatry and the Philosophy of Psychiatry

2.1 Introduction

This chapter outlines psychiatry and philosophy of psychiatry to provide a basis for my later analysis. The first half of this chapter outlines important issues in psychiatry relating to psychiatric diagnoses and the second half outlines important issues in philosophy of psychiatry relating to psychiatric diagnoses. The purpose is to give an overview of where we are and how we got here. In Chap. 3 I will outline neo-Kantianism to provide an alternative philosophical position. This will then cast a new understanding on many of the issues discussed in this chapter when I apply a neo-Kantian approach to psychiatry in later chapters.

In this chapter I primarily outline how psychiatrists and philosophers have differed over whether psychiatry should be understood to involve disease entities. A disease entity is typically understood to be a naturally forming part of the world. A disease entity is where real causes produce a real entity with real characteristics. Which causes and characteristics the entity actually has is determined by the world rather than by psychiatrists. Consequently, disease entities are usually understood as real and scientifically legitimate.

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_2

In this chapter I will outline how psychiatry has oscillated between disease entity and non-disease entity approaches in the twentieth century and into the twenty-first century. Disease entity approaches are currently recognised to face major challenges and this has led to a crisis of confidence in current psychiatry and has led to the development of new approaches which seem to reject disease entities. However, it is to date unclear what psychiatric diagnoses would look like in a future psychiatry which rejected disease entities. Similarly, philosophers of psychiatry have oscillated in how strongly they adhere to a disease entity approach. The most common way in which philosophers of psychiatry have understood psychiatric diagnoses is as natural kinds. Paradigmatic examples of natural kinds are biological species, like tigers, and chemical elements, like oxygen. In its full-blooded form, natural kinds are understood as naturally forming parts of the world, which have associated characteristics which arise due to underlying causes. This understanding of natural kinds seems highly compatible with notions of disease entities. However, philosophers of psychiatry typically appeal to weaker notions of natural kinds which less resemble disease entities. This raises questions about whether they would be real or scientifically legitimate.

In this book I will draw upon neo-Kantianism to address many of these issues. I will outline an alternative approach which does not involve disease entities but which means psychiatric diagnoses can be considered scientifically legitimate. I will partly do this by rejecting a notion of science as describing *how things are* and instead appeal to neo-Kantianism to see science as describing *how things appear to us and are structured by us*. I suggest that notions of disease entities are untenable because they are based on notions of science as describing *how things are*.

I start this chapter by describing the history of psychiatry. Then I outline alternative approaches to psychiatry, outlining categorical, dimensional and person centred approaches. Finally, I outline philosophy of psychiatry.

2.2 The History of Psychiatry

I start this chapter by giving a selective overview of the history of thinking about psychiatric diagnoses in the twentieth and twenty-first century. I describe influential trends which held opposing views on the status of psychiatric diagnoses, presenting them as forerunners of choices which psychiatry currently faces. I aim to draw out two opposing views of psychiatric diagnoses, namely, a disease entity and a non-disease entity understanding of psychiatric diagnoses.

This history will be selective, describing and contrasting significant approaches which have strong parallels to today but without aiming to give a complete history. This history will have limited geographical focus. It will start in Europe and then move to the United States. Additionally, the content relating to Europe will mainly be selected because it is relevant for understanding the history of the United States. The motive for this is twofold. Firstly, my aim is to give the context behind the development of the modern DSM. As such, I selected the history prior to the DSM-III to give context to what the authors of the DSM-III were responding to. Secondly, my work as a historian has focused primarily on the United States, having extensively studied notions of autism and childhood schizophrenia employed between the 1930s to 1970s in the United States (see Fellowes 2024).

2.2.1 Kraepelin

Emil Kraepelin is often considered an early forerunner for the modern style of psychiatric diagnoses. He wrote an influential textbook named *Psychiatrie*. Across eight editions of this textbook (1st edition 1883, 8th edition 1927) he developed a new system of psychiatric diagnoses. Kraepelin was dissatisfied with the imprecision of how most of his contemporaries approached diagnosing patients (Decker 2007, p. 338). He believed the clinical picture (the symptoms exhibited) and the course of the illness (how the symptoms changed over time) of patients diagnosed as psychotic varied significantly. Kraepelin separated some of those patients diagnosed as psychotic into a dementia praecox group (a

diagnosis with significant similarities to the modern diagnosis of schizophrenia) and separated other patients into a manic-depressive group.

The diagnoses Kraepelin formulated were influenced by three beliefs. Firstly, he believed in the importance of using careful observations of patient behaviour as a basis for formulating psychiatric diagnoses. He valued careful observations more than his contemporary psychiatrists did (Berrios and Hauser 1988, p. 815; Decker 2007, p. 338; Jaspers 1997, p. 571). Secondly, he believed psychiatrists should observe rather than interpret the behaviour of patients and should not speculate about the causal origins of behaviour. This contrasts with some of his contemporaries who employed explicit theorising when observing patients and formulating diagnosis (Decker 2007, p. 340). Thirdly, he believed there was a clear distinction between mental illness and mental health. Correctly diagnosed individuals were ill, whereas most people were mentally healthy, and no one was only partly ill (Decker 2007, p. 340).

Kraepelin's approach is commonly understood as being driven by the underlying belief that mental disorder existed in the form of disease entities (Ghaemi 2009, p. 7; Heckers et al. 2021, p. 328; van Praag 2008, p. 32).[1] A disease entity is typically understood as a specific clinical picture with a specific underlying cause. Jaspers described Kraepelin's account of disease entities as follows:

> Clinical pictures of diseases that have *similar causes, a similar basic psychological form, similar development and course, similar outcome and a similar cerebral pathology* and which therefore all present the same over-all picture, are genuine, natural disease entities. (1997, p. 566 emphasis original)

Similarly, Boorse writes that "the strictest definition of a disease entity would be a constellation of signs, symptoms, and pathology with specific etiology and prognosis" (1977, p. 552). Kraepelin followed this general idea when he believed that sufficiently nuanced clinical observations would result in groupings of people with the same underlying disease. Kraepelin hoped his group of dementia praecox patients exhibited a particular set of symptoms due to a specific underlying cause, whilst

[1] Berrios and Hauser (1988) question the adequacy of this popular account.

manic-depressive patients exhibited a different particular set of symptoms due to a different specific cause (even though he felt psychiatrists should not speculate about what those causes were). Though using different terminology, Kraepelin is commonly understood to have endorsed a disease entity approach for major portions of his career. Kraepelin would later in life grow more doubtful about some of his claims (Berrios and Hauser 1988, p. 841). He found some patients had elements of both dementia praecox and manic-depression, and he doubted if there was a sharp separation between mental illness and mental health (Decker 2007, p. 341).

2.2.2 Psychoanalysis

A contrasting system to Kraepelin's arose with the development of psychoanalysis. This certainly did not completely replace Kraepelin's approach but it did produce a significantly divergent approach which had varying levels of dominance in different nations. Additionally, some thinkers were simultaneously influenced by both Kraepelin's approach and psychoanalysis, and tried to implement both approaches with varying degrees of success (Bleuler 1950[1911], p. 390). Austrian physician Sigmund Freud developed a theoretical system for analysing the unconscious drives behind human thoughts and behaviour. He initially focused upon the development of unconscious sexual urges in early childhood and later focused upon how society repressed unconscious urges. Freud primarily saw individuals with relatively minor problems rather than patients of mental hospitals. Consequently, his own views on mental illness are less relevant for my purposes than the views of later psychiatrists who employed his views. His ideas initially had a mixed reception within central Europe but gained in popularity during the interwar period. Of especial interest for understanding the context that the DSM was developed in is how some of Freud's fellow Jewish physicians, psychologists and psychiatrists fled the threatening situation in Europe. They took Freud's ideas, and their own development of Freud's ideas, with them. Psychoanalysis grew significantly in popularity around the early 1940s

and would become the dominant approach to psychiatry in the United States during the 1950s and 1960s (Shorter 1997, p. 170).

Psychoanalysts were not a monolithic group. Individual psychiatrists subscribed to psychoanalysis in varying degrees and psychoanalysts often held divergent views. However, some general trends can be stated. Psychoanalysts saw patients as undergoing a dynamic process. Whether and how symptoms manifested depended upon the stage of the dynamic process within the patient. This process was often believed to be caused by very specific events within an individual's life history. This meant that how the process developed was considered unique to each individual. This focus on individualised dynamic processes had three consequences. Firstly, psychoanalysts expected that the changing dynamic process would result in specific patients manifesting different symptoms at different points in their life. Consequently, psychoanalysts often felt that precisely defined psychiatric diagnoses were too narrow to account for the changing symptoms of individuals. Secondly, psychoanalysts often placed great importance upon theory. They typically categorised individuals based upon the dynamic processes that they thought were present rather than the symptoms they exhibited. Additionally, many psychoanalysts thought that the dynamic processes were the result of many different causal factors rather than a single specific causal factor. Thirdly, psychoanalysts typically rejected a sharp cut off point between illness and normality. They believed some individuals were heavily impaired but also felt that unconscious drives were present in the general population and these could result in minor or moderate neurosis.

This meant that for psychoanalysts "[g]one was any sense of well-defined disease entities" (Shorter 1997, p. 178; see also Wakefield 2022, p. 6). By expecting symptoms to present in a dynamic fashion, by believing multiple causal factors could produce symptoms and by rejecting a sharp demarcation between illness and normality, psychoanalysts typically rejected Kraepelinian notions of disease entities. Psychoanalysts typically primarily focused upon investigating the specific life story of the individual patient rather than primarily focusing upon placing individual patients into precise diagnostic categories. As Shorter comments, psychoanalysts held "contempt for the precise diagnosis of psychiatric illness" (Shorter 1997, p. 178). Based upon my own reading of psychiatrists

during this period, I think it should be emphasised that many psychoanalysts were not anti-psychiatric diagnoses (I show this in relation to figures like Kanner, Bender and Mahler in Fellowes 2024). They just did not think psychiatric diagnoses should be made very precise and that many important things were going on with their patients that cannot be adequately captured by psychiatric diagnoses. Whilst many psychoanalysts still employed psychiatric diagnoses, those primarily influenced by Kraepelin and those primarily influenced by psychoanalysis held opposing views about what constituted a good psychiatric diagnosis and how psychiatric diagnoses related to patients.

Psychoanalysis was far from entirely dominant during this era (especially outside the United States) but it did have a major impact on how people thought about psychiatric diagnoses, be it a complete rejection of precise diagnoses or just placing less value upon precise diagnoses compared to Kraepelin or what would come after psychoanalysis. However, from around the 1950s onwards, there were individuals who were concerned that psychiatric diagnoses were giving far too imprecise descriptions and that greater rigour was needed (Aragona 2014, p. 37). The scientific status of psychoanalysis was increasingly being questioned. Multiple individuals or groups held this concern but the most influential became known as the neo-Kraepelinians. Their work would usher in a new era which would emphasise disease entities and largely eclipse the focus on non-disease entities emphasised by psychoanalysis.

2.2.3 Neo-Kraepelinians

In the 1970s a group named the neo-Kraepelinians (otherwise known as the St. Louis group or the Feighner group) outlined an alternative approach to psychoanalytical approaches. These were psychiatrists based at Washington University in St. Louis, Missouri. They were concerned that what was schizophrenia to one clinician was manic-depression to another clinician. Also, they were unconvinced by the theoretical claims of psychoanalysis and dissatisfied with the lack of findings about biological causes. They believed improving reliability and validity was required

to remedy the problems they perceived with psychiatry. I will now outline reliability and validity.

Reliability is the measure of agreement when different psychiatrists diagnose the same person. Reliability is high when different psychiatrists provide the same diagnosis to a particular patient whereas reliability is low when that patient is assigned different psychiatric diagnoses by different psychiatrists. The neo-Kraepelinians believed that "homogeneous diagnostic grouping provide the soundest base for studies of etiology, pathogenesis, and treatment" (Robins and Guze 1970, p. 108) and believed that reliability was a necessary condition for validity (Spitzer and Fleiss 1974, p. 341). To establish reliability, they proposed that psychiatric diagnoses should have operationalised, specific criteria whereby a clinician would use standardised structured interviews and based diagnoses on standardised symptom checklists. They hoped to establish those standardised symptom checklists through statistical analysis of how symptoms cluster to accurately portray variations in symptoms exhibited in patients. For example, imagine that some patients have symptoms X and Y whilst others have symptoms Y and Z. They thought that psychoanalysts often grouped all those people into one diagnosis with the symptoms X, Y and Z whereas statistical methods would separate out both groups.[2]

The neo-Kraepelinians also placed great importance on validity. It is difficult to provide a common definition of validity because there are multiple notions and often significant aspects are implicit. The general idea is that a validated psychiatric diagnosis is securely based. Most commonly, a validated psychiatric diagnosis is understood as being secured by reality, though reality can be understood in multiple ways and reality is not always taken to be the source of validity. The neo-Kraepelinian notion of validity was primarily concerned with correlating a diagnosis with other factors, such as long-term course of the condition, family studies or laboratory findings. For example, imagine half of all individuals with a particular diagnosis slowly deteriorated over time whereas the other half slowly improved over time. This means the diagnosis is not validated

[2] Reliability was not so important for psychoanalysts because they expect irregular and changing symptoms to present given the dynamic nature of the underlying processes causing the symptoms.

when measured by long-term course. In contrast, if that diagnosis were split in two based upon difference in long-term course, both diagnoses would be validated. This notion of validity seems to be implicitly based on an underlying belief that most mental disorders have sensitive and specific causes. A sensitive cause is where everyone with the diagnosis has the cause and a specific cause is where everyone who does not have the diagnosis does not have the cause. Put another way, a sensitive and specific cause is always present in that psychiatric diagnosis and the cause is only present in people with that psychiatric diagnosis (assuming everyone was accurately diagnosed). This seems to be how the neo-Kraepelinians implicitly thought of validity. For instance, they suggested that correlation between a diagnosis and a validating factor could reach 100% if the diagnostic criteria were sufficiently refined (Robins and Guze 1970, p. 100). As a second example, they associate homogeneity of clinical picture with specific illness and this seems to be conceptualised in terms of causality, that being, high homogeneity of clinical picture means high homogeneity of causal factors (Feighner et al. 1972, p. 57). Explicit statements are largely lacking but it seems like the neo-Kraepelinians adopted a disease entity understanding of validity whereby a valid psychiatric diagnosis covers individuals with homogeneous causes. There was a "trend toward viewing disorders as discrete entities with specific aetiologies" (Moncrieff and Cohen 2005, p. 147; see also Ghaemi 2009, p. 1; Wakefield 2022, p. 14).

Whilst they promoted an innovative approach they largely ignored potential problems with it. In relation to reliability, the system of diagnosis they developed included the category of "Undiagnosed psychiatric illness" which covered anyone taken as being mentally ill but who did not meet the diagnostic criteria of any other diagnosis. Approximately 20% to 30% of patients fell into this category (Spitzer and Fleiss 1974, p. 346). This suggests their system failed to provide operationalised criteria for many individuals. Additionally, some symptoms were not included on diagnostic criteria even though they were considered to be generally present in significant numbers of clinical cases. This was because psychiatrists had difficulty agreeing upon the presence of that symptom in specific cases (Spitzer et al. 1978, p. 774). They preferred to promote high reliability over accurate descriptions of diagnosed individuals. In relation to

validity, they wished to validate diagnosis on long-term course but also recognised some disorders have a variable course. If the same causal factors could lead to a different long-term cause then variations in course does not mean different underlying causal factors. The neo-Kraepelinians recognised this possibility but, without any obvious justification, explicitly state that a diagnosis with a variable long-term course should be considered invalid (Feighner et al. 1972, p. 57).

The neo-Kraepelinian group was hugely influential. They set a research agenda of reliability and validity which soon became dominant and remains so today, though it is becoming increasingly questioned. Their approach also significantly influenced the DSM-III.

2.2.4 DSM-III

Robert Spitzer, the leading figure behind the DSM-III, was not part of the neo-Kraepelinian group but he largely shared their general goals and he significantly based the DSM-III on the key publication of the neo-Kraepelinians, the 1972 paper by Feighner et al, which listed a new set of formal, operationalised diagnostic criteria (see Cooper and Blashfield 2018, p. 15; Wakefield 2022, p. 15). Spitzer and the DSM-III task force took those sets of diagnostic criteria and slightly modified them and also added new diagnoses. These were eventually published as the *Diagnostic and Statistical Manual, 3rd edition*, a document which largely established many of the most well-known currently employed psychiatric diagnoses.

The DSM-III differed from earlier diagnostic approaches. Psychoanalysts and earlier versions of the DSM generally provided relatively few details about which symptoms were required for a diagnosis. They typically preferred to characterise psychiatric diagnoses by providing case studies and vignettes of individuals with the psychiatric diagnoses. In contrast, DSM-III diagnosis had checklists of symptoms. Typically, a diagnosis required the presence of a certain number of symptoms on the checklist rather than the presence of all symptoms. Additionally, differential diagnostic criteria were usually included, specifying how the presence or absence of specific symptoms resulted in one diagnosis rather than another. Diagnoses were also typically accompanied by other

information such as age of onset, prevalence, gender differences and level of presence in family members.

The DSM-III was intended as a set of operationalised criteria, effectively specific instructions for determining which diagnosis an individual should receive. Also, it was a categorical system whereby someone either meets or does not meet the diagnostic criteria. This contrasts with psychoanalytical approaches which placed much less value on neatly placing people into sharply demarcated diagnoses. As Demazeux and Singy write, "[i]n the history of psychiatry, never had a classification been created that was so detailed and exhaustive; never had a classification worried so much about clinical standardisation" (2015, p. xv).

DSM-III diagnosis employed symptom checklists which meant that the symptoms, rather than underlying causes, were the basis of the diagnosis. This contrasts with psychoanalysts who typically diagnosed based on the supposed type of neurotic process they believed was present. The intention was that the DSM-III would be a-theoretical, based purely on observation rather than speculative theoretical claims about underlying causes (Cooper and Blashfield 2018, p. 10). However, there is a general consensus within philosophy of science that theory-neutral operationalism is impossible (Cooper 2005, p. 101; Murphy 2006, p. 226). Additionally, though explicitly intended as theory-neutral, the authors of the DSM-III thought psychiatric diagnosis as having undiscovered biological causes (Blashfield and Keeley 2010, p. 325; Faucher and Goyer 2015, p. 202).

The DSM-III significantly approximated the approach which Kraepelin and the neo-Kraepelinians had advocated for and it rejected the approach most psychoanalysts adopted. Eight years after the publication of the DSM-III Berrios and Hauser wrote that "[p]sychiatry still lives in a Kraepelinian world" (1988, p. 813). The DSM-III was intended to be based on data rather than theoretical claims or in clinical wisdom, though the diagnoses formulated were only empirically informed rather than based on actual empirical studies. There was heavy emphasis on producing operationalised criteria which would increase reliability. Additionally, though explicitly a-theoretical, it was assumed that mental illness was caused by underlying biological factors. For these reasons the pendulum had again swung away from non-disease entity approaches and again

towards disease entity approaches. Disease entity approaches gained prominence with Kraepelin, then became diminished with psychoanalysis, and then became prominent again with the DSM-III.[3]

2.2.5 Later Editions of the DSM

The DSM-III was published in 1980 and would be followed by DSM-III-R in 1987, DSM-IV in 1994, DSM-IV-TR in 2000, DSM-5 in 2013 and DSM-5-TR in 2022 (TR means text revisions, which are intended as minor revisions). The changes of each edition from the DSM-III were relatively small. This is despite the neo-Kraepelinians and Robert Spitzer intending the diagnostic criteria of the DSM-III to be provisional and for them to be revised on relevant new evidence. The DSM-III was intended to have high reliability and this would then facilitate the production of validated psychiatric diagnoses. However, the diagnoses in DSM-III soon became reified, treated as being established disease entities, rather than being refined or replaced over time as Spitzer wished (Wakefield 2022, p. 18).

Across successive editions of the DSM there was an increase in the number of diagnoses. This consequently resulted in an increase in comorbidity (the same individual receiving more than one diagnosis). However, some of these additional diagnoses involved splitting up diagnoses in earlier DSMs rather than being completely new diagnoses.[4] More important were the changes which did not occur. The causal basis of the vast majority of psychiatric diagnoses has not been established. Neither do most psychiatric diagnoses have specific treatment, since treatments are applicable (though usually rarely effective) to multiple diagnoses. Whilst it is generally, though not universally, accepted that the DSM-III and later DSMs improved reliability it was increasingly clear

[3] Other factors also influenced the adoption of the DSM-III. For example, conceptualising psychiatric diagnoses as disease entities fitted the desires of drug companies, who were rapidly developing and marketing new drugs, and insurance companies, who wanted to use the DSM to demarcate who would be given insurance payouts.

[4] This is partly caused by moving diagnoses from subtypes to full diagnoses. Some diagnoses which were only considered subtypes in earlier DSMs, and not given their own unique code in those DSMs, were given unique codes in later DSMs.

that increased reliability had not then delivered increased validity.[5] Concerns over the DSM began to grow and these would explicitly manifest themselves in the debates of the DSM-5.

2.2.6 DSM-5

The perceived lack of progress with successive editions of the DSM led to calls for DSM-5 to take a significantly different approach. These criticisms were made forcibly by Steven Hyman, the ex-director of the National Institute for Mental Health (the main U.S. government funded body for mental health research). He believed that DSM diagnosis had three primary problems. Firstly, over-specification, which means that the criteria employed to receive a diagnosis are too precise. He considered this problematic because individuals typically have heterogeneous clinical pictures and because many individuals have psychiatric issues but do not meet the criteria for any diagnosis (Hyman 2010, p. 167). Secondly, comorbidity, which is where the same individual receives multiple diagnoses (Hyman 2010, p. 167). Thirdly, psychiatric diagnoses are not closely related to causal factors. Hyman suggests these problems stem from the DSM being categorical and that DSM diagnoses have become reified, taken as being correct, rather than modified and updated based upon new evidence (Hyman 2010, p. 168). Hyman argues that "[t]he problems that have emerged within the DSM 'paradigm' (based on operationalized criteria that define a large number of categorical disorders) cannot be fixed by tinkering with existing criteria sets or by adding or subtracting diagnoses at the margins" (2010, p. 171). Rather than take the practically difficult step of making the DSM-5 entirely or largely dimensional, Hyman wished initial steps to focus upon adding dimensional information. A dimensional system is where a particular diagnosis would be rated as present to a particular degree rather than simply rated as being present or absent as currently occurs under a categorical system (see later in this chapter for a detailed discussion of categorical and dimensional approaches). It is worth noting that the most recent version of the ICD,

[5] See Vanheule (2017) for the claim that DSM-III did not increase reliability, rather, the authors of the DSM simply just reduced the level of agreement which was needed for reliability.

the ICD-11, has incorporated dimensionality into the manual, though only in relation to personality disorders.

Initial work on the DSM-5 explored the possibility of implementing a dimensional system. As work progressed the focus moved onto making only personality disorders dimensional. After much dispute it was decided that personality disorders would remain categorical within the main body of the manual but an appendix would include dimensional versions of personality disorders. This fell short of Hyman's suggestions and was even more distant from critics who advocated for the implementation of full dimensional systems (Cuthbert and Insel 2013, p. 4).

2.2.7 Crisis of Confidence over Validity

Currently employed psychiatric diagnoses have not been validated and there is growing suspicion that they will not be validated in the future because they are too causally heterogeneous. This has led to a crisis of confidence over validity (for overview see Zachar and Jablensky 2014). Some believe that this is a problem of currently employed psychiatric diagnoses and that superior psychiatric diagnoses would likely be validated. Others, however, believe the problem lies with notions of validity and that we need to rethink validity itself.

Notions of validity are typically based around sensitive and specific causes (Smith and Combs 2010, p. 206). That being, the cause is present in every member of a particular population and not present in members of any other population. Put another way, "modern understandings of validity encompass aspects of essentialism" (Blashfield and Keeley 2010, p. 334). This is seemingly based in a disease entity approach which presupposes two points. Firstly, that there is a relatively simple relationship between causes and symptoms. Specific causes result in specific symptoms. Secondly, it presupposes that symptoms cluster together in a relatively uniform way. When one symptom is present then so too typically is another set of specific symptoms. However, both these points have been challenged by scientific investigation. Firstly, there is a complicated relationship between causes and symptoms. Scientific investigation suggests that "specific combinations of the same antecedent causes can lead

to different clinical outcomes ('multi-finality') and different antecedent can lead to the same outcome through common developmental pathways ('equi-finality')" (Cloninger 2014, p. 205; see also Kendler 2005a, p. 1247; Murphy 2014a, p. 120; Hoffman and Zachar 2017, p. 70). The same causes can result in different symptoms. Two individuals may have similar causal factors present and yet exhibit different symptoms. Also, different causes can result in similar symptoms. Two individuals may have quite different causal factors present yet have similar symptoms. Secondly, modern evidence suggests that symptoms do not cluster in a neatly demarcated manner. If every individual who exhibited one particular symptom were studied then that one symptom is not always or usually accompanied by a set number of specific symptoms. Rather, it would be accompanied by many different combinations of symptoms in different individuals.

This new understanding of how causes relate to symptoms has led to calls for a rethink of what validating a psychiatric diagnosis should consist of. Rather than a gold standard of finding sensitive and specific causes, a more minimal approach which understands validity in terms of probabilistic relationships between causal factors and symptoms is required.

2.2.8 RDoC, HiTOP and Dimensional Diagnoses

In response to the perceived problems with the DSM, in 2009 a new project known as RDoC (Research Domain Criteria) was started. RDoC is run by the National Institute for Mental Health. This is part of the National Institute of Health, an agency of the United States Department of Health and Human Services. RDoC not only takes currently employed psychiatric diagnoses as flawed but also as holding back causal investigation. Most psychiatric research into causation is based around causally investigating a specific DSM diagnosis. However, if different individuals with the same psychiatric diagnosis have heterogeneous symptoms and causes then it is unsurprising, argues RDoC psychiatrists, that psychiatry has not found the causes of mental illness. RDoC solution to this problem is to largely dispense with currently employed psychiatric diagnoses (Cuthbert and Insel 2013, p. 4). Rather than base studies upon what

psychiatric diagnosis an individual has, RDoC will recruit individuals for studies based upon some other factors, such as a particular symptom. This will typically result in people with many different psychiatric diagnoses taking part in a particular study but this is not considered problematic. "The RDoC initiative is intended to uncouple research questions from traditional diagnostic categories that are of limited validity and/or that are too heterogeneously large for productive validation against biological phenomena of smaller granularity" (Kozak and Cuthbert 2016, p. 288). RDoC indicates what the future may hold. The National Institute for Mental Health intends to steadily increase funding for RDoC research until the point that more funding goes into RDoC research than goes into currently typical DSM-based research.

RDoC takes a dimensional approach rather than a categorical approach when investigating the causal basis of symptoms (I outline categorical and dimensional approaches below). It seems to also generally favour reformulating psychiatric diagnoses dimensionally. Additionally, there is an appreciation of the complicated relationship between causes and symptoms. RDoC intends to chart the dynamic and often non-linear causal relation between genes, neuroscience, psychology, the environment and behaviour. There seems no expectation that sensitive and specific causes will be found for significant numbers of either individual symptoms or clusters of symptoms. Rather, the emphasis seems to be on how dimensional measures of characteristics will relate probabilistically to dimensional causal factors. Given that RDoC generally emphasises a dimensional approach and seems to have abandoned the gold standard of validation on sensitive and specific causes, RDoC generally rejects disease entity models. It is important to note that RDoC is not aiming to produce a new set of psychiatric diagnoses, rather, it is building an evidence basis that could be used in producing new psychiatric diagnoses. Given the general spirit of RDoC, it seems likely that RDoC inspired psychiatric diagnoses would aim to be more homogeneous in symptoms covered and underlying causes than DSM diagnoses.

A similar story can be told of HiTOP (Hierarchical Taxonomy of Psychopathology). This is a new diagnostic system which has been produced by a consortium of researchers. A work in progress version of the diagnostic scheme was published in 2017 (Kotov et al. 2017). Similar to

RDoC, it starts from the position that currently employed psychiatric diagnoses are flawed and need replacing. The new diagnostic "system offers to address problems of arbitrary disorder boundaries (consequences of which include subthreshold and not otherwise specified cases) and substantial unreliability of traditional diagnoses, by characterizing psychopathology in terms of dimensions rather than categories" (Kotov et al. 2018, p. 24). HiTOP focuses upon reorganising how symptoms relate to one another, finding alternative links that are dimensionally related which do not respect current diagnostic boundaries of the DSM and ICD. HiTOP "does not include any of the traditional diagnoses" (Kotov et al. 2017, p. 456). RDoC and HiTOP differ in two main ways. Firstly, RDoC heavily focuses upon causes where HiTOP itself does not itself cover causes (though other projects, such as RDoC, might chart how causes relate to the HiTOP system). Secondly, RDoC covers both biology and psychology but focuses more upon biology, whereas HiTOP does not cover biology and contains much more finegrained psychological demarcations. HiTOP is another example of a significant project which is rejecting the categorical approach of the DSM and ICD in favour of a dimensional approach. It is worth noting that HiTOP has primarily been used in academic research rather than clinical practice, and the project is currently in development. Exactly what clinically usable psychiatric diagnoses HiTOP will settle upon is currently unknown but, like any RDoC inspired psychiatric diagnoses, they will likely be more homogeneous in symptoms covered and underlying causes than DSM psychiatric diagnoses.

2.3 Different Types of Psychiatric Diagnoses and Alternatives to Psychiatric Diagnoses

I have finished outlining a historical narrative which started with Kraepelin and has reached the current situation that psychiatry is in. I will now outline in detail different approaches to psychiatric diagnoses. I will outline categorical psychiatric diagnoses and then outline

dimensional psychiatric diagnoses. I will also describe an alternative to psychiatric diagnoses known as formulations.

The type of psychiatric diagnosis found in the main diagnostic manual, the Diagnostic and Statistical Manual (DSM) and the International Classification of Diseases (ICD), are predominantly categorical, polythetic psychiatric diagnoses. Categorical diagnoses are binary. This means that a particular individual either meets the diagnostic criteria or does not meet the diagnostic criteria. Everyone who is accurately assessed will fall into one of two categories: either they have a particular diagnosis or they do not. Under a categorical system no one only partly has a diagnosis. Many of the categorical diagnoses included in the DSM and ICD are polythetic. This means there is more than one way to meet the diagnostic criteria. It is not the case that an individual needs to exhibit all of the symptoms on the diagnostic checklist to receive a diagnosis. Rather, they only need to exhibit a certain number of them. They are polythetic in the sense that two individuals can receive the same diagnosis despite not exhibiting the same symptoms. Each psychiatric diagnosis has a diagnostic criterion. This typically consists of one or more lists of symptoms. To receive that diagnosis an individual will typically need to exhibit a certain number of symptoms on the diagnostic criteria. This means that individuals with different combinations of symptoms can receive the same psychiatric diagnosis. Note that a diagnosis could be just categorical without being polythetic, or could be polythetic without being categorical, but most currently employed diagnoses are both categorical and polythetic.

There are advantages to employing categorical, polythetic psychiatric diagnoses. Firstly, they demarcate people into various categories which are intended to communicate important information about them. Knowing that someone is autistic suggests they are more likely to behave, think, feel and perceive in certain ways compared to someone who is schizophrenic or a non-diagnosed member of the population (Maung 2019, p. 510; Sadler 2005, p. 66). Secondly, this is useful for administrative matters. Decisions need to be made over who is eligible for treatment, support and benefits. Diagnoses are an easy means to demarcate between those who are eligible and those who are not (Helzer et al. 2006, p. 1673; Hudziak et al. 2007, p. 21). Thirdly, diagnoses can be correlated

with wider factors. Studies can establish the correlation between people with a particular diagnosis and factors like treatment responses, gender ratios, family history and causal factors (Heinrichs 2015; Sadler 2005, p. 66). Fourthly, cultural and political movements can be based around categorical psychiatric diagnoses (Orgota 2013, p. 80). Fifthly, psychiatric diagnoses can act as an explanation of how people act whereby we say a particular person acted in a particular way because they are autistic (Fellowes 2021, p. 4517; Orgota 2013, p. 79).

I now consider nine common criticisms of categorical psychiatric diagnoses.

1. Virtually no categorical psychiatric diagnosis has established causal origins. Typically, a psychiatric diagnosis will only be very weakly correlated with causal factors (Cuthbert and Insel 2013, p. 3; Poland 2014, p. 34).
2. Categorical psychiatric diagnoses are sometimes seen as being solely caused by biological abnormalities. This discounts how the environment can cause, influence or exacerbate symptoms and encourages prescribing drugs which may not be helpful (Johnstone 2018, p. 33; Kinderman et al. 2013, p. 2).
3. Categorical psychiatric diagnoses primarily describe which symptoms someone exhibits. This does not reveal an individual's past life history, their current life situation or aspects of their personality which are not symptoms. It does not reveal what feelings they have towards their symptoms, what understanding they have of them, or whether they have tried to prevent the symptoms occurring or modify how they occur (Johnstone 2018, p. 33; Kinderman et al. 2013, p. 2).
4. Critics often see psychiatric diagnoses as being unhelpfully pathologizing, negatively impacting the self-perception of the diagnosed individual and the perception of others towards the diagnosed individual (Johnstone 2014, p. 275; Vanheule 2017, p. 275).
5. Categorical psychiatric diagnoses only note that an individual either meets the diagnostic criteria or they do not. It fails to describe the degree to which the diagnostic criteria have been met or the degree to which someone fails to meet the diagnostic criteria (Helzer et al. 2006, p. 1673; Hudziak et al. 2007, p. 21).

6. Most people who meet the criteria for one categorical psychiatric diagnosis will typically also meet the criteria for more than one (this is known as being co-morbid). Unless most people with one underlying illness actually also have a second underlying illness then this suggests each diagnosis is failing to adequately describe their underlying illness (Hyman 2010, p. 167; Kirk et al. 2015, p. 69).
7. Currently employed psychiatric diagnoses play a limited role when selecting treatments. Typically, a particular drug or therapy is not effective for everyone with a particular categorical psychiatric diagnosis (Horwitz 2002, p. 226; Moncrieff and Cohen 2005, p. 147).
8. Categorical psychiatric diagnoses may hold back causal research because they typically group together individuals with very different causes (Cuthbert and Insel 2009, p. 989; Poland 2014, p. 46).
9. Being categorical, an individual either meets or does not meet the diagnostic criteria of a categorical psychiatric diagnosis. However, an individual can be subthreshold whereby they have some of the symptoms of the psychiatric diagnosis but not enough to actually meet the diagnostic criteria. This leaves subthreshold individuals in a grey area of both having and not having the psychiatric diagnosis (Jablensky 2016, p. 27; Okasha 2009, p. 130).

The DSM-III was categorical, whereby everyone either had or did not have the psychiatric diagnosis, just like how someone either is, or is not, an instance of a disease entity. However, the DSM-III was also polythetic, whereby there was more than one way to meet the diagnostic criteria of each diagnosis. This means that DSM-III was intended as a stepping stone to finding the correct disease entities which would not be polythetic. The DSM-III was intended to have high reliability and this would then facilitate the production of validated psychiatric diagnoses which would not be polythetic. Later DSMs, including DSM-5, were also almost entirely categorical. However, DSM-III was seen as a stepping stone towards non-polythetic diagnoses, whereas the psychiatric diagnoses in later DSMs soon became reified, treated as being disease entities, rather than being refined or replaced over time as Spitzer wished. This meant that the psychiatric diagnoses in later DSMs became accepted as polythetic.

An alternative to categorical psychiatric diagnoses is dimensional diagnoses. The generally understood difference between categories and dimensions in psychiatry is that categories are either present or absent whereas dimensions are present to a degree. An everyday example which highlights the difference between categorical and dimensional approaches is height (Borsboom et al. 2016, p. 3; Zachar and Kendler 2007, p. 559). Categorically, people are tall or not tall, and they are short or not short. Dimensionally, we can give the specific figure in feet, inches or centimetres that people have. In relation to psychiatry, a "dimensional system classifies clinical presentations based on quantification of attributes rather than the assignment to categories" (Kraemer et al. 2004, p. 17). When people are diagnosed dimensionally, they are given a score or multiple scores, rather than just being given a result of present or absent. "A categorical approach to a diagnosis results in labeling each subject as either having (D+) or not having (D–) a disorder… A dimensional approach results in labeling each subject with an ordinal score (D), with higher scores a stronger indicator of the presence of the disorder" (Kraemer et al. 2004, p. 18). For example, in relation to depression, "patients can either be rated in terms of how many symptoms of depression they have (continuum [dimensional]), or they can be diagnosed as "having" depression or not (category)" (Hopwood et al. 2023, p. 3). Alternatively, rather than giving a single figure, someone might be dimensionally rated for every symptom of depression, resulting in multiple scores conveying how much they exhibit various aspects of depression. Finally, multiple dimensional measures can be employed without then taking these as in some manner relating to one thing. For example, someone can be dimensionally measured for low mood, difficulty sleeping and low appetite, but these three measurements are not then taken as relating to some further thing we call depression. However, researchers sometimes use the word dimension in quite different ways (for discussion see Markon et al. 2011; Hopwood et al. 2023). I shall now outline different ways of understanding dimensions.

Firstly, categories are binary whereas dimensions have a scale with more than two points. Whereas something either is or is not present on a categorical system, on a dimensional system the scale demarcates to a greater degree than two binary points of present or absent. Dimensions

have at least three points on their scale, such as 'not very present', 'moderately present' and 'heavily present' (Kraemer 2007, p. 10). Whilst a dimensional system will have at least three points, there could be hundreds, thousands or more points on the dimensional scale.

Secondly, we can think of categories as measuring demarcated groups whereas dimensions relate to a measurement scale without sharp demarcations (Widiger and Samuel 2005, p. 498). Categories are understood as discontinuous whereas dimensions are understood as continuous. Categories as discontinuous means that everyone is placed somewhere among various discrete positions and there is a gap between those discrete positions. Dimensions as continuous means rating people on a scale which does not have gaps between discrete groups.

Thirdly, we can think of categories and dimensions as relating to qualitative vs quantitative, or difference in kind vs differences in degree (Haslam et al. 2020, p. 1418). A category signifies a qualitative difference, a difference in kind, that exists between people. On this understanding, someone who is depressed is qualitatively different from someone who is not depressed. In contrast, a dimensional approach is a quantitative difference, a difference in degree, between people. On this understanding, everyone who is dimensionally depressed is exhibiting the same thing even though different people can exhibit it to significantly different degrees.

Fourthly, dimensions are typically understood as extending across an entire population but researchers sometimes look for more homogeneous dimensions. Researchers sometimes look for dimensions "that can be distinguished from 'sibling' constructs – meaning constructs that measure similar albeit distinct aspects of psychopathology" (Hopwood et al. 2023, p. 4). The aim is to establish "continuous constructs that are relatively homogenous and empirically distinct within the context of a multidimensional model" (Hopwood et al. 2023, p. 2). Whilst still dimensional, these dimensions might aim to track multiple more specific manifestations of depression rather than simply track one large notion of depression.

Sometimes categories and dimensions are spoken about as being measurement systems and sometimes they are spoken about as being statements about the nature of reality. It is important to demarcate between these. The DSM is a categorical measurement system but a common

complaint is that it does not reflect the reality of psychiatric phenomena. Whilst we can make measurements in a categorical manner we might think that the domain we are measuring actually exists in reality in a dimensional form. Similarly, we can do studies using statistical techniques like factor analysis to chart how properties are distributed among a population. Having done this, we might find studies reveal categories or dimensions, whereby the properties being studied either fall into demarcated groups or do not. Strictly speaking, at this point we are dealing with measurements. However, we might then take sufficiently robust and well tested measurements as revealing how things are in reality. Boorsboom et al. write that "categorical representations line up naturally with an interpretation of disorders as discrete disease entities… while continuum [dimensional] hypotheses are most naturally consistent with the idea that a construct varies continuously in a population" (Borsboom et al. 2016, p. 3; see also Kraemer et al. 2004, p. 17). As such, a demarcation can be made between, firstly, measuring in a categorical or dimensional manner, secondly, a statistical analysis which results in categories and dimensions and, thirdly, a belief that our measurements show that a particular domain exists in reality as categories or dimensions.

Dimensional diagnoses have multiple advantages. Firstly, unlike categorical diagnoses, dimensional diagnoses discriminate between the degree to which symptoms are presented in diagnosed individuals (Hudziak et al. 2007, p. 21; Helzer et al. 2006, p. 1674; Brown and Barlow 2005, p. 552). Secondly, dimensional diagnoses help with the problem of comorbidity. A person can be rated dimensionally for any number of symptoms, regardless of which categorical psychiatric diagnosis those symptoms are associated with (Helzer et al. 2006, p. 1673; Widiger and Samuel 2005, p. 501). Thirdly, dimensional diagnoses avoid the problems of NOS diagnosis. A dimensional approach can rate dimensionally each symptom and so there is no issue whether they meet some pre-determined cut off or instead have to be given a less specific NOS diagnosis (Potuzak et al. 2012, p. 2; Widiger and Samuel 2005, p. 495). Psychoanalysts were typically closer to a dimensional approach, unlike Kraepelinian notion which is closer to a categorical approach.

There are disadvantages to dimensional diagnoses. Rating an individual upon various different factors, rather than simply noting that they

categorically either do or do not have a psychiatric diagnosis, has some practical implications. Firstly, categorical diagnoses are easier to communicate compared to dimensions since it is easier to describe someone as, say, autistic rather than list each score for various attributes (Hudziak et al. 2007, p. 21; Widiger and Samuel 2005, p. 501). Secondly, a dimensional system does not by itself provide a cut off for deciding who is sufficiently ill for treatment, support, medication, hospitalisation and insurance coverage whereas a categorical system does (Helzer et al. 2006, p. 1673; Widiger and Samuel 2005; Wakefield 2022, p. 18).

An alternative to categorical and dimensional approaches is to stop providing psychiatric diagnoses altogether. Rather than saying that someone has a particular categorical or dimensional diagnosis some advocate for dispensing with such diagnoses and instead advocate for going directly to the individual. Rather than simply saying an individual is, say, autistic, an alternative approach is to instead say they have the symptom of low social skills, the symptom of repetitive behaviour and so on. Additionally, those advocating for such approaches typically believe focus should also go upon many other aspects of people, such as their personality, desires, past history and current situation. Such approaches could be broadly described as 'person centred'. Here I will discuss a particular person centred approach called formulations (sometimes also called psychological formulations, psychosocial formulation or a clinical case formulation). A formulation is the

> process of co-constructing a hypothesis or 'best guess' about the origins of a person's difficulties in the context of their relationships, social circumstances, life events, and the sense that they have made of them… Formulation draws on two equally important sources of evidence: the clinician brings knowledge derived from theory, research, and clinical experience, while the service user brings expertise about their own life and the meaning and impact of their relationships and circumstances. (Johnstone 2018, p. 32; see also Vanheule 2017, p. 172)

Formulations are designed to fit specific individuals (Aveline 1999, p. 207; Johnstone 2018, p. 32; Vanheule 2017, p. 104). The formulation will describe how a specific individual is behaving and feeling. Additionally,

the formulation will explain why the individual is exhibiting unusual behaviour or feeling distressed. A formulation will aim to describe the specific behaviours an individual exhibits and describe their feelings. Additionally, a formulation aims to give an account of why an individual acts in unusual ways or feels distress. Related to this, formulations aim to avoid seeing the origins of behaviour and distress as being biomedical in nature or as exclusively biomedical in nature. Rather than simply saying that mental distress is caused by an illness with underlying biological causes a formulation places emphasis on factors that are specific to individuals such as their personality, psychology, life history, social circumstances and life events (Johnstone 2018, p. 33; Vanheule 2017, p. 172).

Advocates of formulations argue that they avoid many problems of categorical diagnosis and some problems with dimensional diagnoses. Firstly, unlike categorical and dimensional diagnoses, formulations cover many specific aspects of the individual, situating an individual within wider contexts such as social situation and life history. Secondly, psychiatric diagnoses can lead to a loss of meaning and significance for the diagnosed individual because they believe their actions stem from an illness. In contrast, formulations often portray unusual actions and mental distress as understandable reactions to adverse situations and so helps restore a sense of meaning. Formulations typically assume that seemingly pathological behaviour is actually an understandable response to bad circumstances (Johnstone 2018, p. 42). Thirdly, categorical diagnoses do not cover the qualitative nature of symptoms. The diagnosis indicates that symptoms are present but does not indicate the degree to which any particular symptom is present or how a particular individual may exhibit symptoms in an idiosyncratic way. In contrast, formulations will tailor descriptions of symptoms to particular individuals (Aveline 1999, p. 207; Johnstone 2018, p. 32; Vanheule 2017, p. 104).

There are potential problems with formulations. Firstly, formulations are typically informed by psychological theory. Constructing a formulation involves applying psychological theory to specific people. This is not necessarily problematic in itself and can be a strength of formulations. However, advocates of formulations admit that different practitioners can produce quite different formulations based upon their theoretical inclinations (Johnstone 2014, p. 183; Vanheule 2017, p. 173). On one

approach, it could be assumed that one theoretical approach can give the correct formulation (given a specific aspect of a specific patient under investigation) whereas other theoretical approaches will give incorrect formulations. If so, then this raises the problem of working out which theoretical approach is the correct one and the danger that the correct theoretical approach is either not used or is mistakenly taken as being an incorrect approach. Alternatively, as Vanheule (2017, p. 198) argues, each different theoretical approach could be seen as revealing different aspects of a person so no single theoretical approach gives the correct formulation. However, this still leaves the problem of ensuring that all the relevant theoretical approaches are used and the problem of how to integrate together different theoretical approaches. Secondly, formulations take considerable time to do (Johnstone 2014, p. 183; see also Vanheule 2017, p. 203). Additionally, formulations are best constructed not by a single practitioner but by a team of professionals. This helps increase the likelihood that good formulations will be produced (Johnstone 2014, p. 219). This again significantly increases the time required to produce good quality formulations. An advocate of formulations would argue that you cannot understand something as complex as someone in mental distress without spending time, but this creates problems when little time is available. This means that under current health care systems there may be limited opportunity to actually carry out formulations.

There is a complicated relationship between categorical, dimensional and person centred approaches. Whilst in theory they are easy to demarcate between, in practice they can blur into one another. For example, imagine in a clinical setting that a clinician has someone in front of them who they know are categorically autistic. They will likely very quickly get a rough idea of whether the individual is heavily autistic or mildly autistic. Whilst much less precise than a specific value given by a dimensional system, doing this already starts to convey dimensional information. Similarly, given time they will hopefully get an idea of which symptoms the individual exhibits and what is going on in their life. Whilst much less detailed than a psychological formulation, this would start to convey person centred information. Similarly, a clinician who knows that someone scores a particular figure on a dimensional scale for autism can then

start to work out which specific symptoms they exhibit and other details of their life. Finally, even if they try not to, some clinicians who are providing psychological formulations might still have at the back of their mind that some people seem autistic, or autistic to a degree.

Categories and dimensions can be combined in a more explicit manner. Having imposed a categorical boundary which only some people meet we can then dimensionally rate everyone with that categorical diagnosis. An example of this is some understandings of the autistic spectrum. On some understandings, only some people are autistic but among those people we can rate the degree to which they are autistic. Also, we can also rate everyone dimensionally and demarcate categorical cut offs, such as how blood pressure is dimensional but we impose a categorical cut off of hypertension upon it (Aftab et al. 2024). Categories and dimensions can be combined in multiple ways. We might dimensionally study the entire population and find clusters of symptoms, form categories from those dimensions, dimensionally study those categories and find subdivisions, form subtypes from those subdivisions, and then dimensionally study those subtypes and so on. On this approach, someone might be told they score a particular figure on the autism dimension, then told they meet the categorical criteria for autism, then told they meet a subtype, then told what degree they meet the subtype. This sort of combination of categorical and dimensional seems to me to be a significant improvement upon simply diagnosing people with DSM 5 autism. It is potentially harder to directly combine categorical or dimensional approaches with person centred approaches. Having rated someone categorically, dimensionally, or a combination of the two, we can then do a further person centred analysis of someone, and I think doing so is a very good idea, but this is more doing something additional rather than directly combining it with categories or dimensions.

Two key aims of this book, as outlined in the introduction, are to argue that polythetic categorical psychiatric diagnoses are scientifically legitimate despite not describing entities in the mind independent world and to give an indication of what a good categorical diagnosis would consist of. At the same time, I also think dimensions and person centred approaches should be employed. This then leads to the other two key aims. I will argue that we should still employ categorical psychiatric

diagnoses even if we converted all DSM diagnoses into being dimensions. Also, I will argue that there still could be good reason to employ categorical diagnoses even if RDoC and HiTOP ultimately produce relatively homogeneous dimensional groupings.

2.4 Summary

Psychiatry has oscillated between disease entity and non-disease entity approaches. Kraepelin established a system of diagnoses based upon notions of disease entities. Psychoanalysts generally downplayed or rejected disease entities. The neo-Kraepelinians and the DSM-III implicitly endorsed disease entities. Both also conceptualised reliability and validity in terms of disease entities. However, modern evidence suggests that currently employed psychiatric diagnoses are not disease entities. Additionally, modern evidence is significantly compatible, and may even suggest, that there are no, or few, disease entities out their waiting to be found. The areas of the world studied by psychiatry may not take the form of disease entities. This has led to the crisis of confidence over validity and RDoC. Whilst the future of psychiatric diagnoses is unclear there appears to be a movement away from disease entities given the desire for a dimensional approach. Meanwhile, others advocate for abandoning both categorical and dimensional diagnoses.

All this has significant philosophical implications. I outlined in the introduction chapter how critics of psychiatric diagnoses claim they are not real or scientifically legitimate because they are not real entities. Such concerns seem to be adopting a disease entity model whereby psychiatric diagnoses need be disease entities to be real and scientifically legitimate. Critics are concerned that psychiatrists are treating psychiatric diagnoses as though they are disease entities when actually they are not. Do such concerns stand if psychiatrists are actively moving away from a disease entity approach? Even if disease entities were dropped as a standard for assessing psychiatric diagnoses, what other standard should psychiatric diagnoses be judged on? In future chapters I will uncouple questions of scientific legitimacy from disease entities and provide an alternative

foundation to see psychiatric diagnoses as scientifically legitimate. I now turn to showing how philosophers of psychiatry developed their own accounts of how psychiatric diagnoses should be assessed.

2.5 Philosophy of Psychiatry Which Orientates Towards Analytical Philosophy of Science

There are many different areas of philosophy of psychiatry. Topics which philosophers of psychiatry might research include the demarcation between health and illness, phenomenology, enactivism, identity, meaning and values. In this book I primarily focus upon a particular branch of philosophy of psychiatry, namely that which is primarily based around analytical philosophy of science. Analytical philosophy of science is a branch of philosophy with long historical roots featuring individuals like Mach, Poincaré, Carnap, Popper, Quine, Hempel, Kuhn and Feyerabend. Broadly speaking, common themes include the relationship between theory and evidence, the philosophical status of aspects of science (theoretical claims, observations, laws, causes and entities), the role of non-scientific factors in scientific research and the reality of scientific postulations. Though philosophy of science has a long history, with various approaches growing and waning in popularity, philosophy of psychiatry which is primarily orientated towards analytical philosophy of science is largely a post-2000 phenomenon.

Philosophers of psychiatry have put forward varying views about what psychiatric diagnoses are and how they relate to reality. I will now outline how philosophers of psychiatry have outlined varying views on whether, or under what circumstances, psychiatric diagnoses can be considered real and scientifically legitimate. The most common approach has two distinct elements. Firstly, psychiatric diagnoses can in principle be real and scientifically legitimate. On this understanding psychiatrists could potentially formulate psychiatric diagnoses in a manner that would mean they are real and scientifically legitimate. Secondly, philosophers of psychiatry typically doubt that currently employed psychiatric diagnoses are the real ones. On this understanding, currently employed psychiatric

diagnoses are not real or scientifically legitimate but some alternative ones could be.

In the remainder of this chapter I focus on four issues.

1. What is the target of psychiatric diagnoses? This would establish what psychiatric diagnoses could describe.
2. Is that target real? This would establish whether psychiatric diagnoses could describe something real.
3. How and under what circumstances would this process deliver scientific legitimacy? This would establish whether we should consider psychiatric diagnoses to be scientifically legitimate.
4. How well do currently employed psychiatric diagnoses (or a specific psychiatric diagnosis) describe that target? This would establish how well they describe whatever degree of reality that target has.

In this chapter I outline the dominant thinking by philosophers on these issues to date, and in later chapters I will draw upon neo-Kantianism to outline alternative approaches to these issues.

2.5.1 The Target of Psychiatric Diagnoses

The most common view held by philosophers of psychiatry is that the target of psychiatric diagnoses should be natural kinds. Tekin writes that "the scientific legitimacy of mental disorders has hinged on their status as natural kinds" (Tekin 2016 p.148). Natural kinds have a long historical basis in philosophy (Kripke 1972; Putnam 1975) and philosophy of science (Boyd 1991; Hacking 1991). They are typically thought of as entities which, according to various philosophical accounts, have significant levels of homogeneity. They are typically seen as the sort of things which scientists would stereotypically be believed to study. We have a stereotypical notion that scientists study things which actually exist in the external world. Stereotypical examples of things which scientists study are things like lions, argon and gravity. These are things which are typically believed to exist in the external world which scientists can discover facts about. Examples of natural kinds in psychiatry which philosophers of psychiatry

have suggested are Down syndrome (Cooper 2012, p. 38) and schizophrenia (Beebee and Sabbarton-Leary 2010, p. 23). Within philosophy of science, some accounts are essentialist whereby each entity has the same cause and same properties (Ellis 2001, p. 1). Other accounts only require that each instance of the entity have the same broadly homogeneous underlying causal mechanisms (Boyd 1991, p. 142; Dupré 1993, p. 5; Psillos 1999, p. 289).

There are slightly different accounts of what sort of entity would be the target of psychiatric diagnoses. Typically, accounts either emphasise similarity of underlying causes or similarity of overt characteristics. These will often map onto one another. An example of a psychiatric diagnosis which fits this notion of similarity of underlying causes and overt characteristics would be Down's syndrome. Individuals diagnosed with Down's syndrome have the same underlying cause and quite similar overt characteristics. It is an easy step to think of Down's syndrome as a naturally forming part of the world, like tigers and oxygen, which scientists can study. Cooper employs Down's syndrome as an example of a diagnosis which she considers to be a natural kind and uses this position to argue that Down's syndrome is real (Cooper 2012, p. 39). However, it is worth noting something can be considered a natural kind with only similarity in overt characteristics or only similarity in underlying causes. Both are not required. I show this by drawing upon Haslam's exposition of four different accounts of kinds in psychiatry and then compare his approach to two other approaches.

Haslam outlines essentialist natural kinds where a "specific, causally efficacious pathological process, mechanism or structure is present" (Haslam 2002, p. 212). All instances of the kind have the cause whereas all instance of other kinds lack that cause. Additionally, there are clear boundaries between the symptoms of essentialist natural kinds and other kinds. No one is only partly an instance of the kind, rather, either someone is or is not an instance of the kind. Essentialist accounts emphasise both similarity of underlying causes and overt symptoms.

Haslam outlines discrete kinds which, like essentialist kinds, have clear boundaries but, unlike essentialist kinds, do not have essentialist causes. Multiple causes can produce a discrete set of symptoms. Haslam considers such kinds as determined by the world since they have discrete

boundaries at the level of symptoms even though they do not have all the same causes (Haslam 2002, p. 210).

There are also fuzzy kinds which lack the essentialist causes of essentialist kinds and lack the discrete boundaries of essentialist and discrete kinds. Some individuals fully meet the criteria for the fuzzy kind and many individuals meet none of the criteria, however, unlike discrete and essentialist kinds, the same individual could only partly be an instance of the fuzzy kind. They only have some, but not all, of the attributes required to count as an instance of that kind. Haslam considers them as having objective boundaries, writing that "[f]uzzy kinds have real, if indistinct, category boundaries" (Haslam 2002, p. 214). They have real boundaries, since some members of the fuzzy kind meet all the criteria for the fuzzy kind, but they also have indistinct boundaries since some members do not meet the criteria for fuzzy kinds. On my reading of Haslam, the world objectively determines the boundaries of those members of the fuzzy kind who meet all the criteria whereas psychiatrists make an imposition by considering those other members who do not meet the full criteria as members of the fuzzy kind. Like discrete kinds, fuzzy kinds are grouped by overt characteristics rather than underlying entities.

Finally, there are practical kinds where there is no clear separation at all between members of the practical kind and non-members. Whilst some members of a particular fuzzy kind are clearly separated from non-members (even if some other individuals are only partly instances of that fuzzy kind), there is no clear boundary between every member of the kind and non-members. Rather, the boundary is "imposed for practical reasons [rather than being] detected because it is in some respect objectively present" (Haslam 2002, p. 214). The boundary conditions of a practical kind are imposed rather than being determined by the world. On Haslam's approach, the boundaries of essentialist, discrete kinds and fuzzy kinds is determined by the world whereas psychiatrists determine the boundaries of practical kinds.

A different account is taken by philosopher Kendell and philosophically informed psychiatrist Jablensky in their account of validity. They argue that valid psychiatric diagnoses are those with sensitive and specific causes whereas psychiatric diagnoses without sensitive or specific causes are "arbitrary" (2003, p. 7). Though not discussing natural kinds, they

take, as Murphy suggests (2014b, p. 64), a natural kind based approach in all but name. They appear to believe that valid psychiatric diagnoses reflect objective groupings in the world whereas psychiatric diagnoses without sensitive and specific causes depend upon the views of psychiatrists rather than reflecting true groupings. This contrasts with Haslam who considers discrete kinds and fuzzy kinds to be objective groupings despite lacking sensitive and specific causes.

Cooper also takes a different approach to natural kinds. Based on Dupre's account of natural kinds, she writes that "[i]f we take cases of mental disorders as our domain and plot them onto a multidimensional quality space (as in cluster analysis) then we will find clusters of similar cases" (2012, p. 62). On this account, there are objective facts about how characteristics of entities cluster together. Despite this, there can still be more than one way in which psychiatrists can form psychiatric diagnoses. The properties of mental disorders are objectively grouped together but they are objectively grouped together in multiple different ways, depending upon which set of properties are focused upon. Psychiatrists need make a decision about which of those objective groupings to focus upon when formulating psychiatric diagnoses. This means psychiatrists have a level of choice over which psychiatric diagnoses to formulate. There might be two clusters which merge into one another, thus where the boundary of one diagnosis begins and the other ends depends upon which properties psychiatrists choose to focus upon when grouping. Whether a diagnosis is real to Cooper seems to depend upon the strength of the underlying cause of the chosen grouping and not, in contrast to Kendell & Jablensky and Haslam, on whether psychiatrists have a choice over where to put the boundaries.

Some philosophers of psychiatry have placed much less emphasis on natural kinds. For example, Murphy considers discussions of natural kinds to be relatively unimportant in psychiatry.[6] He does, however, largely focus upon the same factors of homogeneity of underlying causes and overt characteristics when assessing psychiatric diagnoses. He writes that "there are causes of mental illness [that] are genuinely out there in the structure of the world, waiting to be discovered" (Murphy 2014b,

[6] At least in this 2006 book, he places more emphasis on them in other publications (Murphy 2014b).

p. 62). On this account, causes are real parts of the world and psychiatric diagnoses are also real when they reflect those causes. Murphy's account is arguably less restrictive compared to philosophers of psychiatry who emphasise natural kinds since Murphy considered lower levels of shared cause and similarity of overt characteristics to be compatible with objectivity. Despite not endorsing natural kinds there is the same emphasis on causes and characteristics.

Philosophers of psychiatry typically see natural kinds as a means of attaining scientific legitimacy. In general, science is taken to attain legitimacy only if it describes something real rather than simply just describe the opinions of scientists. Since natural kinds are real, a scientific theory which describes natural kinds would be scientifically legitimate. More specifically, two central parts of science are inductions and causal claims. An induction is where we can predict the likelihood of a particular event occurring given that the event occurs to a particular frequency in similar situations. Scientific legitimacy requires that there is good reason to believe that the event will occur again given the same situation. Natural kinds supply this good reason. Since natural kinds are identical or similar we have good reason to expect the same event to occur when different instances of a natural kind are placed in the same situation. A causal claim is where one situation results in another situation occurring. The first situation causes the second. Scientific legitimacy requires a reason to believe that one situation results in another. As with inductions, natural kinds provide this reason. Since natural kinds are similar or identical then we can expect the same event to be caused when instances of the same natural kind are placed in the same situation (Cooper 2007, p. 45). It is also worth noting that natural kinds have also played a role in debates over notions of disorder. If psychiatric diagnoses were natural kinds then claiming that psychiatric diagnoses are also disorders would have some basis in science, although additional value judgements are needed to work out which natural kinds are disorders and which are not (Cooper 2007, 39).

This survey has addressed three of the four questions I posed earlier. Firstly, the target of psychiatric diagnoses is typically considered to be natural kinds or groupings in the world with similarity of causes and overt characteristics. Secondly, such a target potentially means psychiatric

diagnoses can be real because causes and overt characteristics are real parts of the world. Thirdly, such a target would also deliver scientific legitimacy because causes and characteristics are the sorts of things which scientists could potentially gain accurate knowledge of. However, understanding natural kinds as real and scientifically legitimate becomes more questionable the further that notions of natural kinds are moved away from essentialist notions. On an essentialist reading, both the causes and characteristics of the natural kind are determined by the objective world. The ratio to which the natural kind is formulated by human views, rather than by reality, seems to increase when notions of natural kinds move further from essentialism. Also, on an essentialist reading the same causes and characteristics are present in each instance of the entity. This means they can support inductive practises, law like generalisations and causal claims. This gives reason to consider them as scientifically legitimate. However, the further natural kinds move from essentialism means that a particular instance of a natural kind will have less shared causes and characteristics with any other instances of that natural kind. This seemingly is less amenable to inductive, law like and causal generalisations. This raises the question of why they are scientifically legitimate. The way in which philosophers of psychiatry have weakened notions of natural kinds may thus have important implications for the status of psychiatric diagnoses. This problem is compounded because, as I now show, very few currently employed psychiatric diagnoses appear to be natural kinds even under this weakened notion of natural kinds.

2.5.2 How Well Do Currently Employed Psychiatric Diagnoses Describe Their Target?

Having outlined various accounts of how philosophers of psychiatry see psychiatric diagnoses potentially targeting the real world I will now outline whether currently employed psychiatric diagnoses are considered to successfully describe that target. Generally, philosophers of psychiatry are concerned that currently employed psychiatric diagnoses are significantly flawed, failing to accurately describe the world (Kendell and Jablensky 2003, p. 5; Meehl 2001, p. 509; Murphy 2014a, p. 120).

The biggest reason for doubting the adequacy of currently employed psychiatric diagnoses is the failure to establish their causal basis. Bluhm writes that there is a "growing sense that the symptom-based categories are not the right sort of categories to guide the search for the neurobiological correlates of mental disorder" (2017, p. 146). As described above, modern research has uncovered very high levels of causal heterogeneity. This concern can take two forms. Firstly, that an alternative psychiatric diagnosis could be formulated which would not be causally heterogeneous. The causal heterogeneity which is currently detected is concealing more homogeneous causes. Psychiatric diagnoses would be superior if their causal basis were those homogeneous causes (Kendell and Jablensky 2003, p. 7). Secondly, an alternative approach would be to assume that the domain of psychiatry is actually largely composed of causal heterogeneity and that psychiatrists have already found most or all homogeneous causes already, there are not many more yet to be found. As Bolton writes, "there is not much prospect that the science of the etiology of psychiatric conditions will deliver a single, optimal classification scheme – the reason being that the last few decades of research has uncovered systemic complexity, rather than reductionist simplicity" (2012, p. 6). Consequently, psychiatric diagnoses should be assessed based on how well they fit heterogeneous causes rather than homogeneous causes. Whilst significant homogeneity might be impossible there is still concern that currently employed psychiatric diagnoses still cover individuals with too high heterogeneity (Murphy 2006, p. 204). Reducing that heterogeneity would result in superior diagnoses.

Another concern is that the theories employed within psychiatry are epistemologically weak. False or inadequate theories might cause psychiatrists to make false observations, theory influencing and thus distorting their observations (Cooper 2005, p. 77). The theoretical assumptions employed when observing symptoms could be weak, as eliminative materialists (Churchland 1979) and phenomenologists (Parnas and Sass 2008) often argue. There also could be weak theoretical assumptions being employed when symptoms are grouped together to form a diagnosis, such as the downplaying of some symptoms without good reason (Cooper 2005, p. 150). Understanding and interpretation of data relating to causation could also be weak, such as some of the speculative inferences and

uncertainty involved in neuroscience (Sullivan 2014, p. 265). Even if many psychiatric theories are not flawed, there may be flawed understanding of how all those theories should work together. Since very few psychiatric diagnoses have one simple, clear underlying cause we must employ multiple theories, being sure each theory is applied only to its relevant domain alongside needing understand how phenomena described by different theories interact, the "difficulty of which is hard to overestimate" (Kendler 2005b, p. 438).

Finally, there are concerns over what can be called non-scientific interests. This is where scientific research is influenced by issues not typically considered relevant for understanding the world. For example, there are concerns that the insurance industry and drug manufacturers have a negative influence upon psychiatric research. As Jablensky writes, psychiatry "remains easily influenced by ideological, political and market forces" (Jablensky 1999, p. 138; also Cooper 2005, p. 150; Poland 2014, p. 35). Psychiatrists might investigate some parts of the world rather than others, or interpret their findings, based upon such influences rather than purely scientific concerns.

2.5.3 The Pragmatist Alternative

Some philosophers of psychiatry have developed a pragmatist position which contrasts with the realist position of most philosophers of psychiatry who I have described. Pragmatists typically think of reality in a manner which is intended to recognise the limited point of view of humans. Either we cannot access reality as it is or reality cannot be separated from the human point of view. This means that pragmatists can take a different approach to thinking about whether psychiatric diagnoses are real or scientifically legitimate (Schaffner 2012, p. 176; Zachar 2014, p. 228).

Pragmatism is typically associated with taking truth as usefulness (James 1979 [1907]). Useful beliefs should be considered as true and there is nothing more to truth than those beliefs which are useful. Pragmatists think there are significant constraints on what counts as useful. Which scientific theories are useful partly depends upon how the world answers back to our inquiries. Despite this, the world does not determine which theories are useful and thus real. Rather, this partly depends upon a significant human element.

Pragmatist philosopher of psychiatry, Peter Zachar, sees the basis of knowledge as the world answering back in the form of sensation we receive from the external world (2014, p. 108). This is where the world constrains our belief. Anything which humans postulate above this will not be solely determined by the world but will instead depend on us. Despite this, Zachar considers things which are beyond that level to be potentially real. He describes how pragmatists see true and real as important but the very notions of true and real are themselves human centred (2014, p. 52).

To a pragmatist, psychiatric diagnoses can be potentially real on a pragmatist notion of reality. They are not part of what comes to us from the world since they cannot be found in things like sense perception. Rather, they would involve a significant contribution from us rather than from the world. They could, however, be real under a pragmatist notion of reality if they have sufficient utility. Zachar lists important factors for assessing the utility, and thus reality, of psychiatric diagnoses. He writes that

> These goals include among others scientific priorities (finding causally homogeneous kinds), measurement priorities (selecting indicators that are both sensitive and specific), professional priorities (distinguishing good from bad therapy responders), economic priorities (not treating conditions that will remit on their own), and sociopolitical priorities (reducing stigmatization). (Zachar 2014, 154–155)

Unlike realist philosophers who typically consider only similarity of underlying causes or characteristics as meaning a psychiatric diagnosis is real, for pragmatists all those factors are important but they also appeal to a wider range of factors. Additionally, none of the factors, including similarity of causes or characteristics, reveal some ultimate reality that is independent of our values and questions of utility. Also, perhaps because they typically appeal to a wider number of factors when assessing psychiatric diagnoses compared to realists, pragmatists sometimes assess currently employed psychiatric diagnoses in a relatively favourable light (Kendler 2012 p.263; Zachar 2014, p. 227). Whereas realists typically have doubts about currently employed psychiatric diagnoses, pragmatists sometimes consider some currently employed psychiatric diagnoses to be real and scientifically legitimate.

There are some important, arguably outstanding questions, for a pragmatist philosophy of psychiatry which I will now briefly mention. Firstly, when applied to psychiatric diagnoses it is unclear exactly how greatly it differs from non-essentialist realist positions. Pragmatists wholeheartedly accept that psychiatric diagnoses are partly dependent upon us rather than existing independently in the world whilst realists who reject essentialist natural kinds also accept that psychiatric diagnoses partly depend upon us. This makes it unclear how different their positions actually are. Secondly, if pragmatists are correct that the world does make a contribution, that it answers back, it is not entirely clear how it answers back or how answering back influences our knowledge. Thirdly, if humans make a significant contribution to knowledge, if we rather than the world decide which psychiatric diagnoses there are, then it is unclear that any knowledge we have should be considered scientifically legitimate. There is a concern that pragmatism undermines science.

I will forge an explicit middle point between realism and pragmatism, it will show how both the world and ourselves make a contribution to our knowledge and it will show how this can result in scientifically legitimate knowledge.

2.6 Conclusion

Psychiatry has oscillated between disease entity and non-disease entity approaches. Kraepelin effectively endorsed disease entities, they were rejected by psychoanalysts and then disease entities seemed to be triumphant with the fall of psychoanalysis and the influence of the DSM-III. However, after decades of emphasising validation in terms of sensitive and specific causes, there is growing demand to move away from disease entity approaches. This can be seen in calls for notions of validity which are not sensitive and specific and in calls for dimensional diagnoses. Both these approaches are emphasised within RDoC. Post 2000 philosophy of psychiatry has also moved, in varying degrees, from disease entities. Rather than employing essentialist notions of natural kinds which would see psychiatric diagnoses as being entities with sensitive and specific causes, weaker notions of natural kinds are employed.

Philosopher of psychiatry Kathryn Tabb suggests that philosophers may have mistakenly clung onto natural kinds even though psychiatrists are starting to move beyond them. Given the advent of projects like RDoC and HiTOP, which effectively do not endorse a natural kind approach, "diagnostic kind models [natural kinds] may no longer be the most appropriate, or the most interesting, site for analysis by philosophers of psychiatry" (2019, p. 2192). Whether philosophers of psychiatry will or should move away from seeing natural kinds as making up the domain of psychiatry is debatable. However, it is certainly the case that those favouring a natural kind approach are further pressed to justify their position given the current direction of psychiatry.

This movement away from disease entities and stronger notions of natural kinds has important consequences for thinking about psychiatric diagnoses. I outlined in the introduction chapter how critics of psychiatric diagnoses claim they are not real or scientifically legitimate because they are not real entities. Such concerns seem to be adopting a disease entity model whereby psychiatric diagnoses need be disease entities to be real and scientifically legitimate. Since currently employed psychiatric diagnoses are not natural kinds then they are not real or scientifically legitimate. Also, critics are concerned that psychiatrists are treating psychiatric diagnoses as though they are disease entities when actually they are not. What should be made of such criticisms, however, if psychiatrists are moving away from disease entities and philosophers are appealing to weaker notions of natural kinds? Does this mean that the critics have effectively won the argument and the next step should be to abandon psychiatric diagnoses? Not in the eyes of many. Although the future is unclear, projects like the RDoC and realist philosophers of psychiatry are not calling for the abandonment of psychiatric diagnoses. Rather, they hope superior psychiatric diagnoses will be produced. However, the status of these future, supposedly superior psychiatric diagnoses is unclear. To what degree will the world play a role in determining which psychiatric diagnoses are the correct ones? It seems intuitively clear that the world plays a very significant role under a disease entity and essentialist natural kind approach but what role does it play if a non-disease entity approach and a weak notion of natural kinds are employed? And would these future psychiatric diagnoses be scientifically legitimate, especially if by moving away from disease entities the world is playing a relatively small role?

In the following chapters I will employ neo-Kantianism to provide answers to all these questions. I will show that psychiatric diagnoses can be scientific legitimate if a notion of reality as *how things are* is rejected and a notion of *how things are to us* is adopted as a basis for science. Additionally, I will show that some currently employed psychiatric diagnoses could also be scientifically legitimate.

References

Aftab, Awais, Konrad Banicki, Mark Ruffalo, and Allen Frances. 2024. Psychiatric diagnoses: A clinical guide to navigating diagnostic pluralism. *Journal of Nervous and Mental Disease* 212 (8): 445–454. https://doi.org/10.1097/NMD.0000000000001791.

Aragona, Massimiliano. 2014. Validation of psychiatric classifications: The psychobiological model of personality as an exemplar. In *Alternative perspectives on psychiatric classification, DSM, ICD, RDoC, and beyond*, ed. Peter Zachar, Drozdstoj St Stoyanov, Massimiliano Aragona, and Assen Jablensky, 27–46. Oxford: Oxford University Press.

Aveline, Mark. 1999. The advantages of formulation over categorical diagnoses in explorative psychotherapy and psychodynamic management. *The European Journal of Psychotherapy, Counselling & Health* 2 (2): 199–216. https://doi.org/10.1080/13642539908402327.

Beebee, Helen, and Nigel Sabbarton-Leary. 2010. Are psychiatric kinds 'real'? *The European Journal of Analytical Philosophy* 6 (1): 11–27. https://doi.org/10.31820/ejap.

Berrios, German E., and R. Hauser. 1988. The early development of Kraeplin's ideas on classification: A conceptual history. *Psychological Medicine* 18: 813–821. https://doi.org/10.1017/S0033291700009740.

Blashfield, Roger, and Jared Keeley. 2010. A short history of a psychiatric diagnostic category that turned out to be a disease. In *Contemporary Directions in Psychopathology*, ed. Theodore Millon, Robert F. Krueger, and Erik Simonsen, 324–336. Gilford: Gilford University Press.

Bleuler, Eugen. 1950. *Dementia Praecox, or, The group of schizophrenias*. New York: International Universities Press.

Bluhm, Robyn. 2017. The need for new ontologies in psychiatry. *Philosophical Explorations* 20 (2): 146–159. https://doi.org/10.1080/13869795.2017.1312498.

Bolton, Derek. 2012. Classification and causal mechanisms: A deflationary approach to the classification problem. In *Philosophical issues in Psychiatry II: Nosology*, ed. Kenneth S. Kendler and Josef Parnas, 6–11. Oxford: Oxford University Press.

Boorse, Christopher. 1977. Health as a theoretical concept. *Philosophy of Science* 44 (4): 542–573. https://doi.org/10.1086/288768.

Borsboom, Denny, Mijke Rhemtulla, Angelique O. Cramer, Hans L.J. van der Maas, Martin Scheffer, and Conor V. Dolan. 2016. Kinds versus continua: A review of psychometric approaches to uncover the structure of psychiatric constructs. *Psychological Medicine* 46 (8): 1567–1579. https://doi.org/10.1017/S0033291715001944.

Boyd, Richard. 1991. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* 61: 127–148. https://doi.org/10.1007/BF00385837.

Brown, Timothy A., and David H. Barlow. 2005. Dimensional versus categorical classification of mental disorder in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders and beyond: Comments on the special section. *Journal of Abnormal Psychology* 114 (4): 551–556. https://doi.org/10.1037/0021-843X.114.4.551.

Churchland, Paul. 1979. *Scientific realism and the plasticity of mind*. Cambridge: Cambridge University Press.

Cloninger, Robert. 2014. Validation of psychiatric classifications: The psychobiological model of personality as an exemplar. In *Alternative perspectives on psychiatric classification, DSM, ICD, RDoC, and beyond*, ed. Peter Zachar, Drozdstoj St Stoyanov, Massimiliano Aragona, and Assen Jablensky, 201–223. Oxford: Oxford University Press.

Cooper, Rachel. 2005. *Classifying madness: A philosophical examination of the diagnostic and statistical manual of mental disorders*. Dordrecht: Springer.

———. 2007. *Psychiatry and the philosophy of science*. Stocksfield: Acumen.

———. 2012. Progress and the calibration of scientific constructs: A new look at validity. In *Philosophical issues in Psychiatry II: Nosology*, ed. Kenneth S. Kendler and Josef Parnas, 35–40. Oxford: Oxford University Press.

Cooper, Rachel, and Roger Blashfield. 2018. The myth of Hempel and the DSM-III. *Studies in History and Philosophy of Biological and Biomedical Sciences* 70: 10–19. https://doi.org/10.1016/j.shpsc.2018.05.008.

Cuthbert, Bruce N., and Thomas R. Insel. 2009. Endophenotypes: Bridging genomic complexity and disorder heterogeneity. *Biological Psychiatry* 66: 988–989. https://doi.org/10.1016/j.biopsych.2009.10.008.

———. 2013. Towards the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine* 11 (126): 1–8. https://doi.org/10.1186/1741-7015-11-126.

Decker, Hannah. 2007. How Kraepelinian was Kraepelin? How Kraepelinian are the neo-Kraepelinians? – From Emil Kraepelin to DSM-III. *History of Psychiatry* 18 (3): 337–360. https://doi.org/10.1177/0957154X07078976.

Demazeux, Steeves, and Patrick Singy. 2015. Introduction. In *The DSM-5 in perspective: Philosophical reflections on the psychiatric Babel*, ed. Steeves Demazeux and Patrick Singy, i–xxiv. Dordrecht, the Netherlands: Springer.

Dupré, John. 1993. *The disorder of things: Metaphysical foundations of the disunity of science*. Massachusetts: Harvard University Press.

Ellis, Brian. 2001. *Scientific essentialism*. Cambridge: Cambridge University Press.

Faucher, Luc, and Simon Goyer. 2015. RDoC: Thinking outside the DSM box without falling into a reductionist trap. In *The DSM-5 in perspective: Philosophical reflections on the psychiatric Babel*, 199–224. Netherlands: Springer.

Feighner, John, Eli Robins, Samuel B. Guze, Robert A. Woodruff, George Winokur, and Rodrigo Munoz. 1972. Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry* 26 (1): 57–63. https://doi.org/10.1001/archpsyc.1972.01750190059011.

Fellowes, Sam. 2021. How autism shows that symptoms, like psychiatric diagnoses, are 'constructed': Methodological and epistemic consequences. *Synthese* 199 (1): 4499–4522. https://doi.org/10.1007/s11229-020-02988-3.

———. 2024. A history of childhood schizophrenia and consequences for autism. *History and Philosophy of the Life Sciences* 46 (29). https://doi.org/10.1007/s40656-024-00627-5.

Ghaemi, S. Nassir. 2009. Nosologomania: DSM & Karl Jaspers' critique of Kraepelin. *Philosophy, Ethics, and Humanities in Medicine* 4 (10): 1–8. https://doi.org/10.1186/1747-5341-4-10.

Hacking, Ian. 1991. A tradition of natural kinds. *Philosophical Studies* 61: 109–126. https://doi.org/10.1007/BF00385836.

Haslam, Nick. 2002. Kinds of kinds: A conceptual taxonomy of psychiatry categories. *Philosophy, Psychiatry, and Psychology* 9 (13): 203–217. https://doi.org/10.1353/ppp.2003.0043.

Haslam, Nick, Melanie JMcGrath, Wolfgang Viechtbauer, and Peter Kuppens. 2020. Dimensions over categories: A meta-analysis of taxometric research. *Psychological Medicine* 50 (9): 1418–1432. https://doi.org/10.1017/S003329172000183X.

Heckers, Stephan, Eric J. Engstrom, and Kenneth S. Kendler. 2021. "Manifestations of insanity": Kraepelin's final views on psychiatric nosology

in their historical context. *Molecular Psychiatry* 27: 328–334. https://doi.org/10.1038/s41380-021-01232-9.

Heinrichs, Douglas. 2015. Model-based science and the ethics of ongoing treatment negotiation. In *The Oxford handbook of psychiatric ethics*, ed. John Z. Sadler, Werdie Van Staden, and K.W.M. Fulford. Oxford: Oxford University Press.

Helzer, John, Helena C. Kraemer, and Robert F. Krueger. 2006. The feasibility and need for dimensional psychiatric diagnoses. *Psychological Medicine* 32: 1671–1680. https://doi.org/10.1017/S003329170600821X.

Hoffman, Ginger, and Peter Zachar. 2017. RDoC's metaphysical assumptions: Problems and promises. In *Extraordinary Science and Psychiatry*, ed. Jeffery Poland and Şerife Tekin, 59–86. Massachusetts: MIT Press.

Hopwood, Christopher, Leslie C. Morey, and Kristian E. Markon. 2023. What is a psychopathology dimension? *Clinical Psychology Review* 106: 102356. https://doi.org/10.1016/j.cpr.2023.102356.

Horwitz, Allan. 2002. *Creating mental illness*. University of Chicago Press.

Hudziak, James, Thomas M. Achenbach, Robert R. Althoff, and Daniel S. Pine. 2007. A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research* 16 (S1): 16–23. https://doi.org/10.1002/mpr.217.

Hyman, Steven. 2010. The diagnosis of mental disorder: The problem of reification. *Annual Review of Clinical Psychology* 6: 155–179. https://doi.org/10.1146/annurev.clinpsy.3.022806.091532.

Jablensky, Assen. 1999. The nature of psychiatric classification: Issues beyond ICD-10 and DSM-IV. *Australian and New Zealand Journal of Psychiatry* 33: 137–144. https://doi.org/10.1046/j.1440-1614.1999.00535.x.

———. 2016. Psychiatric classifications: Validity and utility. *World Psychiatry* 15 (1): 26–31. https://doi.org/10.1002/wps.20284.

James, Williams. 1979. *Pragmatism*. Cambridge, MA: Harvard University Press.

Jaspers, Karl. 1997. *General psychopathology*. Vol. 2. JHU Press.

Johnstone, Lucy. 2014. Controversies and debates about formulation. In *Formulation in psychology and psychotherapy: Making sense of people's problems*, ed. Lucy Johnstone and Rudi Dallos, 2nd ed., 260–289. Hove: Routledge.

———. 2018. Psychological formulation as an alternative to psychiatric diagnosis. *Journal of Humanistic Psychology* 58 (1): 30–46. https://doi.org/10.1177/0022167817722230.

Kendell, Robert, and Assen Jablensky. 2003. Distinguishing between the validity and utility of psychiatric diagnosis. *American Journal of Psychiatry* 35: 139–144. https://doi.org/10.1176/appi.ajp.160.1.4.

Kendler, Kenneth. 2005a. "A Gene for…": The nature of gene action in psychiatric disorders. *American Journal of Psychiatry* 162: 1243–1252. https://doi.org/10.1176/appi.ajp.162.7.1243.

———. 2005b. Towards a philosophical structure for psychiatry. *American Journal of Psychiatry* 162: 433–440. https://doi.org/10.1176/appi.ajp.162.3.433.

———. 2012. Phenomenology, nosology and prototypes. In *Philosophical issues in Psychiatry II: Nosology*, ed. Kenneth S. Kendler and Josef Parnas, 260–264. Oxford: Oxford University Press.

Kinderman, Peter, John Reed, Joanna Moncrieff, and Richard P. Bentall. 2013. Drop the language of disorder. *Evidence-Based Mental Health* 16 (1): 1–3. https://doi.org/10.1136/eb-2012-100987.

Kirk, Stuart, David Cohen, and Tomi Gomory. 2015. DSM-5: The delayed demise of descriptive diagnosis. In *The DSM-5 in perspective: Philosophical reflections on the psychiatric Babel*, 63–81. Netherlands: Springer.

Kotov, Roman, Robert F. Krueger, David Watson, Thomas M. Achenbach, Robert R. Althoff, R. Michael Bagby, Timothy A. Brown, William T. Carpenter, Caspi Avshalom, Lee Anna Clark, Nicholas R. Eaton, Miriam K. Forbes, Kelsie T. Forbush, David Goldberg, Deborah Hasin, Steven E. Hyman, Masha Y. Ivanova, Donald R. Lynam, Kristian Markon, Joshua D. Miller, Terrie E. Moffitt, Leslie C. Morey, Stephanie N. Mullins-Sweatt, Johan Ormel, Christopher J. Patrick, Darrel A. Regier, Leslie Rescorla, Camilo J. Ruggero, Douglas B. Samuel, Martin Sellbom, Leonard J. Simms, Andrew E. Skodol, Tim Slade, Susan C. South, Jennifer L. Tackett, Irwin D. Waldman, Monika A. Waszczuk, Thomas A. Widiger, Aidan G.C. Wright, and Mark Zimmerman. 2017. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology* 126 (4): 454–477. https://doi.org/10.1037/abn0000258.

Kotov, Roman, Rober F. Krueger, and David Watson. 2018. A paradigm shift in psychiatric classification: The Hierarchical Taxonomy of Psychopathology (HiTOP). *World Psychiatry* 17 (1): 24–25. https://doi.org/10.1002/wps.20478.

Kozak, Michael J., and Bruce N. Cuthbert. 2016. The NIMH Research Domain Criteria Initiative: Background, issues, and pragmatics. *Psychophysiology* 53: 286–297. https://doi.org/10.1111/psyp.12518.

Kraemer, Helena. 2007. DSM categories and dimensions in clinical and research contexts. *International Journal of Methods in Psychiatric Research* 16 (S1): S8–S15. https://doi.org/10.1002/mpr.211.

Kraemer, Helena C., Art Noda, and Ruth O'Hara. 2004. Categorical versus dimensional approaches to diagnosis: Methodological challenges. *Journal of Psychiatric Research* 38 (1): 17–25. https://doi.org/10.1016/S0022-3956(03)00097-9.

Kripke, Saul. 1972. *Naming and necessity, semantics of natural language*. Dordrecht: Reidel.

Markon, Kristian, Michael Chmielewski, and Christopher J. Miller. 2011. The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin* 137 (5): 856. https://doi.org/10.1037/a0023678.

Maung, Hane. 2019. The function of diagnoses in medicine and psychiatry. In *The Bloomsbury companion to philosophy of psychiatry*, ed. Şerife Tekin and Roybn Bluhm, 507–526. London: Bloomsbury.

Meehl, Paul. 2001. Comorbidity and taxometrics. *Clinical Psychology: Science and Practice* 8 (4): 507–519. https://doi.org/10.1093/clipsy.8.4.507.

Moncrieff, Joanna, and David Cohen. 2005. Rethinking models of psychotropic drug action. *Psychotherapy and Psychosomatics* 74 (3): 145–153. https://doi.org/10.1159/000083999.

Murphy, Dominic. 2006. *Psychiatry in the scientific image*. Massachusetts: Massachusetts Institute of Technology.

———. 2014a. Natural kinds in fold psychology and in psychiatry. In *Classifying psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 105–122. Massachusetts: MIT Press.

———. 2014b. Validity, realism and normativity. In *Alternative perspectives on psychiatric classification, DSM, ICD, RDoC, and beyond*, ed. Peter Zachar, Massimiliano Aragona, Drozdstoj St Stoyanov, and Assen Jablensky, 60–75. Oxford: Oxford University Press.

Okasha, Ahmed. 2009. Would the use of dimensions instead of categories remove problems related to subthreshold disorders? *European Archives of Psychiatry and Clinical Neuroscience* 259: 129–133. https://doi.org/10.1007/s00406-009-0052-y.

Orgota, Frances. 2013. Cerebralizing Autism with the neurodiversity movement. In *Worlds of Autism*, ed. Joyce Davidson and Michael Orsini, 73–96. Minneapolis: University of Minnesota Press.

Parnas, Josef, and Louis A. Sass. 2008. Varieties of 'phenomenology': On description, understanding, and explanation. In *Philosophical issues in psychiatry: Explanation, phenomenology and nosology*, ed. Kenneth S. Kendler and Josef Parnas, 239–278. Baltimore: Johns Hopkins University Press.

Poland, Jeffrey. 2014. Deeply rooted sources of error and bias in psychiatric classifications. In *Classifying psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 29–64. Massachusetts: MIT Press.

Potuzak, Mellisa, Caitlin Ravichandran, Kathryn E. Lewandowski, Dost Ongür, and Bruce Cohen. 2012. Categorical vs dimensional classifications of psychotic disorders. *Comprehensive Psychiatry* 53 (8): 1118–1129. https://doi.org/10.1016/j.comppsych.2012.04.010.

Psillos, Stathis. 1999. *Scientific realism: How science tracks truth*. Oxon: Routledge.

Putnam, Hilary. 1975. The meaning of "meaning". *Minnesota Studies in the Philosophy of Science* 7: 131–193.

Robins, Eli, and Samuel B. Guze. 1970. Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry* 126 (7): 983–987. https://doi.org/10.1176/ajp.126.7.983.

Sadler, John. 2005. *Values and psychiatric diagnosis*. Oxford: Oxford University Press.

Schaffner, Kenneth. 2012. A philosophical overview of the problems of validity for psychiatric disorders. In *Philosophical issues in Psychiatry II: Nosology*, ed. Kenneth S. Kendler and Josef Parnas, 169–189. Oxford: Oxford University Press.

Shorter, Edward. 1997. *A short history of psychiatry*. Wiley.

Smith, Gregory, and Jessica Combs. 2010. Issues of construct validity in psychological diagnoses. In *Contemporary Directions in Psychopathology*, ed. Theodore Millon, Robert F. Krueger, and Erik Simonsen, 205–222. Gilford: Gilford University Press.

Spitzer, Robert, and Joseph L. Fleiss. 1974. A re-analysis of the reliability of psychiatric diagnosis. *The British Journal of Psychiatry* 125 (587): 341–347. https://doi.org/10.1192/bjp.125.4.341.

Spitzer, Robert, Jean Endicott, and Eli Robins. 1978. Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry* 35 (6): 773–782. https://doi.org/10.1001/archpsyc.1978.01770300115013.

Sullivan, Jacqueline. 2014. Stabilizing mental disorders: Prospects and problems. In *Classifying psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 257–282. Massachusetts: MIT Press.

Tabb, Kathryn. 2019. Philosophy of psychiatry after diagnostic kinds. *Synthese* 196: 2177–2195. https://doi.org/10.1007/s11229-017-1659-6.

Tekin, Şerife. 2016. Are mental disorders natural kinds?: A plea for a new approach to intervention in psychiatry. *Philosophy, Psychiatry, and Psychology* 23 (2): 147–163. https://doi.org/10.1353/ppp.2016.0013.

van Praag, Hermann. 2008. Kraepelin, biological psychiatry, and beyond. *European Archives of Psychiatry and Clinical Neuroscience* 258: 29–32. https://doi.org/10.1007/s00406-008-2006-1.

Vanheule, Stijn. 2017. *Psychiatric diagnosis revisited: From DSM to clinical case formulation*. Cham: Palgrave Macmillan.

Wakefield, Jerome. 2022. Klerman's "credo" reconsidered: Neo-Kraepelinianism, Spitzer's views, and what we can learn from the past. *World Psychiatry* 21 (1): 4–25. https://doi.org/10.1002/wps.20942.

Widiger, Thomas, and Douglas B. Samuel. 2005. Diagnostic categories or dimensions? A question for the diagnostic and statistical manual of mental disorders. *Journal of Abnormal Psychology* 114 (4): 494. https://doi.org/10.1037/0021-843X.114.4.494.

Zachar, Peter. 2014. *A metaphysics of psychopathology*. Massachusetts: Massachusetts Institute of Technology Press.

Zachar, Peter, and Assen Jablensky. 2014. Introduction: The concept of validation in psychology and psychiatry. In *Alternative perspectives on psychiatric classification, DSM, ICD, RDoC, and beyond*, ed. Peter Zachar, Drozdstoj St Stoyanov, Massimiliano Aragona, and Assen Jablensky, 27–46. Oxford: Oxford University Press.

Zachar, Peter, and Kenneth S. Kendler. 2007. Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry* 164 (4): 557–565. https://doi.org/10.1176/ajp.2007.164.4.557.

3

Neo-Kantianism

3.1 Introduction

In this book I shall employ a neo-Kantian approach to build a new way of understanding psychiatric diagnoses. In this chapter I shall give an overview of Kantianism. I then apply this new understanding to psychiatric symptoms, psychiatric diagnoses and causation in psychiatry in later chapters. I will sometimes talk of Kantianism and sometimes talk of neo-Kantianism. By Kantianism I mean someone who holds general principles which were inspired by the work of Immanuel Kant. By neo-Kantianism I refer to a set of specific thinkers who I draw upon to build the position I adhere to. This demarcation is certainly imperfect but I hope to convey the difference between adhering to a general set of principles and set of more specific philosophical positions which aim to manifest those principles (there are other potential ways of manifesting those general principles, such as German idealism, which I would not endorse).

Kantians typically believe that there are parts of the mind-independent world that are inaccessible to us. We can only generate knowledge by applying something to the world which is not present in the world. That thing which is applied to the world comes from us. It is what makes

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_3

knowledge possible. As such, all knowledge in some way depends on us rather than reflecting the mind-independent world. In this chapter I shall outline various views Kantians have held on what comes from us when formulating knowledge. This gives a picture of how we go about formulating scientific knowledge which will elucidate what should be considered scientific knowledge. Crucially, on a Kantian account what counts as scientific knowledge does *not* depend on knowledge claims reflecting the mind-independent world. Kantians employ other standards, which I shall outline, when demarcating good science from bad science. Employing this Kantian framework means that, as I show in later chapters, psychiatric diagnoses need not describe actual people to be considered good science.

Philosophers holding a significant variety of views have called themselves Kantians or neo-Kantians. These views often diverge significantly from one another and from the views of Kant. Additionally, Kant's position can be interpreted in multiple ways. As such, I will outline different views and show in what sense they aim to be distinct from realism, anti-realism, empiricism, pragmatism, operationalism and constructivism. I will also outline which views I favour.

3.2 Immanuel Kant

Kant (1742–1804) was born and lived his life in Königsberg, then part of Prussia, now part of Russia (see Kuehn 2001 for bibliography). He spent most of his career as a professor at Königsberg University. He wrote on an extensive range of topics. I focus upon his epistemology, metaphysics and philosophy of science but he also wrote extensively on many other topics such as ethics, political philosophy, physics, cosmology and anthropology.

Kant was reacting to, and trying to find a middle ground between, two seemingly opposing philosophical traditions of rationalism and empiricism (Dudley and Engelhard 2014, p. 3; Gardner 1999, p. 1). On one hand, Kant is responding to rationalists who believe that all knowledge ultimately stems from thought. For rationalists, sensation is not required to attain knowledge of the world. Kant felt this led to metaphysical excesses, leading philosophers to make wide ranging claims about the

nature of God, the soul and the fundamental nature of substance based upon a few non-empirical principles. Kant had severe reservations about the possibility of this approach giving genuine knowledge of the world. He was also responding to empiricists who believed that all knowledge came from the senses. Kant was concerned that this knowledge was without any firm foundation. Many crucial elements, such as causation, on which our knowledge depends are not found in sensation. If key aspects of knowledge depend on things which are not found in sensation then this would undermine our ability to know the world. Kant proposed a position which involved both empiricism and rationalism. It is partly rationalist because all knowledge of the external world depends on us but it is partly empiricist because what comes from us needs to be applied to perception to generate knowledge of the world. He hoped to attain the security given by rationalism and the relevance to the world given by empiricism.

To do this Kant employed the notion of the *synthetic a priori*. To understand this we need to introduce some technical terminology. Knowledge which is not based upon experience is *a priori* and knowledge which is based upon experience is *a posteriori*. Typically, *a priori* knowledge is understood as *analytic*. An *analytic* statement is one which is true in virtue of its meaning. For example, the claim that 'all bachelors are unmarried' is *analytic* because the meaning of 'bachelor' is the same as 'unmarried'. Also, we know that 'all bachelors are unmarried' through *a priori* means because we do not need experience to determine if the statement is true (the statement would be true even if everyone throughout human history was married or if no one ever got married). In contrast, *a posteriori* statements, which are based upon experience, are typically understood as *synthetic*. A *synthetic* statement is true or false depending upon the way the world is, rather than just upon its meaning. For example, the statement that 'guinea pigs like eating cucumber' is *synthetic* because we need make observations of the external world to determine if it is true. So there seems to be a clear link between *analytical* (truth based upon meaning) and *a priori* (knowledge without experience). There also seems to be a clear link between *synthetic* (truth based upon the world) and *a posteriori* (knowledge based upon experience). Kant's innovation was to propose *synthetic* (truth based upon the world) *a priori* (knowledge

without experience) knowledge. Kant argues that we have something *a priori* which is not based upon experience but which is applied to experience. Something that is independent of experience is applied to experience to give experience a form.

Kant argues that we can have knowledge of the external world but that the knowledge is limited and it partly depends upon us. Knowledge of the external world comes "from the human standpoint" (Kant 1998, A27/B43).[1] Knowledge is limited because there are domains which our knowledge cannot extend to. Kant famously makes the distinction between *how things are in themselves* and *how things appear to us* (Kant 1998, A27/B43 & A35/B52). This distinction is often taken to mirror his distinction between phenomena and noumena but philosophers have questioned how closely they map onto one another so I will avoid talking of Kant's notion of phenomena and noumena.[2] Kant argues that how things are in themselves is different to how they appear to us. How something is in itself is how something is when considered independently from us (or any conscious being except God). Since humans can only consider things from the human perspective we can have no knowledge of how things are in themselves.[3]

What then is this human perspective and what exactly does it mean that things *appear to us*? Kant argues that many things which humans typically assume to exist in the external world actually come from us (Kant 1998, A646/B674; see also Ameriks 1992, p. 259; Guyer 1980, p. 207; Ostaric 2009, p. 161; Rubenstein 2001, p. 85). His clearest example is when he argues that things in themselves do not exist in space and time. Rather, space and time are imposed upon the world and we cannot help but see them as existing in space and time. Space and time

[1] I use typical conventions for citing Kant, namely: giving the A/B edition paragraph numbers of the *Critique of Pure Reason* (1998) and the Academy Edition page number for both the *Critique of the Power of Judgement* (2000) and the *Prolegomena* (2002).

[2] For example, Langton (1998) takes these as identical, whereas for Allison (1983) noumena are how things are thought of rather than something external.

[3] Strictly speaking, the claim that we know nothing about things in themselves might be not fully accurate. We arguably do know some things about them such as that things in themselves exist and that we know nothing of them beyond that they exist.

are concepts we apply to the world.[4] In the words of a contemporary philosopher, Kant thinks that the "fundamental forms for sensory representation" stem from the "structure of our own minds" (Guyer 2006a, p. 50; see also Buchdahl 1967, p. 211; Kant 1998, A26/B43 & A34/B51 & A80/B106; Pilot 1990, p. 123). Any attempt to see things which exist outside us will require imposing space and time upon them. Consequently, we cannot see them *as they are in themselves* but only *as they appear to us*. Kant extends the idea that we impose forms upon the world to many other areas other than space and time. He often claims that causation, matter, substance, laws and objects are all imposed upon the world (Kant 1998, A23/B38 & A92/B124; see also Banham 2013, p. 19; Forster 2008, p. 43). The meaning of these claims is more controversial compared to his discussion of space and time. For example, Kant often talks of how causes and objects are projected onto the world rather than existing in the external world but on very rare occasions he, seemingly contradictorily, tentatively suggests there may be objects which are independent of what we project and they do have a causal influence on us (Kant 1998, A496/B524).

Kant sees the faculties of our mind as providing a standard of truth. The application of these faculties means that we can never see things as they are in themselves. Any standard of truth that we employ cannot be based around how things are in themselves. Additionally, we cannot know of anything except through those faculties. The raw perception that we receive from the world is not knowledge since it does not contain many important elements of knowledge such as space, time and causation. Rather, by structuring raw perception to produce sense data, bringing sense data under concepts to produce entities and causes, these faculties *make knowledge possible* (Kant 1998, Bxviii; Butts 1984, p. 7; Rubenstein 2001, p. 85). This means that we cannot use 'how things are independent of our concepts' as a standard of truth. Kant says that things in themselves are "nothing for us" (Kant 2002, 4:336). Rather, any

[4] Strictly speaking, Kant considered space and time as intuitions rather than concepts. This distinction is not usually accepted by modern Kantians and I shall not mention it again beyond in future footnotes.

standard of truth must involve our faculties because we cannot know anything outside of them.

It is important to clarify what Kant is committing himself to. Kant is generally taken to believe that we bring what comes to us from perception under notions like laws, causes, matter, substance and objects. However, there is disagreement over what, if anything, Kant thinks actually exists outside of us. There are many statements in Kant that looks like he thinks anything external to us are simply ideas, meaning he is effectively just holding an idealism in the sense of Bishop Berkeley (Wilkerson 1976, p. 28).[5] Kant heavily denied he thought the external world was simply ideas but those following Kant typically took him as holding some strong type of idealism, and then either criticised or broadly endorsed Kant on these grounds. This, however, is not a common interpretation of Kant by modern Kantians. They typically assume that Kant thought there was something external to the human mind which was not simply ideas, though Kant also thought that external thing was in some sense inaccessible.

It is also important to counter some stereotypical but false views of Kant. Kant is stereotypically pictured as believing that the human mind structures the world in an unconscious manner purely from *a priori* principles. On this understanding humans have no ability to control or influence this process. However, Kant is very explicit that we do have full control over some of the ways we apply principles to structure the world. Kant posits that we employ the faculty of the understanding to structure the world into entities and causes. We then employ the faculty of reason to posit different *types of* entities and causes (Kant 2000, 5:188; Butts 1990, p. 2; Ostaric 2009, p. 165; Meerbote 1991, p. 133; Neiman 1994, p. 60). The faculty of the understanding requires that we see the world in terms of entities but we need to employ something more, the faculty of reason, to say that those entities take the form of electrons or beetles. Kant believed that the faculty of reason involves employing principles which are based in reason, describing them as having a "logical objective

[5] Bishop Berkeley was an idealist philosopher who believed that to exist is to be perceived, whereby objects in the external environment around us existed in the form of non-physical ideas in the mind of God.

necessity" (2000, 5:182 see also Cassirer 1953a [1910], p. 269; Cassirer 1957 [1929], p. 361; Natorp 2004 [1921], p. 348; Buchdahl 1969, p. 516; Kitcher 1986, p. 229; Morrison 1989, p. 162; Reichenbach 1965 [1920], p. 53; Rickert 1986 [1902], p. 227; Windleband 2015a [1910], p. 32; Vaihinger 1935 [1911], p. 159). It is the case that we must employ concepts like causation and substance but we can use those principles of reason to decide *which particular* causes and substances there are. We have no choice over structuring the world into causes and substances but we do have a choice over how to employ principles when deciding what causes and substances there are. Additionally, because those principles of reason are rooted in reason they are "the *touchstone of truth*" (Kant 1998, A647/B675). They provide a standard for assessing truth from falsity.

This summary of Kant has led to five key claims. Firstly, how things are in themselves is inaccessible to us. There are things which we can never know. Secondly, all knowledge of the external world is based in how things appear to us. Thirdly, all knowledge must come through various faculties. Fourthly, we have a level of control over how some of those faculties generate knowledge. Fifthly, those faculties provide a standard of truth for Kant.

3.3 Historical Interpretations of Kant

There have been many different interpretations of Kant. Before outlining my own approach it is helpful to briefly outline some varying historical interpretations of Kant to show different ways Kant has been understood. German philosophy post Kant largely split into the extremes of an idealistic a priorism, in the form of Hegel and German Romanticism, and a positivistic empiricism (Köhnke 1991, p. 62). This influenced a call to return to Kant who was seen as a middle ground between these. The advantage of idealism was that it emphasised the role of the mind in the formulation of objects of our perception but it also struggled to account for how the external world influences our perceptions. The advantage of a positivist materialism was that it placed great emphasis on the external world but it struggled to account for how our minds shaped our

perceptions. Kant's notion of bringing perceptions under concepts was considered a middle ground between these positions and 1850 saw the first neo-Kantian book published which attempted to merge idealism and materialism (Lange's *History of Materialism*). This eventually led to the "neo-Kantian tradition, which was without a doubt the most broadly influential movement of European philosophy between approximately 1850 to 1918" (Luft 2015a, p. xx; see also Heis 2018; Köhnke 1991, p. 280) (I thank an anonymous reviewer for suggesting that this claim about the influence of neo-Kantianism could be exaggerated since potentially philosophers like Kierkegaard, Schopenhauer, Nietzsche and Marx were equally or more influential in post-Hegelian German philosophy). However, during this period many neo-Kantians significantly developed Kant's position in ways which often did not fully respect Kant's original position (see Glock 2015, p. 63; Luft 2015a, p. xxi). Indeed, Köhnke remarks upon "how narrowly circumscribed was the interest in Kant felt by so many 'neo-Kantians'" (1991, p. 163). Those advocating for a return to Kant stated that "understanding Kant means going beyond Kant" (Windelband in Kinzel 2021, p. 579).

One popular interpretation in the second half the 1850s was to take a scientific interpretation of Kant (Luft 2015a, p. xxii; Köhnke 1991, p. 234). On this approach, which is typically rejected by modern neo-Kantians, the faculties of the mind which Kant posits should be understood in scientific terms. For example, Lange's *History of Materialism*, which was the most well read of the early neo-Kantians, took the faculties as being biological in nature, believing that our knowledge was constrained by the biological nature of our sense organs. Others saw these faculties as psychological in nature, in a somewhat analogous manner to how cognitive psychologists often see psychological mechanisms as hard-wired into us. Additionally, some took a Darwinian evolutionary account, believing those biological or psychological faculties were the product of evolution.

However, such scientific understandings of the *a priori* arguably lack sophistication because we need to employ those faculties to gain biological and psychological knowledge. We cannot establish neutral biological or psychological knowledge which is independent of the faculties of the mind to then describe those faculties. Such accounts were rejected by the

two main schools behind neo-Kantianism, the Marburg school and the Baden (or Southwestern) school (Luft 2015a, p. xxii).[6] These were major philosophical movements in mainland Europe between around 1870 and 1920 (Makkreel and Luft 2010, p. 1; see also Köhnke 1991, p. 280). Both believe that what came from us ultimately stemmed from reason rather than our biological or psychological capacities. They believed that seeing the faculties as biological or psychological would just be a subjectivist position whereas they felt that Kantianism entailed some type of objectivity (Cassirer 1957 [1929], p. 400; Natorp 2015a [1912], p. 183; Windleband 2015b [1883], p. 283).

Each school took a different approach. The Marburg school took an approach whereby they demarcated a specific fact and then aimed to find the contribution from us which made that fact possible. The specific fact was the best scientific theories in any particular field (Cassirer 1957 [1929], p. 408; Krijnen 2015, p. 118; Ferrari 2015, p. 264; Luft 2010, p. 63; Kühn 2010, p. 113; Natorp 2015a [1912], p. 182). They believed that for science the principles could be traced back to reason. For example, whichever mathematical principles make possible theories in physics is what comes from us and these are based in reason.

The Baden school took a different approach, believing that values are required for achieving scientific knowledge. They thought that even the most basic perceptions possible which we could be aware of, such as structuring perceptions into space and time, were value-laden (Beiser 2009, p. 14; Centi 2015, p. 135; Crowell 2010, p. 152; Rickert 1986 [1902], p. 206; Windleband 2015b [1883], p. 280). To accomplish certain ends in science certain values were needed and reason establishes the required values. Both schools thought that biological and psychological interpretations of Kant entailed subjectivity and instead they sought to give Kantianism an objective foundation by seeing what comes from us as based in reason.

I think there is much to be learned from both schools and I will draw upon both. Ultimately, however, I favour the Marburg school. Whilst I find Rickert, a key member of the Southwestern school, provides an

[6] Note that there were neo-Kantians during this period who did not fall into either school, such as Vaihinger.

excellent account of abstraction and idealisation in the natural and historical science, deeper problems with his philosophy were unearthed by his student Emil Lask. Rickert posits two different domains, that of values and that of facts. This then raises the question of how values connect with facts. Lask argues that Rickert cannot bridge the gap between values and facts. This is not a question the Marburg school faces because they only posit one domain, that of experience and the *a priori* principles which make experience possible. Lask aimed to produce a neo-Kantian alternative to Rickert. Unfortunately, he died aged forty during the Great War. As such, his account is very much incomplete. Also, very little of his work has been translated into English, and when reading those few works I can only agree with Beiser who describes Lask as "formidably difficult, not only in substance but also in style" (Beiser 2008, p. 284). My understanding from the secondary literature is that Lask significantly deviated from the Kantian project. To bridge the divide between value and fact, Lask started to see values as being structured by objects (Beiser 2008, p. 293). Also, Lask thinks that we can have non-conceptual knowledge of objects (Beiser 2008, p. 293). Both these claim seem the opposite of Kant's key insight that we cannot know of objects except through what we apply to them and that what we apply comes from us rather than coming from the objects. As Beiser writes, "Lask was stretching Kant to the breaking point" (Beiser 2008, p. 291). Given the problems with Southwestern school's divide between values and facts, and given that Lask's solution to them seems incompatible with key Kantian notions, I find myself inclined towards the Marburg camp.

The neo-Kantian movement fell out of popularity in the 1920s (Makkreel and Luft 2010, p. 7). This is partly due to the perceived problems with Kantianism arising from Einstein's work. Kant is often taken as believing that what we apply to the world entails Newtonian physics but Einstein showed that there is an alternative to Newton. To some observers this was fatal to Kantianism. Also, around this time new philosophies were gaining in popularity, such as phenomenology and logical positivism, whilst the split between analytical and continental philosophy harmed the neo-Kantian movement because it covered both approaches.

3.4 Forging a Neo-Kantian Middle Ground

Given that few Kantians or neo-Kantians fully subscribe to Kant's exact claims it is worth outlining a general Kantian position, showing how Kantianism can be considered a middle ground between realism and relativism. I will also outline the specific neo-Kantian position which I subscribe to.

Firstly, Kantianism is a human centred position. Knowledge of the world is in some way limited by us. We can imagine that there could be a being, such as a God, who could see the world as it is (Kant 2002, 5:407; Butts 1990, p. 11; Guyer 2006b, p. 438) but humans cannot access that view of the world. Instead, our knowledge of the external world is limited by something that we contribute to knowledge. What we contribute is not present in the world so we cannot see the world as it is but instead we can only see it as it appears to us. There is considerable disagreement over what exactly we contribute to the world. As discussed above, some have seen what comes from us as being our biological, physiological and psychological capacities. Others have argued that what comes from us is not based upon such capacities but rather stem from fundamental principles of reason. Other philosophers believe that both these approaches are possible, that a Kantian should see some of what we contribute as coming from our psychological capacities and other parts from reason ([Heinz] Cassirer 1978, p. 64; Kitcher 2007, p. 52; Schlick 2015 [1921], p. 453). Secondly, Kantians are typically taken as believing that some type of truth and objectivity is possible (see Buchdahl 1969, p. 631; Cassirer (1957 [1929], p. 385; Natorp 2015b [1887], p. 165; Rickert 1986 [1902], p. 227; Windleband 2015a [1910], p. 323). Some or all of what we contribute to the world involves some type and level of truth and objectivity.[7] They will not be truths about some mind-independent reality of how things are in themselves. Despite this, some

[7] Some neo-Kantians not only spoke of truth and objectivity as being possible but also an accurate understanding of reality as being possible. However, this is very much a redefinition of reality, whereby what is considered real is not the objects in the world around us but an ideal structure of objects that is based in reason. Additionally, other neo-Kantians would speak of reality as being what we see in the world around us, whereby the aim of science was precisely not to describe that reality but to move beyond it.

or all of what comes from us is not seen as being subjective or simply conventions. It is considered to have a higher status than this.[8] Additionally, the concepts which come from us are thought of as legitimating scientific knowledge. Science is not seen as arbitrary, subjective, or based on conventions because what we contribute when formulating scientific knowledge involves some level of objectivity.

How exactly can Kantians hold a position which sees things in themselves as inaccessible whilst also seeing scientific theories as potentially true and objective? This is a complicated and controversial question but here I shall outline two possible reasons. Firstly, Kant recognised that humans have an innate desire to seek the unconditioned truth, that is, the things in themselves which are independent of our concepts. He thought that philosophy, through a literal critique of reason, could provide an antidote to such an untenable goal but at the same time he recognised that humans need to talk of notions of truth. Consequently, on this interpretation Kant redefined notions of truth whereby things are only true within the human perspective (see Buchdahl 1969, p. 640; Butts 1984, p. 141; Grier 1997, p. 17; O'Neill 1992, p. 291; Okruhlik 1986, p. 318). From our perspective, some scientific claims should be considered true and others considered false, even if no scientific claim resembles the things in themselves. Secondly, there are principles which we can employ when formulating scientific theories which give scientific theories an elevated status. For example, simplicity is a common standard for judging scientific theories on. When comparing two theories then, all other things being equal, the simpler one should be considered more truthful. On this interpretation, such standards relate to truth and objectivity because those standards stem from, or are entailed by, reason (Buchdahl 1969, p. 516; Cassirer 1953a [1910], p. 18; Cassirer 1957 [1929], p. 447; Kitcher 1986, p. 229; Morrison 1989, p. 162; Natorp 2004 [1921], p. 348; Reichenbach 1965 [1920], p. 104; Rickert 1986 [1902], p. 227; Windleband 2015a [1910], p. 323). This means those standards are objective and so entail truth. On all these accounts, what is

[8] It is harder to understand our contribution as being objective if that contribution is taken to be our biological, physiological and psychological capacities rather than stemming from reason. As such, Kantians are typically closer to the latter position.

considered truthful is not a mere convention. Rather, we have reason to consider some claims as truthful and other claims as mere conventions.

Combining these points together, Kantianism provides an alternative way to think about truth and objectivity. A scientific claim does not need to reflect how things are in themselves, does not need to resemble a God's eye view of the world, for it to be objective or be legitimate. A scientific claim can involve some contribution from us and still be objective and be legitimate.

I endorse the second option listed above, namely that reason provides a set of standards for formulating scientific claims and those standards deliver objectivity because reason should be considered objective and a standard of truth. Scientists can produce idealised models which idealise and transform much of what we receive from the world. Additionally, scientists can model in very different ways, meaning that scientists can produce quite diverse models to cover the same part of the world. However, we can use principles which stem from reason to judge those models. This means we can take some scientific claims as false, and others as objective, but without judging those scientific claims on whether they resemble anything in the external world. We can also take them as being scientifically legitimate without considering if they resemble anything in the external world. Rather, they are judged upon how well they conform to principles which stem from reason. However, someone who preferred the first approach of considering truth as relating to the human perspective might still have full sympathy with my argument. My argument will appeal to principles that confer a higher status upon science. I see these principles as stemming from reason but if those principles were simply seen as conveying truthfulness from a human perspective rather than stemming from reason then my argument still works. If it were accepted that truth from a human perspective should consider, for example, simpler theories as more truthful even though we cannot know if reality itself is simple then my arguments should be fully applicable.

Exactly what reason is and why it relates to objectivity is something that Kant and the neo-Kantians are far from explicit about. My understanding is that Kant sees the objectivity of reason as based in a transcendental argument. By transcendental I mean the necessary presuppositions that humans must make to have any knowledge at all. For humans to

have knowledge of the external world we must presuppose that the world takes a form which is amiable to human knowledge. We cannot prove this assumption because it goes beyond all possible experience. However, for humans to make knowledge claims we have to presuppose that it is true (Kant 2000, 5:186; [Patricia] Kitcher 1990, p. 19; Neiman 1994, p. 64). Additionally, since this presupposition posits that the world must be amiable to human knowledge, we can specify fundamental principles that are required by human knowledge by reflecting upon the production of human knowledge. By finding the elements that are required to have knowledge means we find the fundamental principles which make knowledge possible.

Such reflection shows that it is a necessary presupposition for experience to be possible it must take various forms, namely it must take the form of causation, laws and objects as formulated by the faculty of the understanding. However, as outlined above, this is insufficient to form knowledge because knowing that things must have shape does not specify if we are looking at an apple or a red ball. Neiman writes that "[t]he organisation of the material provided by the understanding requires principles of selection and choice, notions of relevance and appropriateness" (1994, p. 71). This very much fits the account of science that I will heavily emphasise whereby we need make choices over what factors are considered relevant and which are considered irrelevant when assessing data. As such, it is a necessary presupposition of knowledge that we posit some type of selection principles. These are the principles of reason which Kant employs. In this regard knowledge claims gain objectivity from reason because reason must presuppose various principles to make any knowledge claims at all.

3.5 Formulating Scientific Knowledge

I now outline the two steps which Kantians see as required to formulate scientific knowledge. These are, firstly, bringing data under concepts and, secondly, building a system of knowledge (known as systematising). In this section I will outline these as distinct steps. However, whilst all the

neo-Kantians I draw upon think both steps are required, not all of them saw a sharp demarcation between them.

On a typical Kantian picture, our observations of the world do not give us scientific knowledge. Our observations of the world are not raw, unfiltered perceptual data but instead are conceptually constituted. We cannot see the world independently of applying some concepts. Even if we could somehow access the raw perceptual data that we receive from the world it would just be an amalgamation of unstructured sensation without any form. We need to apply concepts such as space, time, causation and substance to that unstructured sensation. This produces what I shall call the observable world. What we apply to the world are "forms of knowing that mold private sensation into publicly available objects of possible experience" (Butts 1984, p. 7; see also Buchdahl 1969, p. 476; Cassirer 1957 [1929], p. 194; Meer 2018, p. 349). These concepts make possible an observable world whereby we might see two people talking or an apple fall from the tree.

This, however, is still not scientific knowledge. Such observations lack the precision of scientific knowledge. Rather, we need to bring observations under further concepts to formulate scientific knowledge. The observable world is too coarse grained and involves too many interacting factors to produce scientific knowledge. For example, neo-Kantian philosopher of science Michela Massimi shows how in Galileo's experiments

> the goal of the inclined plane experiment was to extract from the appearance (motion of a bronze ball along an inclined plane) the property of uniform acceleration…. we should not think that what we observe, say, a free-falling object, is the rough-and-ready observable phenomena… If we stick to the level of observable[s]… Galileo may seem no more right than Aristotle. (Massimi 2008, p. 25)

In this example, Galileo did not observe uniform acceleration for the simple reason that bronze balls do not have uniform acceleration when travelling down an inclined slope due to things like air resistance and friction. Consequently, Galileo conceptualised this notion of uniform acceleration by abstracting away air resistance and friction. Following this, Massimi sees "conceptually determined appearance or *phenomena* as the

proper object of scientific knowledge" (Massimi 2008, p. 14, emphasis original). The appearance of the ball rolling on the inclined plane is not the same as uniform acceleration, rather, scientists obtain the concept of uniform acceleration by bringing the appearance under concepts. Similarly, Cassirer writes that it is only through our input

> that the chaos of sensory impressions begin to clear and take on fixed form for us. The fluid impression assumes form and duration for us only when we *mould* it by symbolic action [what we apply] in one direction or another… the product of their activity [what we apply] in no way resembles the mere *material* with which they began. (1953b [1923], p. 107, emphasis original)

For Cassirer, we do not observe measurable properties in the world around us. Rather, measurable properties only arise once we have started to abstract away from what we observe. More bluntly, Rickert says that we "must destroy the immediacy of reality [by reality he means what we observe around us]" (1986 [1902], p. 54), Natorp says "disregarding all fluctuations of the senses" is required (2004 [1921], p. 161) and Vaihinger says objects arise by "cutting away definite portions of the given sensory material" (1935 [1911], p. 157). For these neo-Kantians, science is not based upon describing what we observe in the conceptually constituted world around us but instead transforming what we observe.

Most people believe that, when done properly, the results scientists come up with accurately portray the external world. Neo-Kantians reject this notion, believing our scientific views of the external world must involve some contribution from us. Massimi sees scientists as taking appearances and bringing them under concepts in a manner which creates scientific phenomena rather than describes how the world actually is independent of us. In this sense, "phenomena are something that… we *make*, rather than something that comes to us as ready-made in nature" (Massimi 2008, p. 8, emphasis original). There are important differences between the appearance of the ball and the property of uniform acceleration; they are not the same thing. Massimi primarily appeals to a modelling process when describing how appearances are turned into phenomena. Massimi believes there is some freedom in this process, writing that the

"phenomena we infer depends on the way we have carved and 'massaged' those data" (Massimi 2011, p. 104). Scientists employ numerous modelling techniques and theoretical principles when moving from appearances to phenomena. This is a key aspect of the neo-Kantian approach I adopt. Whilst typically using different terminology, the key neo-Kantian point is that science involves building models and those models are not the same as what we observe (Butts 1984, p. 187; Cassirer 1953a [1910], p. 130; Cassirer 1953b [1923], p. 108; Natorp 2015b [1887], p. 170; Natorp 2004 [1921], p. 210; Rickert 1986 [1902], p. 39; Windleband 2015b [1883], p. 272; Vaihinger 1935 [1911], p. 88).[9]

It is important to consider how we build scientific models on a Kantian approach. One aspect of this is the Kantian *synthetic a priori* (something that is independent of experience but which is applied to experience). Kant famously argued that the concepts we apply to sensation are *a priori*. Kant's notion of *a priori* seems to include three elements. Firstly, not involving empirical evidence from the external world. Secondly, fixed for all time whereby humans have always had one set of concepts and will always have them. Thirdly, universal whereby no human could lack those concepts. Kant is not clear upon where these *a priori* concepts come from but it seems like he thinks they stem from reason.[10] However, there is good reason to reject Kant's understanding of the *a priori*. Historical evidence suggests that Kant was being influenced by Newtonian science when formulating his *a priori* concepts (Friedman 2001, p. 14; Wilkerson 1976, p. 93). This suggests that those concepts were actually influenced by empirical evidence and that Kant was mistaken to see them as purely non-empirical. Additionally, Kant thought that those concepts entailed Newtonian physics. If those concepts were true for all time then so must

[9] Note that by observations I do not mean neutral observations of the external world. Rather, I mean conceptually constituted observations whereby what we see is already structured into space, time, entities and causes. Also note that the classic neo-Kantians typically talked of abstractions rather than models.

[10] Kant takes two quite different approaches to the *a priori* in *The Critique of Pure Reason* and a short summery of the book he published two years later. That short summary, which was partly published in response to critics of *The Critique of Pure Reason*, is named the *Prolegemena to Any Future Metaphysics*. In *The Critique of Pure Reason* Kant takes the synthetic approach whereby he takes the concepts to stem from reason whereas in the *Prolegemena* he takes the analytical approach whereby he searches for the concepts which would give justification to our best science without then trying to justify in turn those concepts.

Newtonian physics be true for all time. However, this stance is untenable given the development of Einstein's theories of relativity. Einstein not only provided an alternative to Newtonianism, based upon an alternative account of space and time to that which Kant thought of as *a priori*, but Einstein also offered a superior alternative to Newtonianism. Einstein used different concepts and generated empirical results which covered all empirical results derived from Newtonianism and additional empirical results which Newtonianism could not account for. Einstein's relativity arguably shows that Kant's approach to the *a priori* is untenable. It showed that the concepts Kant took as *a priori* could be challenged by empirical evidence and it showed that they were not fixed for all time or universal since Einstein offered alternative concepts.

Many philosophers took the development of Einstein's physics as a decisive disproof of Kantianism. However, whilst Kant's particular approach to *a priori* concepts seems untenable, some neo-Kantians have developed alternative accounts of the *a priori* which aim to retain some central Kantian insights. Some neo-Kantians developed these alternative accounts of the *a priori* in the aftermath of Einstein's physics (Cassirer 1953a [1921], p. 433; Natorp 2015a [1912], p. 183; Reichenbach 1965 [1920], p. 88) and some recent philosophers have further developed them (Friedman 2001, p. 45). The notion that *a priori* concepts are true for all time and universal is abandoned. However, the notion that we need apply *a priori* principles to sensation to generate knowledge is retained. On Reichenbach's redefinition of *a priori*, "'A priori' means 'before knowledge,' but not 'for all time' and not 'independent of experience'" (1965 [1920], p. 105; see also Cassirer 1953a [1910], p. 140; Rickert 1986 [1902], p. 225).[11] The key idea is that we employ principles that are independent of experience, in that they are required for experience to be possible, but we can find a conflict between our concepts and experience. Whereas Kant thought the world necessarily must conform to Newtonian principles because we constitute it as Newtonian, later neo-Kantians recognised there might be aspects of the world which resist being constituted in a Newtonian manner. An example of this is the orbit

[11] Reichenbach is typically associated with the logical positivists rather than the neo-Kantians, however, Reichenbach's first book was neo-Kantian and only later did he move to logical positivism.

of Mercury. As such, experience gives us reason to search for alternative *a priori* principles. Einstein then produced an alternative set of *a priori* principles which were independent of empirical data and made a new set of knowledge possible. Neo-Kantians thus see *a priori* principles as prior to knowledge, even if experience sometimes gives us good reason to seek out new *a priori* principles. The key point for my argument is that we need *a priori* principles to constitute the world and we have a level of choice over which *a priori* principles to employ (Cassirer 1953a [1921], p. 365; Reichenbach 1965 [1920], p. 87). As I will show later, however, I will depart from some neo-Kantian approaches to choosing which *a priori* to employ.

The second aspect of formulating scientific knowledge is systematicity. This is where individual knowledge claims are placed into a system by establishing how different knowledge claims relate to one another. This works in multiple ways. Firstly, given the modelling process directly described above, any particular results can be considered instances of the same phenomena or of a different phenomenon. For example, if scientists melt two pieces of lead and record different melting points because of differences in ambient air temperature then does this count as one phenomenon which produces two different results or as two different phenomena? A system of science would specify how those two singular results would relate to one another. Secondly, any two phenomena could be considered causally related where one phenomenon is the cause of another phenomenon. A system of science should specify which phenomena cause which other phenomena. Thirdly, entities can be posited which causally produce those phenomena. Entities are assigned properties which are taken as being causally responsible for particular phenomena. A system of science would specify that there are various different types of objects, specify which properties each type of entity has and specify how the properties of entities interact to produce other entities or phenomena.

This systematicity process is Kantian because it is taken as being the product of our minds rather than reflecting how things are in the mind-independent unstructured world. As neo-Kantian philosopher of science Philip Kitcher writes "[t]he causal structure of the world, the division of things into kinds, the objective dependencies among phenomena are all generated from our efforts at organisation" (Kitcher 1993, p. 172). This

is similar to my earlier discussion of Kant whereby we impose causation and objects onto the sensation we receive from the external world (Kant 1998, A23/B38 & A92/B1240). Doing this produced what Kant calls (on Meer's translation (2018, p. 348)) thought objects, whereby it is composed of thought rather than sensation. Similarly, Cassirer talks of objects as "intellectual constructions" (1953a [1921], p. 361), Natorp says that we know of objects through "intellectual construction" (2004 [1921], p. 373) and Vaihinger talks of objects as "ideational construct[s]" (1935 [1911], p. 157). Relatedly, Massimi describes natural kinds as "functionally relevant clusters of properties" (2014, p. 428) whereby entities are formulated to play a functional role in explaining a range of phenomena. There are multiple reasons to think that a system of science depends on us rather than it reflecting some mind-independent world. Firstly, as described earlier, the mind-independent world is inaccessible to us because all knowledge needs to involve our sensory capacities. This means all knowledge involves something which is not present in mind-independent reality but comes from us. Secondly, our need to formulate causes and entities does not determine which causes and entities there are. Rather, we need make choices when building a system of science. This again suggests that such a system depends on us rather than reflecting something mind-independent.

Kantians have taken different approaches to constructing objects. Kant outlines various principles for building a system of science. These principles have a different nature to the concepts which Kant believes we must apply to perception. Kant believed that we must apply concepts like space, time, substance and causation to the world. He then believed that we can use principles to order the product of that process into a system of science, specifying which objects and causes there are. Kant outlined three principles of reason for building a systematised science. These are, firstly, "sameness of kind… the principles of homogeneity" (A658/B686). This principle reduces differences by formulating general classes which do not demarcate between specific differences of instances of that general class. Secondly, there is "variety of what is same in kind… [the principle of] specification" (A658/B686). This principle highlights differences in generalised classes by formulating more specific classes which do include those differences. Thirdly, there is the "affinity of all concepts, which offer

a continuous transition from every species to every other through a graduated increase of varieties… [the principle of] continuity of forms" (A658/B686). This principle orders together the more generalised classes produced by the principle of homogeneity and the more specific classes produced by the principle of specification into a system, whereby the more specific classes are more detailed versions of the more general classes. The key point for now is that there are principles we can employ and those principles come from us rather than the world. These principles of reason determine the type of objects which we can know (Kant 2000, 5:181; Abela 2006, p. 421; Brittan 1991, p. 178; Butts 1984, p. 7; Kraus 2018, p. 78; Ostaric 2009, p. 165; Pilot 1990, p. 127; Morrison 1989, p. 160; Wartenberg 1992, p. 417; Zuckert 2017, p. 102).

Other Kantians have outlined alternative, if somewhat related, principles for formulating a system of science. They often talk of constructing objects as an act of unification (Cassirer 1953a [1910], p. 307; Kitcher 1989, p. 499; Natorp 2004 [1921], p. 386; Rickert 1986 [1902], p. 110). Cassirer describes entities as "fixed thing-unities" and a "cogitated point of unity" (1957 [1929], p. 142 & p. 317, see also pp. 287–288). He sees entities arising from transforming our observations of the world, abstracting away from what we observe to ideal entities which do not resemble what we observe. Rather, we take things that initially appear dissimilar and abstract away the dissimilarities to unify them into an ideal object. For example, in relation to chemistry, we establish various reactions between various substances. We then unify those results by positing various entities with various properties as causally producing those results. We posit elements and compounds with varying number of atoms, protons and neutrons. Doing this takes two seemingly dissimilar results, whereby coloured substances result from interactions, and unify those results by saying that they are produced by entities with various properties. In doing so we see "connections, common patterns, in what initially appeared to be different situations" (Kitcher 1989, p. 432). Which entities there are and which properties they have is not based upon reflecting mind-independent entities. Rather, it is based upon applying principles that stem from reason (Cassirer 1953a [1921], p. 365; Kitcher 1986, p. 229; Natorp 2004 [1921], p. 348; Reichenbach 1965 [1920], p. 87; Rickert 1986 [1902], p. 142; Windleband 2015b [1883], p. 272;

Vaihinger 1935 [1911], p. 157). For example, Cassirer thinks we should build a system of science based upon the principles of generality and simplicity, whereby we cover as much conceptually constituted data as possible (aiming to both incorporate all the data under theories and reduce the gap between theory and data) whilst simultaneously positing as few entities as possible. The most unified science would find a balance between those two principles of generality and simplicity, whereby we intellectually construct a system of scientific entities which best covers the data whilst being sensitive to the demands of simplicity. The key issue is that, as with Kant, these principles come from us. This is another key element of the neo-Kantian position I advocate for. The principles we employ in building scientific models come from us and do not reflect the mind-independent world.

I am not entirely sure why the principles of reason need take the form of homogeneity, specificity and affinity for Kant (see Neiman 1994, p. 77, for discussion). I find a defence of specific principles to be clearer in Cassirer. On my interpretation of Cassirer, to talk of knowledge we need to posit a target. That target is a world of intersubjective objects which is consistent and which different humans can communicate about. This would be the final complete science. Such intersubjective objects cannot be based upon reflecting what the world is like independent of our concepts. As such, we need an alternative standard to establish what those intersubjective concepts should be. For Cassirer, the correct intersubjective objects are those which are part of a maximally unified system. The reason why maximally unified matters is because this has to be the form the final complete science would take. Science completes itself by accounting for as much phenomena as possible whilst unifying that phenomena within a system which posits a hierarchical structure of explanations whereby one explanation explains other explanations, and those other explanations in turn explain other explanations, leaving as little as possible unexplained. As such, a unified view of the world entails regulative principles like generality, covering as much phenomena as possible, and simplicity, finding a smaller number of explanations which explain other explanations. Regulative principles like simplicity and generality are necessary presuppositions for us to talk of a maximally unified system of objects and therefore to meaningfully talk of intersubjective objects at

all (Cassirer 1953a [1910], p. 308; Heis 2014, p. 18). Knowledge demands intersubjective objects, intersubjective objects demand unity, and unity demands regulative principles. As such, Cassirer sees reason as entailing the regulative principles of simplicity and generality from which can generate different constitutive *a priori*.

I now combine together the above discussion to outline the specific neo-Kantian position which I endorse. Firstly, our observations of the external world are conceptually constituted. We do not see the world as it is but only as we interpret it through our faculties. Secondly, scientists take conceptually constituted observations (such as experimental data) and apply *a priori* constitutive principles. We choose which *a priori* constitutive principles to employ which then constitute scientific models in particular ways. Thirdly, the way we build those conceptually constituted models and the way we assess them depends upon how we employ principles which come from us and which do not produce a reflection of the external world. These points mean that we can produce a system of science in many different ways and we should take none of those ways as reflecting the external world. Despite this, we still have a means of judging good science from bad science even though it does not involve comparing science to the external world. Rather, we judge how well scientific theories, which have been constructed from experimental data, fit the concepts and principles which stem from reason. Given the status of reason as a source of objectivity, we can produce a science which we should take as objective and legitimate.

I now consider how my account differs from other philosophies of science. I demarcate it from scientific realism, anti-realism, instrumentalism and constructivism.

3.6 How My Position Differs from Scientific Realism

I now demarcate between a neo-Kantian position and notions of scientific realism found in philosophy of science. Note that some neo-Kantians see their position as realist but this is based upon consciously redefining

realism whereas scientific realism has much in common with traditional notions of realism.

A scientific realist position posits that the external world is potentially knowable and that some scientific theories have succeeded in correctly describing it. In philosophy of science, scientific realism is theory specific whereby we are scientific realists towards one theory and not scientific realists towards another. Thus, someone might be a scientific realist towards dinosaurs but not electrons (Stanford 2011, p. 893). Many scientific realists only see the most well confirmed and theoretically robust theories as meriting an ascription of realism. This typically means theories in physics which make extremely precise predictions, whose predictions have been proved accurate in many different experiments and which have undergone a significant period of theoretical development (Psillos 1999, p. 105). Consequently, scientific realists are typically highly selective over what parts of science they consider to be realist.

Scientific realism typically involves claims about theories and entities. Theories typically posit the existence of entities with various properties. For example, theories relating to electrons make various predictions which are well supported by experimental data. Scientific realists believe the reason why those theories are so predictively successful is because the entities posited by the theory actually exist. That being, there really are entities out there in the world which have the properties posited by theories about electrons. Put simply, the entities described by a theory which is realist actually exist (Psillos 2009a, p. 44).[12]

Scientific realists typically focus upon the status of unobservables. Philosophers of science sometimes demarcate between the observable world and the unobservable world. The observable world is the things that we can see in the world around us, such as trees, chairs and cats. Positing that observable entities exist is relatively uncontroversial for a scientific realist. The scientific realist thinks that we can see them. In contrast, positing the existence of unobservable entities, which we cannot see, is more controversial (Van Fraassen 2001, p. 162). The unobservable world is things which we cannot see. For example, electrons are too small

[12] There are two scientific realist notions which only partly subscribe to these claims. Structural realism posits that the theories accurately describe the world but the entities posited by the theory do not exist (Ladyman 1998). Entity realism is the opposite whereby theories do not accurately describe the world but the entities posited by the theory do exist (Cartwright 1999).

to see, tectonic plates are too far under the earth to see and societies are too big to see. We need make an inference which we could get wrong. However, the scientific realist claims that this inference is justified if our theories are sufficiently good. Given that theories about electrons make such good predictions we should believe that electrons exist even though we cannot see electrons (Psillos 2011, p. 24).[13]

Scientific realists typically think of theories corresponding with the world. Correspondence is where our ideas correspond with things in the real world. Our theories about electrons posit that electrons have various properties. If there actually exists an entity in the world which has those properties then our theories about electrons correspond to the world. In contrast, if no entity exists which has those properties then our theories about electrons do not correspond. However, scientific realists typically only commit to approximate truth rather than full truth (Psillos 1999, p. 276). They think that the best scientific theories are very close to accurately describing real entities but might not be fully correct in literally every detail. Understood in terms of correspondence, they might think that our theories about electrons accurately cover almost all properties of real electrons. This is compatible with our theories positing a property which real electrons actually lack, or real electrons having a property which our theories of electrons miss out. Similarly, our theories might posit a property which real electrons do have but the description of that property is slightly inaccurate in some very specific details. By committing to only approximate truth rather than full truth, a scientific realist thinks our best theories are basically correct whilst accepting some of the fine details might need tuning.[14]

[13] Philosophers of science have debated how to demarcate between the observable and the unobservable, and some have doubted if such a distinction can be drawn. For example, prior to the telescope the moons of Jupiter were unobservable whereas now they seem observable given that we have landed satellites, and could land people, upon them. Also, sometimes observables are seen as being theory-neutral whereas the unobservable requires theoretical presuppositions. We typically do not draw upon a range of theories when we notice a nearby tree. However, other philosophers of science think that theories are required to justify claims about observables, though they typically think those theories are good theories so we should believe in observables (Psillos 2009b, p. 88).

[14] Some scientific realists only think properties posited by theories which are well confirmed across multiple iterations of the theory and are central to the predictions, the so-called working posits, are the ones we should think of as corresponding (Psillos 1999). Also, some scientific realists reject correspondence theories of truth whereby they might adopt coherence instead. Such coherentist scientific realists are much closer to the neo-Kantian position I advocate for.

A neo-Kantian would reject or significantly clarify the claim that the world is knowable.[15] Neo-Kantians believe all knowledge depends upon applying something that is not present in the world. This means that any knowledge we have does not depend upon it reflecting the external world because we need to apply something that is not in the external world. To a neo-Kantian, we cannot see the world as it is, rather, we can only see it as it appears to individuals like us. Humans cannot have a God's eye view of the world. This means that our scientific theories do not correspond to the external world or only correspond in a non-typical way. Neo-Kantians recognise that our scientific theories posit entities with various properties. However, for two reasons, neo-Kantians believe those properties do not reflect the real properties of entities. Firstly, properties posited by theories involve something that comes from us, so those properties cannot reflect the real properties of objects in the mind-independent world. Secondly, our minds play a role in grouping together the properties we assign to entities so this does not reflect the real entities because those real entities have not been grouped together by our minds. Scientific realists believe that theories which make sufficiently impressive, well confirmed predictions do correspond, whereas the neo-Kantian thinks that even in very predictively successful theories our minds have added something on which makes reflection impossible. There is significant debate over what theory of truth Kant employed. Whilst it is not the main interpretation, Kant has sometimes been interpreted as holding a correspondence theory of truth, though usually one which has been significantly modified compared to typical correspondence theories (Allais 2003, p. 376). However, key neo-Kantian thinkers I draw from reject correspondence (Cassirer 1953a [1910], p. 36; Natorp 2004 [1921], p. 139; Rickert 1986 [1902], p. 45; Vaihinger 1935 [1911], p. 8). Additionally, neo-Kantians hold an alternative standard of truth and objectivity than scientific theories

[15] A Kantian might be a metaphysical anti-realist. Metaphysical realism posits that there exists a mind-independent world which exists independently of our ideas. A metaphysical anti-realist might claim that such a mind-independent world does not exist or, more moderately, deny that such a mind-independent world could ever even in principle be knowable (or, alternatively, never in principle be knowable to us but could potentially be knowable to God).

reflecting external objects in the world. Neo-Kantians typically relate truth and objectivity to our act of conceptually constituting the world and then systematising by placing that conceptually constituted world into a system. As discussed earlier, the place of the theory within the system, rather than the theory reflecting external objects, is how we should judge the truth of scientific theories.

3.7 How My Position Differs from Anti-Realism, Pragmatism and Operationalism

An alternative to scientific realism is anti-realism. An anti-realist accepts that there is a mind-independent world and that it is in principle knowable. Our scientific theories could potentially describe the external world whereby our scientific theories correspond to the external world. However, an anti-realist argues we have reason to doubt that our scientific theories do actually correspond (or, alternatively, we have no good reason to think they do correspond). Anti-realists can point to the history of science whereby there have been scientific theories which make very good predictions but posited entities which later turned out not to exist (Laudan 1981, p. 33). Consequently, our current best scientific theories might also posit entities which one day turn out not to exist despite making such good predictions. Anti-realists can also point to the possibility that scientists might think only one theory explains the phenomena but that there could be multiple theories, including ones not yet formulated, that explain the phenomena equally well (Stanford 2006, p. 171). These possibilities mean anti-realists think we should not believe that our best theories correspond to the world.

A neo-Kantian position disagrees with the anti-realist position by denying that scientific theories could accurately describe the external world. Whatever is going on in the external world is in some sense inaccessible to us because scientific knowledge always involves applying something to the world which is not in the external world. Whereas both the scientific realist and anti-realist think science can in principle

accurately describe the world and disagree over whether they think our best theories actually succeed in describing the world, a neo-Kantian will typically deny or significantly qualify the claim that scientific theories can accurately describe the world. Additionally, whereas the anti-realist does not think we should see our best scientific theories as being true, only believing they make excellent predictions, neo-Kantians judge theories as true based upon how well they contribute to a systematised view of the world.

I will now contrast my position with the pragmatist position I outlined in Chap. 2. A pragmatist position can also be considered a human centred philosophy and it accepts notions of human centred truths. As such, there is some similarity between pragmatism and neo-Kantianism. However, in Chap. 2 I outlined three potential problems with a pragmatist position. Here I will show how a neo-Kantian position provides answers to those problems. Firstly, I mentioned that it is unclear how much pragmatism differs from a non-essentialist realist position since both pragmatism and non-essentialist realists accept that we need make a contribution to forming a view of the world. In contrast, a neo-Kantian position is different from a non-essentialist realist by insisting that all knowledge of the external world depends upon something which we apply to the world. It is not the case that we need to add something which comes from us in some places but not in others; rather, all sensation we receive from the external world is brought under something which comes from us. Secondly, pragmatists believed that the world answers back to us but it is unclear in what way the world answers back. On a neo-Kantian position the world answers back through the sensation we receive from the external world. We have no control and do not influence which sensations we receive from the external world, even if we must bring that sensation under something which comes from us, and thus alter it, before we can ever access it. Thirdly, pragmatists believe we need to make a contribution when formulating scientific knowledge but it is unclear why we should value that scientific knowledge if it is based upon us rather than the world. I outlined above two different ways in which we should understand truth on a neo-Kantian position. As such, different neo-Kantians will hold different reasons for considering science as generating knowledge.

Another position which denies that scientific theories describe the external world is operationalism (see Chang 2021). It sees science as providing instructions or procedure for interacting with the observable world. A scientific theory is literally just taken to be a statement about what is expected to happen in the observable world under various conditions. For example, a theory about electrons is taken only to be a description of what is likely to be observed when passing electricity through a cloud chamber. However, to an operationalist, the scientific theory is nothing more than a set of instructions or procedures. Statements are only considered meaningful if they relate to the observable world. Claims about electrons are only meaningful in the sense that they help us make predictions in the observable world. Consequently, claims that there exists a real entity with the properties described by theories about electrons are taken as meaningless. Scientific theories should not be understood as aiming to give a truthful description of the unobservable world, rather, the meaning of scientific terms is nothing more than the way in which it helps us to make predictions.[16] In contrast, a neo-Kantian is happy talking about unobservables and thinking that unobservables are meaningful parts of scientific theories. They just think that unobservables, just like the observable world, are conceptually constituted and are systematised. Whether dealing with observables or unobservables, science judges the truth of scientific theories based upon how well they contribute to a system of science. To a neo-Kantian, unlike an operationalist, there is no demarcation between observables as meaningful and unobservables as meaningless.

3.8 How My Position Differs from Conceptual Constructivism and Social Constructivism

The general idea of conceptual constructivism is that we need construct a view of the world (for discussion see Hacking 1999; Kukla 2000). This rejects the notion that we can have a neutral interpretation of the world.

[16] Scientific claims are understood to be not truth-apt. They are not the sort of thing which we should think of as true or false. This position is known as semantic anti-realism.

We need to take a standpoint or a viewpoint when we aim to interpret the world. We can interpret the world in many different ways and we have no means of establishing which is the correct way to interpret the world. Since we cannot see the world as it is we need to use our minds to construct an image of the world. The image we construct may certainly have some basis and connection to the world but ultimately we always need to add something on when viewing the world. A related position is social constructivism whereby this construction process is taken as being socially embedded. Sometimes, social constructivists seem to suggest that there is no reality beyond what we construct. Reality is itself constructed, so reality itself changes as our social conventions and beliefs change. On this very strong reading there is no mind-independent reality above and beyond what we socially construct. A more moderate social constructivist position is critical realism. Critical realism aims to combine realism and social constructivism. There is a mind-independent reality but our view of it is socially constructed (Pilgrim and Bentall 1999, p. 262). Critical realism is ontologically realist but epistemologically relativist. Our description of the mind-independent world is influenced by our underlying assumptions, our culture and our values. The image of the world which we construct is at least partly dependent upon the historical era we are in, whereby it is dependent upon or influenced by conscious or unconscious values and assumptions held in that era. As such, rather than producing an image of the world which actually reflects the world, our image of the world reflects our values and assumptions. This might also have problematic ethical implications if our values are based upon prejudices.

A neo-Kantian position also accepts that our image of the world does not reflect the actual world. Additionally, a neo-Kantian can accept that scientists will often be influenced by conscious or unconscious values that are historically embedded whereby the values will change between different historical eras and this then influences the view of the world scientists produce. However, the key neo-Kantian notion is that at least some part of what we apply to the world has an elevated status. We can construct an image of the world in different ways but some ways of constructing an image of the world have a higher status providing they can be based in something which has that higher status. If we can understand our

scientific concepts as taking our observations of the world and transforming them into something based in reason then we have avoided relativism. Scientists may be influenced by many different underlying concepts but when we assess science we should consider those scientific claims which can be based in reason as in some sense truthful, whereas other approaches should not be considered truthful.

As a work in analytical philosophy of science, this text might at times seem disconnected from social constructivist concerns. In relation to psychiatry, a critical realist approach emphasises that psychiatric research can be driven by harmful prejudices and social stereotypes. Also, it emphasises that psychiatry is often reductionistic whereby it only focuses upon the biological and not the social (Pilgrim 2013, p. 344). I am more interested in establishing what a good psychiatric diagnoses would consist of rather than discuss the well-known problem of prejudices in psychiatric research and care. However, my approach actually is relevant for mitigating these issues. Firstly, all entities and causes that we can know about are conceptually constituted by us. This means biological and social entities, and biological and social causes, are conceptually constituted. This stance consequently rejects notions of biology as being more real or more scientifically legitimate than the social. Secondly, I emphasise that good scientific standards are those that have some basis in reason. Our best science does not entail sexism or racism. Positing one gender or race as being inferior to another would not be the best theory when measured by simplicity and generality (it may be simple but is not general given how much data it would not account for). As such, we have good reason to think they are not scientifically legitimate. Thirdly, some of my arguments for employing categorical psychiatric diagnoses are that they can help us track the causal impact of environments, help us track how people who exhibit psychiatric symptoms often end up in different (often harmful) environments and help us track the effects of self-conception on the diagnosed person. In some situations categorical psychiatric diagnoses are better able to do these compared to dimensional diagnoses.

Neo-Kantianism also takes a stance upon the notion that major changes in thinking might occur as cultures change and as different historical eras occur. A potential concern is that knowledge is relativised to culture and historical era. This would then undermine objectivity. The

neo-Kantian response is that there are standards which are independent of culture and historical era, namely those derived from reason. The Baden and Southwestern schools recognise that at different periods throughout history we might have an incomplete understanding of reason, and we might more successfully formulate and apply principles stemming from reason in some eras than others (see Cassirer 1957 [1929]). Despite this, reason is independent of history and culture so it can act as a standard of objectivity. Commentators upon psychiatry are tempted to deny that there is one ultimate reason (e.g. Rose and Kalathil 2019, p. 5). However, I feel this leads to relativism and that a neo-Kantian position provides an important non-relativist counter to such claims. Additionally, the Marburg and Southwestern schools developed a philosophy of culture whereby they aimed to find the rational elements within culture (Cassirer 1961 [1942]; Windleband 2015a [1910]). As such, they saw some types of culture as being more rational and therefore more objective than others. Applying this notion to psychiatry would require a project in its own right and perhaps one day I will give its feasibility consideration. I am not going to discuss this outside of a dedicated study because, on one hand, it might be a useful way of assessing which parts of our culture have a foundation in justice and equality but, on the other hand, rating cultures as better and worse clearly has some extremely bad historical precedences. The only thing to note is that those interested in what a neo-Kantian should think about the demarcation between health and illness, or normality and disorder, might consult their writing on philosophy of culture.

3.9 Problems with Kantianism and Neo-Kantianism

I consider the position I outline to be Kantian even though it does not reflect Kant's position. As discussed, there is disagreement over what Kant's position was and many neo-Kantians who followed him were consciously aware that their position significantly deviated from the position that they interpreted Kant as holding. Similarly, I think I need modify

not only Kant's position but also some aspects of the neo-Kantian position I have described.

Kant believed we could not see the world as it is because we had to add something onto the world which was not present in the world. This, however, did not compromise objectivity because the thing we added on was objective. Specifically, Kant thought (at least on common readings) that what we applied was (1) universal among humans, (2) it was fixed for all time, (3) it created necessary relationships, (4) it was based on reason, and (5) it was prior to all experience. The neo-Kantians of the 1870s to 1920s felt some of these claims were problematic so dropped some of them. Recognising that Newtonian physics is not true for all time, they dropped the claims about (1) universal among humans and (2) being fixed for all time. They still, however, believed that science would (3) form a system with necessary relationships, (4) that was based on reason, and (5) contained elements that were prior to experience. In this sense, they dropped two of the criteria Kant employed when seeking objectivity but saw the remaining criteria as sufficient to attain objectivity.

I will be dropping a further criterion, specifically, that what we apply (3) creates necessary relationships. Whereas the neo-Kantians emphasised a science governed by necessary relationships, I believe this is difficult to hold in relation to most sciences and even harder to hold in relation to psychiatry. The neo-Kantians can be considered to have been partly whiggish, whereby they believed science was progressing in a necessary manner (Friedman 2000, p. 89). Specifically, they thought that the claims made by earlier scientists could be embedded within claims made by later scientists. For example, when describing the history of chemistry, Cassirer talks of "the great logical lines of direction of their advance" (1953a [1910], p. 204). Cassirer seems to think that there was a path that chemistry either had to take or at least should have taken. He says that all the different approaches taken to chemistry all eventually fitted together. He writes that the "totality of all these different points of view, mutually confirming and correcting each other, that finally after many experiments gives a unitary table of atomic weights, and thus lays the basis of a definite system of chemical formulae" (1953a [1910], p. 207). The same is true in relation to physics. On their account, Newtonian physics entailed various predictions and Einsteinian physics entailed all those same

predictions and additional predictions. In this sense Einsteinian physics was a further logical development of Newtonian physics which entailed new predictions. There is a very real sense in which the neo-Kantians thought Einsteinian physics was the next step on the path after Newtonian physics, not taking seriously the possibility or likelihood that something significantly different to Einsteinian physics might have followed Newton. This allows them to see Newtonian physics and Einsteinian physics as having a necessary relationship, whereby one entails the other because they are both on the same logical path. This then allows the move from necessity to objectivity.

This view is not popular with modern philosophers of science. The notion of science as a logical and linear development was heavily challenged by sociologist and philosopher of science Thomas Kuhn (1996).[17] He used the history of science to argue that science does not progress in a linear manner, that successive theories do not simply neatly follow from one another and that science could have taken a variety of different paths. Put simply, science did not have to necessarily take the path that it did. Significantly inspired by Kuhn, philosophers of science started paying more attention to the history and sociology of science, aiming to understand how science actually develops rather than aiming to produce a rational reconstruction of science. Modern philosophers of science often emphasise the disunity of science whereby all scientific knowledge does not neatly fit together (Dupré 1993). Physicists do their own thing, chemists do their own thing and biologists do their own thing. Additionally, a physicist working upon quantum gravity might have little interest in understanding how their work relates to the works of scientists studying relativity. Most scientists are typically not interested in getting the whole of science to fit together and are unconcerned if it does not fit together. For example, a study suggested that there are thirty different models of how nuclear physics works and they all contradict one another (Morrison 2011). Providing each of the models produces useful predictions then most scientists would not be concerned by this. Most science does not involve finding logical relationships between theories whereby one theory entails another theory. Additionally, if we took this as a goal

[17] Ironically, Kuhn was inspired by the 1870s to 1920s neo-Kantians (Friedman 2008, p. 243).

for science it is unclear that science is currently meeting that goal very well. Put simply, if the objectivity of science is based upon there being necessary relationships between different parts of science then it looks like science is not objective. We might be able to judge which theories are better than other theories by appealing to a static set of criteria for theory change like simplicity and generality but we should not expect that scientific theories should follow one another in a manner where they have increasing degrees of simplicity and generality. If this is problematic for sciences like physics then it is even more problematic for psychiatry. It seems completely implausible to say currently employed psychiatric diagnoses are necessarily related to earlier psychiatric diagnoses. It seems completely mistaken to claim that earlier diagnoses can be logically derived from current diagnoses or believe that future diagnoses will logically entail current diagnoses.

Whilst I abandon necessity as a criterion of objectivity, I still retain the notion that (4) science involves building a system that is based on reason, and (5) science contains elements that are prior to experience. In this regard I hope to situate psychiatric diagnoses within a neo-Kantian framework which provides greater levels of objectivity compared to that entailed by other philosophies of science. If, ultimately, readers are unconvinced that this is sufficiently Kantian to be considered Kantian then I hope I will still give a novel reading of psychiatric diagnoses which will contain novel and nuanced philosophical insights into what psychiatric diagnoses are and what they should be.

Another potential problem is that I heavily draw upon Cassirer but is questioned whether Cassirer actually is a neo-Kantian. There is a possibility that Cassirer is Hegelian. On a typical Kantian account, reason is fixed for all time. We might discover more about reason across time but there just is one fixed rationality. This was challenged by Hegel. Rather than seeing reason as static, Hegel posited that reason unfolded across time as history developed. Contradictions in thought would be dialectically resolved as history developed, ultimately resulting in reason itself developing. Reason is something that develops rather than being fixed for all time for Hegel. Cassirer can be seen as Hegelians because he also posits that human history goes through stages in his three volume work, the philosophy of the symbolic forms, which was published in the 1920s. We

first understand the world through myth. The myths we posit do not reflect what occurs in the external world but they still are quite close to basic sensation. As history develops we started employing language in a manner that is more abstract than myth, moving further away from basic sensation. Eventually, we developed science which is even more abstract than language whereby the scientific entities we posit are very distant from sensation. Whilst the notion that scientific concepts do not resemble sensation is very neo-Kantian, the notion that this has been reached through a series of historical developments may be Hegelian.

A Kantian interpretation of Cassirer can certainly accept that at different historical eras we have a different understanding of reason. The deeper problem is that reason itself changes in different historical eras. Cassirer believes that philosophy should find the *a priori* principles which make knowledge possible in any particular era. These *a priori* principles change as we move from myth to language and then to science. This can be understood as reason itself changing, making reason itself dynamic. This would make Cassirer more Hegelian rather than Kantian. Whether this interpretation is correct is debatable. Some scholars think that Cassirer is more Hegelian than Kantian (Verne 2004, p. 21), others think Cassirer is both Hegelians and Kantian (Luft 2015b, p. 223), others think that it is unclear whether Hegelian or Kantian strands are stronger in Cassirer's philosophy (Richardson 2010, p. 288) and other philosophers think that Cassirer is Kantian and not Hegelian (Ferrari 2010, p. 298). There is also disagreement about whether the symbolic forms are transhistorical (Friedman 2004, p. 241) or not transhistorical (Rudolph 2008, p. 9). I now justify why my use of Cassirer fits my Kantian aims.

Firstly, the main work by Cassirer I draw upon is his work *Substance and Function* which was published in 1910. This has been described as "the standard work of Neo-Kantian epistemology" (Ferrari 2015, p. 266). This was published before his three volumes on the *Philosophy of the Symbolic Forms.* Heis suggests work on the symbolic forms "differs in some fundamental ways" (Heis 2018) from his contemporary neo-Kantian thinkers. If we assume his philosophy of the symbolic forms was Hegelian this does not appear to the case with his earlier work.

Secondly, Cassirer posits three different symbolic forms, namely myth, language and science. My aim in this book is to argue that psychiatry can be profitably understood through neo-Kantian notions of science. If

reason is dynamic to Cassirer then I am aiming to understand psychiatry through the science stage of reason. Conveniently, science may be considered the most developed symbolic form. In some publications Cassirer seems to emphasise that none of the symbolic forms have priority over the others whereas in other publications he seems explicit that myth is the lowest stage and science is the highest stage (see Freudenthal 2008, p. 207; Luft 2015b, pp. 234–235). As such, I am potentially analysing psychiatry through the highest symbolic form.

Thirdly, whilst the *a priori* employed changes throughout different historical eras, there might still be something underlying the *a priori* which is itself static. This can be understood as regulative principles (Dupré 2004, p. 42; Friedman 2001, p. 65; Heis 2014, p. 14; Ryckman 2015, p. 75). Whereas Kant thought the *a priori* we apply was distinct from regulative principles, Cassirer can be understood to think that which *a priori* we apply depends upon applying regulative principles to sensation. For example, the sensation produced by Newton during his experiments combined with static regulative principles produced the Newtonian *a priori* whereas the sensation produced by different experiments in the early 1900s combined with static regulative principles produced the Einsteinian *a priori*. Put another way, Cassirer thinks there are dynamic *a priori* principles, such as those that condition Newtonianism and Einsteinianism, and there are static *a priori* principles which govern theory selection (Heis 2014, p. 17). Cassirer talks of Einsteinian science being superior to Newtonian because it is more complete and simpler (1953a [1921], p. 365). Cassirer recognised that constitutive *a priori* principles change over time but to retain the notion of one final complete science Cassirer needs to posit some static *a priori* principle which guides selecting the changing constitutive *a priori* principles. These are the principles of simplicity and generality which are required by a maximally unified science.

3.10 Conclusion

In this chapter I have outlined a neo-Kantian approach to formulating scientific knowledge. Neo-Kantians believe that things in themselves are inaccessible to us and that we can only know of how things appear to us. Additionally, how things appear to us, what we perceive in the external world, is not itself scientific knowledge due to the lack of precision and because it is not a system of science. To formulate scientific knowledge we need to bring data under concepts and build a system of science. We have various principles we can use when employing concepts and building a system of science. These principles provide a standard of legitimate, objective and truthful science and these standards do not rely upon resembling the mind-independent world.

I will employ a neo-Kantian position to develop a new understanding of psychiatric diagnoses. I will show how psychiatric diagnoses have often been assessed by considering whether they resemble anything in the mind-independent external world. In contrast, by drawing upon a neo-Kantian position, I argue that this is an untenable standard for assessing psychiatric diagnoses. Instead, I will draw upon the standards that neo-Kantians use when assessing scientific theories and then employ these standards to assess psychiatric diagnoses. Doing this will show that, under certain conditions, psychiatric diagnoses are much more scientifically legitimate compared to what many critics claim. Also, I shall suggest that some (though certainly not all) currently employed psychiatric diagnoses either meet those standards or are close to doing so. Additionally, some neo-Kantians believe that those standards for assessing scientific theories can deliver truth and objectivity, and I will also suggest that, when they meet those standards, it can make sense to talk of psychiatric diagnoses as being objective.

References

Abela, Paul. 2006. The Demands of Systematicity: Rational Judgment and the Structure of Nature. In *A Companion to Kant*, ed. Graham Bird, 408–422. Blackwell.

Allais, Lucy. 2003. Kant's Transcendental Idealism and Contemporary Anti-realism. *International Journal of Philosophical Studies 11* (4): 369–392. https://doi.org/10.1080/0967255032000136489.

Allison, Henry E. 1983. *Kant's Transcendental Idealism: An Interpretation and Defense*. New Haven, CT: Yale University Press.

Ameriks, Karl. 1992. The Critique of Metaphysics: Kant and Traditional Ontology. In *The Cambridge Companion to Kant*, ed. Paul Guyer, 249–279. Cambridge: Cambridge University Press.

Banham, Gary. 2013. Regulative Principles and Regulative Ideas. In *Kant und die Philosophie in weltbürgerlicher Absicht*, ed. Stefan Bacin, Alfredo Ferrarin, Claudio La Rocca, and Margit Ruffing, 15–24. Boston: De Gruyter.

Beiser, Frederick C. 2008. Emil Lask and Kantianism. *The Philosophical Forum* 39 (2): 283–295.

———. 2009. Normativity in Neo-Kantianism: Its Rise and Fall. *International Journal of Philosophical Studies 17* (1): 9–27. https://doi.org/10.1080/09672550802610941.

Brittan, Gordon G. 1991. Systematicity and Objectivity in the Third Critique. *Southern Journal of Philosophy 30* (S1): 167–186. https://doi.org/10.1111/j.2041-6962.1992.tb00664.x.

Buchdahl, Gerd. 1967. The Relation Between 'Understanding' and 'Reason' in the Architectonic of Kant's Philosophy. *Proceedings of the Aristotelian Society 67*: 209–226. https://doi.org/10.1093/aristotelian/67.1.209.

———. 1969. *Metaphysics and the Philosophy of Science; the Classical Origins, Descartes to Kant*. Oxford: Basil Blackwell.

Butts, Robert E. 1984. *Kant and the Double Government Methodology: Supersensibility and Method in Kant's Philosophy of Science*. Dordrecht: D. Reidel Publishing Company.

———. 1990. Teleology and Scientific Method in Kant's Critique of Judgment. *Noûs 24* (1): 1–16. https://doi.org/10.1007/978-94-015-8188-2_6.

Cartwright, Nancy. 1999. *The Dappled World*. Cambridge: Cambridge University Press.

Cassirer, Ernst. 1953a. *Substance and Function and Einstein's Theory of Relativity*. Dover Publications.

———. 1953b. *The Philosophy of Symbolic Forms Volume One: Language*. Yale University Press.

———. 1957. *The Philosophy of Symbolic Forms Volume Three: The Phenomenology of Knowledge*. Yale University Press.

———. 1961. *The Logic of the Humanities*. Yale University Press.

Cassirer, Heinz. 1978. *Kant's First Critique: An Appraisal of the Permanent Significance of Kant's Critique of Pure Reason*. New York: Humanities Press.

Centi, Beatrice. 2015. The Validity of Norms in Neo-Kantian Ethics. In *New Approaches to Neo-Kantianism*, ed. Nicholas de Warren and Andrea Staiti, 127–146. Cambridge: Cambridge University Press.

Chang, Hasok. 2021. Operationalism. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/fall2021/entries/operationalism/. Accessed April 10 2024.

Crowell, Steven G. 2010. Transcendental Logic and Minimal Empiricism: Lask and McDowell on the Unboundedness of the Conceptual. In *New Approaches to Neo-Kantianism*, ed. Nicholas de Warren and Andrea Staiti, 113–131. Cambridge: Cambridge University Press.

Dudley, Will, and Kristina Engelhard. 2014. Introduction. In *Immanuel Kant Key Concepts*, ed. Will Dudley and Kristina Engelhard, 1–12. Oxon: Routledge.

Dupré, John. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.

Dupré, Louis. 2004. Cassirer's Symbolic Theory of Culture and the Historicization of Philosophy. In *Symbolic Forms and Cultural Studies*, ed. Cyrus Hamlin and John Michael Krois, 35–48. New Haven: Yale University Press.

Ferrari, Massimo. 2010. Is Cassirer a Neo-Kantian Methodologically Speaking? In Neo-Kantianism in Contemporary Philosophy. In *Neo-Kantianism in Contemporary Philosophy*, ed. Rudolf A. Makkreel and Sebastian Luft, 293–314. Bloomington: Indiana University Press.

———. 2015. Cassirer and the Philosophy of Science. In *New Approaches to Neo-Kantianism*, ed. Nicholas de Warren and Andrea Staiti, 261–284. Cambridge: Cambridge University Press.

Forster, Michael. 2008. *Kant and Skepticism*. Princeton, NJ: Princeton University Press.

Freudenthal, Gideon. 2008. The Hero of Enlightenment. In *The Symbolic Construction of Reality*, ed. Jeffrey Andrew Barash, 189–214. Chicago, IL: The University of Chicago Press.

Friedman, Michael. 2000. *A Parting of the Ways: Carnap, Cassirer, and Heidegger*. Open Court.

———. 2001. *The Dynamics of Reason*. Stanford, CA: CSLI Publications.

———. 2004. The Davos Disputation and Twentieth-Century Philosophy. In *Symbolic Forms and Cultural Studies*, ed. Cyrus Hamlin and John Michael Krois, 227–243. New Haven, CT: Yale University Press.

———. 2008. Ernst Cassirer and Thomas Kuhn: The Neo-Kantian Tradition in History and Philosophy of Science. *Philosophical. Forum* 39 (2): 239–252. https://doi.org/10.1111/j.1467-9191.2008.00293.x.

Gardner, Sebastian. 1999. *Routledge Philosophy GuideBook to Kant and the Critique of Pure Reason*. London: Routledge.

Glock, Hans-Johann. 2015. *Neo-Kantianism and Analytical Philosophy*. In *New Approaches to Neo-Kantianism*, ed. Nicholas de Warren and Andrea Staiti, 59–81. Cambridge: Cambridge University Press.

Grier, Michelle. 1997. Kant on the Illusion of a Systematic Unity of Knowledge. *History of Philosophy Quarterly 14* (1): 1–28.

Guyer, Paul. 1980. Kant on the Apperception and *A Priori* Synthesis. *American Philosophical Quarterly 17* (3): 205–212.

———. 2006a. *Kant (Routledge Philosophers)*. London: Routledge.

———. 2006b. Bridging the Gulf: Kant's Project in the Third Critique. In *A Companion to Kant*, ed. Graham Bird, 423–440. Blackwell.

Hacking, Ian. 1999. *The Social Construction of What*. Cambridge, MA: Harvard University Press.

Heis, Jeremey. 2014. Realism, Functions, and the a priori: Ernst Cassirer's Philosophy of Science. *Studies in the History and Philosophy of Science* 48: 10–19.

Heis, Jeremy. 2018. Neo-Kantianism. In *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), ed. Edward N. Zalta. https://plato.stanford.edu/archives/sum2018/entries/neo-kantianism/.

———. 1998. *Critique of Pure Reason (The Cambridge Edition of the Works of Immanuel Kant)*. Cambridge: Cambridge University Press.

———. 2000. *Critique of the Power of Judgment (The Cambridge Edition of the Works of Immanuel Kant)*. New York: Cambridge University Press.

Kant, Immanuel. 2002. *Theoretical Philosophy After 1781 (Kant, Immanuel, 1724–1804. Works. English. 1992)*. Cambridge: Cambridge University Press.

Kinzel, Katherina. 2021. Historical Thought in German Neo-Kantianism. *British Journal for the History of Philosophy* 29 (4): 579–589. https://doi.org/10.1080/09608788.2021.1932411.

Kitcher, Philip. 1986. Projecting the Order of Nature. In *Kant's Philosophy of the Physical Sciences*, ed. Robert E. Butts, 201–238. Dordrecht: D. Reidel Publishing Company.

———. 1989. Explanatory Unification and the Causal Structure of the World. In *Scientific Explanations*, ed. Philiphp Kitcher and Wesley C. Salmon, 410–506. Minnesota: Minneapolis.

Kitcher, Patricia. 1990. *Kant's Transcendental Psychology*. Oxford: Oxford University Press.

Kitcher, Philip. 1993. *The Advancement of Science*. New York: Oxford University Press.

———. 2007. A Priori. In *The Cambridge Companion to Kant and modern philosophy*, ed. Paul Guyer, 28–60. Cambridge: Cambridge University Press.

Köhnke, Klaus. 1991. *The Rise of Neo-Kantianism: German Academic Philosophy Between Idealism and Positivism*. Cambridge: Cambridge University Press.

Kraus, Katharina T. 2018. The Soul as the 'guiding idea' of Psychology: Kant on Scientific Psychology, Systematicity, and the Idea of the Soul. *Studies in History and Philosophy of Science Part A 71*: 77–88. https://doi.org/10.1016/j.shpsa.2017.11.010.

Krijnen, Chrisian. 2015. Philosophy as Philosophy of Culture? In *New Approaches to Neo-Kantianism*, ed. Nicholas de Warren and Andrea Staiti, 111–126. Cambridge: Cambridge University Press.

Kuehn, Manfred. 2001. *Kant: A Biography*. Cambridge: Cambridge University Press.

Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions*. 3rd ed. Chicago: University of Chicago Press.

Kühn, Manfred. 2010. Interpreting Kant Correctly: On the Kant of the Neo-Kantians. In *Neo-Kantianism in Contemporary Philosophy*, ed. Rudolf A. Makkreel and Sebastian Luft, 113–131. Bloomington: Indiana University Press.

Kukla, Andre. 2000. *Social Constructivism and the Philosophy of Science*. London: Routledge.

Ladyman, James. 1998. What Is Structural Realism? *Studies in History and Philosophy of Science Part A 29* (3): 409–424. https://doi.org/10.1016/S0039-3681(98)80129-5.

Langton, R. 1998. *Kantian Humility: Our Ignorance of Things in Themselves*. Oxford: Oxford University Press.

Laudan, Larry. 1981. A Confutation of Convergent Realism. *Philosophy of Science* 48 (1): 19–49. https://doi.org/10.1086/288975.

Luft, Sebastian. 2010. Reconstruction and Reduction: Natorp and Husserl on Method and the Question of Subjectivity. In *Neo-Kantianism in Contemporary Philosophy*, ed. Rudolf A. Makkreel and Sebastian Luft, 59–91. Bloomington: Indiana University Press.

———. 2015a. Editors Introduction. In *The Neo-Kantian Reader*, ed. Sebastian Luft, xx–xxxi. London: Routledge.

———. 2015b. The Philosophy of the Marburg School: From the Critique of Scientific Cognition to the Philosophy of Culture. In *New Approaches to Neo-Kantianism*, ed. Nicholas de Warren and Andrea Staiti, 221–239. Cambridge: Cambridge University Press.

Makkreel, Rudolf A., and Sebastian Luft. 2010. Introduction. In *Neo-Kantianism in Contemporary Philosophy*, ed. Rudolf A. Makkreel and Sebastian Luft, 1–24. Bloomington: Indiana University Press.

Massimi, Michela. 2008. Why There Are No Ready-Made Phenomena: What Philosophers of Science Should Learn from Kant. In *Kant and Philosophy of Science Today*, ed. Michela Massimi, 1–35. Cambridge: Cambridge University Press.

———. 2011. From Data to Phenomena: A Kantian Stance. *Synthese* 182: 101–116. https://doi.org/10.1007/s11229-009-9611-z.

———. 2014. Natural Kinds and Naturalised Kantianism. *Nous*. 48 (3): 416–449.

Meer, Rudolf. 2018. Immanuel Kant's Theory of Objects and Its Inherent link to Natural Science. *Open Philosophy 1*: 342–359. https://doi.org/10.1515/opphil-2018-0025.

Meerbote, Ralf. 1991. Systematicity and Realism in Kant's Transcendental Idealism. *The Southern Journal of Philosophy 30* (2): 129–137. https://doi.org/10.1111/j.2041-6962.1992.tb00661.x.

Morrison, Margaret. 1989. Methodological Rules in Kant's Philosophy of Science. *Kant-Studien 80* (2): 155–172. https://doi.org/10.1515/kant.1989.80.1-4.155.

———. 2011. One Phenomenon, Many Models: Inconsistency and Complementarity. *Studies in History and Philosophy of Science Part A 42* (2): 342–351. https://doi.org/10.1016/j.shpsa.2010.11.042.

Natorp, Paul. 2004. *Plato's Theory of Ideas, An Introduction to Idealism*. Academia Verlag Sankt Augustin.

———. 2015a. Kant and the Marburg School. In *The Neo-Kantian Reader*, ed. Sebastian Luft, 180–197. London: Routledge.

———. 2015b. On the Objective and Subjective Grounding of Knowledge. In *The Neo-Kantian Reader*, ed. Sebastian Luft, 164–179. London: Routledge.

Neiman, Susan. 1994. *The Unity of Reason, Rereading Kant*. New York: Oxford.

O'Neill, Onora. 1992. Vindicating Reason. In *The Cambridge Companion to Kant*, ed. P. Guyer, 280–308. Cambridge: Cambridge University Press.

Okruhlik, Kathleen. 1986. Kant on Realism and Methodology. In *Kant's Philosophy of Physical Science: Metaphysische Anfangsgründe der Naturwissenschaft 1786–1986*, ed. Robert E. Butts, 307–329. Dordrecht: Springer Netherlands.

Ostaric, Lara. 2009. Kant's Account of Nature's Systematicity and the Unity of Theoretical and Practical Reason. *Inquiry 52* (2): 155–178. https://doi.org/10.1080/00201740902790235.

Pilgrim, David. 2013. The Failure of Diagnostic Psychiatry and Some Prospects of Scientific Progress Offered by Critical Realism. *Journal of Critical Realism* 12 (3): 336–358. https://doi.org/10.1179/1476743013Z.0000000004.

Pilgrim, David, and Richard Bentall. 1999. The Medicalisation of Misery: A Critical Realist Analysis of the Concept of Depression. *Journal of Mental Health* 8 (3): 261–274. https://doi.org/10.1080/09638239917427.

Pilot, Harold. 1990. Kant's Theory of the Autonomy of Reflective Judgment as an Ethics of Experiential Thinking. *Noûs* 24 (1): 111–135. https://doi.org/10.2307/2215616.

Psillos, Stathis. 1999. *Scientific Realism: How Science Tracks Truth*. Oxon: Routledge.

———. 2009a. Scientific Realism and Metaphysics. In *Knowing the Structure of Nature: Essays on Realism and Explanation*, ed. Stathis Psillos, 34–47. Basingstoke: Palgrave Macmillan.

———. 2009b. Tracking the Real: Through Thick and Thin. In *Knowing the Structure of Nature: Essays on Realism and Explanation*, ed. Stathis Psillos, 84–98. Basingstoke: Palgrave Macmillan.

———. 2011. The Scope and Limits of the No Miracles Argument. In *Explanation, Prediction, and Confirmation Volume II*, ed. Maria Carla Galavotti and Friedrich Stadler, 23–35. Dordrecht: Springer.

Reichenbach, H. 1965. *The Theory of Relativity and Apriori Knowledge*. Berkeley: University of California Press.

Richardson, Alan. 2010. Ernst Cassirer and Michael Friedman: Kantian Or Hegelian Dynamics of Reason. In *Discourse on a New Method, Reinvigorating the Marriage of History and Philosophy of Science*, ed. Mary Domski and Michael Dickson, 279–294. Chicago: Open Court.

Rickert, Heinrich. 1986. *The Limits of Concept Formation in Natural Science: A Logical Introduction to the Historical Sciences*. Cambridge: Cambridge University Press.

Rose, Diana and Jayasree Kalathil. 2019. Power, Priviledge and Knowledge: the Untenable Promise of Coproduction in Mental "Health". *Frontiers in Sociology*. 57 (4): 1–11.

Rubenstein, Eric M. 2001. Rethinking Kant on Individuation. *Kantian Review 5*: 73–89. https://doi.org/10.1017/S1369415400000650.

Rudolph, Enno. 2008. Symbol and History, Ernst Cassirer's Critique of the Philosophy of History. In *The Symbolic Construction of Reality*, ed. Jeffrey Andrew Barash, 3–16. Chicago: The University of Chicago Press.

Ryckman, Thomas. 2015. A Retrospective View of Determinism and Indeterminism in Modern Physics. In *The Philosophy of Ernst Cassirer*, ed. J. Tyler Friedman and Sebastian Luft, 65–102. De Gruyter: Berlin.

Schlick, Moritz. 2015. Critical Or Empiricist Interpretation of Modern Physics. In *The Neo-Kantian Reader*, ed. Sebastian Luft, 447–456. London: Routledge.

Stanford, Kyle. 2006. *Exceeding Our Grasp: Science, History and the Problem of Unconvinced Alternatives*. Oxford: Oxford University Press.

Stanford, Kyle P. 2011. *Damn the Consequences: Projective Evidence and the Heterogeneity of Scientific Confirmation.* 78 (5): 887–988. https://doi.org/10.1086/662283.

Vaihinger, Hans. 1935. *The Philosophy of 'as if,': A System of the Theoretical, Practical and Religious Fictions of Mankind*. 2nd ed. London: Routledge & K. Paul.

Van Fraassen, Bas. 2001. Constructive Empiricism Now. *Philosophical Studies* 106: 151–170.

Verne, Donald Phillip. 2004. Cassirer's Concept of a Philosophy of Human Culture. In *Symbolic Forms and Cultural Studies*, ed. Cyrus Hamlin and John Michael Krois, 19–27. New Haven: Yale University Press.

Wartenberg, Thomas. 1992. Reason and the Practice of Science. In *The Cambridge Companion to Kant*, ed. Paul Guyer, 228–248. Cambridge: Cambridge University Press.

Wilkerson, Terrance Edward. 1976. *Kant's Critique of Pure Reason: A Commentary for Students*. Oxford: Oxford University Press.

Windleband, William. 2015a. Philosophy of Culture and Transcendental Idealism. In *The Neo-Kantian Reader*, ed. Sebastian Luft, 371–324. London: Routledge.

———. 2015b. Critical Or Genetic Method? In *The Neo-Kantian Reader*, ed. Sebastian Luft, 271–286. London: Routledge.

Zuckert, Rachel. 2017. Empirical Scientific Investigation and the Ideas of Reason. In *Kant and the Laws of Nature*, ed. Michela Massimi and Angela Breitenbach, 89–107. Cambridge: Cambridge University Press.

4

Neo-Kantian Understanding of Symptoms

4.1 Introduction

It is generally assumed that symptoms are real aspects of the mind-independent world. People exist, they behave in certain ways and some of those ways constitute a symptom. Some philosophers and psychiatrists believe that which behaviour constitutes a symptom is determined by underlying causes, whilst others believe that we need to consider our ethical values to determine which behaviour is merely normal behaviour and which is a pathological symptom. In this second case the behaviour is real even if psychiatrists need to then decide which behaviour constitutes a symptom. Thus symptoms are in some sense mind-independent.[1] This is not to suggest that it is particularly easy to accurately detect and describe these mind-independent symptoms. Psychiatrists plausibly often inaccurately describe symptoms. Rather, the claim is that the correct symptoms

[1] It could be objected that psychiatry must be mind-dependent because psychiatry studies the mind and the brain. I use the terms mind-independent and mind-dependent in the philosophical senses of the term. Mind-independent means what things are like independent of our views and mind-dependent means something depends upon our views.

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_4

are parts of the mind-independent world and that psychiatrists could potentially accurately describe those symptoms.

A neo-Kantian would not see symptoms as mind-independent. Neo-Kantians believe that the objects of scientific knowledge are in some sense constructed and do not reflect the mind-independent world. Neo-Kantians see the objects of scientific knowledge as the product of a transformation process. What we see in the observable world is transformed by bringing observations under concepts and placing them into a system of science. Thus, a neo-Kantian would consider symptoms as not mind-independent. Rather, symptoms are constructed.

The notion that symptoms are mind-independent is typically taken to play an important epistemic role. On this notion the mind-independent world determines the correct symptoms. If, however, as neo-Kantians argue, the objects of scientific knowledge are constructed, then alternative criteria will be required to determine what counts as a good symptom. A neo-Kantian will see the correct symptoms as those which have been constructed in a particular way, taking what we observe and idealising it in accordance with epistemological standards.

In this chapter I firstly outline how psychiatrists and philosophers of psychiatry see symptoms as mind-independent. Then I outline neo-Kantian notions of phenomena and show that this is how symptoms should be understood. This will show that symptoms are constructed. I then discuss the standards neo-Kantians can employ to give an account of how we should formulate symptoms.

4.2 Symptoms Are Typically Seen as Mind-Independent

Most people would typically consider symptoms to be part of the mind-independent world. For example, many autistic people are considered to exhibit the symptom low eye contact. It intuitively seems like it is a fact that either eye contact is given or it is not given (or, to be more nuanced, it is given to a certain degree). Eyes seem to exist and where they look and for how long also seem to be a fact. So too, intuitively, many other

symptoms would be considered mind-independent. Whether someone sees imaginary beings, and thus hallucinates, or whether someone is feeling an all crushing oppression, and thus is depressed, seems to be a fact. Such real parts of the world are something psychiatrists wish to find out about.

Symptoms as mind-independent can be understood as characteristics. It is widely accepted that there are different things in the world and those different things will have various characteristics. Horses have one set of characteristics whilst octopuses have a different set of characteristics. It seems such characteristics are part of the world. We can discover that horses have bones and live on land whilst octopuses do not have bones and live under water. Similarly, so too do different people have characteristics. Some are tall, some are short and some hallucinate or have literal thinking. That these people have these characteristics is a fact, it is not dependent upon the views of psychiatrists. It is typically considered a fact because there are real causes which produce those characteristics. Those real causes can be ones internal to the organism, external to the organism or a combination of the two. Such real causes determine, for example, the height of an individual and how they think. Similarly, such real causes would determine whether someone hallucinates or has literal thinking. Characteristics are real, mind-independent things because they have real causes. Thus symptoms are part of the pre-grouped, mind-independent world due to those pre-grouped causes. The world has already determined what symptoms there are, psychiatrists need only discover them. Of course, psychiatrists could be (and plausibly often are) mistaken in what symptoms they attribute to people but it is a fact that those characteristics are out there whether psychiatrists correctly describe them or not.

There are two potential challenges to these views which are worth discussing so that my claims here can be demarcated from them. Firstly, it could be argued that symptoms are not mind-independent because they involve value judgements. To have a symptom means someone has an illness. Illnesses are things that we consider to be bad. We have made a judgement that the characteristics forming the illness are a bad thing. We judge according to our values that hallucinations are bad things and thus we consider them instances of an illness. Since values are not part of the mind-independent world this means that symptoms are not

mind-independent parts of the world. When I talk of construction in psychiatry I do not have this concern in mind. I think there are two ways to respond to this problem. Firstly, it might be claimed that what constitutes an illness is actually determined by the world. For example, an illness might be considered parts of an organism which fail to function properly. The world determines the correct functioning of an organism and thus the world determines what symptoms there are (Boorse 1977, p. 555). Secondly, it could be argued that the characteristics that we consider to be symptoms are real parts of the world even if it is not determined by the world that they should count as symptoms. On this approach, there are mind-independent characteristics in the world. Having discovered these mind-independent characteristics we then decide using our ethical values which of those mind-independent characteristics are symptoms. Thus symptoms are mind-independent characteristics even if which mind-independent characteristics are symptoms needs to be determined by our values (Beebee and Sabbarton-Leary 2010, p. 25; Cooper 2005, p. 50; Murphy 2006, p. 24).

A second problem, which needs distinguishing from my concerns in this chapter, is that symptoms appear to have fuzzy boundaries. It can be difficult to tell if an individual has a particular symptom because they only partly manifest it. For example, it might seem intuitively obvious who is hallucinating and who is not but actually the hearing of voices is quite prevalent within the general population (Bentall 2014, p. 40; Rashed 2010, p. 189). Thus it is unclear whether those people have hallucinations. This problem can take three general forms. Firstly, some behaviour might be very mild compared to more concrete instances of the symptom such as hearing voices only rarely. Secondly, some behaviour may only be partially like paradigmatic instances of the symptoms such as the voices always being friendly. Thirdly, the behaviour might contain elements not associated with the paradigmatic instances of the symptoms such as religious content. If it is unclear what constitutes hallucinations then it can be argued that this needs to be determined by psychiatrists and thus symptoms are not mind-independent. These sorts of issues are only a small part of why I see psychiatry to involve construction. Also, I believe there are possible responses to these problems which would still allow symptoms to be mind-independent. Firstly, both the

concrete instances of a symptom and the ambiguous cases are mind-independent characteristics even if there is ambiguity over whether they constitute the same mind-independent characteristic or not. Secondly, only the concrete instances of the symptom are a mind-independent symptom whereas the more ambiguous cases are not mind-independent. Rather, psychiatrists must decide whether they count as instances of the mind-independent cases and this thus renders those ambiguous cases not mind-independent. These responses mean that most or all symptoms can be mind-independent in spite of these problems. I shall draw upon other reasons to think that symptoms are mind-dependent.

The neo-Kantian approach I advocate for sees all symptoms as mind-dependent. It does so for reasons other than those two problems just discussed. Thus these responses to the issues of our values and the issues of ambiguous boundaries will not block the reasons why my neo-Kantian position takes symptoms as mind-dependent. In the following sections I will outline a neo-Kantian account of symptoms, showing they are not mind-independent whilst showing that there are better and worse ways of constructing symptoms.

4.3 Bringing Behaviour Under Concepts

Neo-Kantians do not see the objects of knowledge as mind-independent. Rather, the objects of knowledge are systematised appearances. This is where appearances are brought under concepts and those concepts are then fitted within a wider body of knowledge (Kant 1998, A658/B686). In this section I shall describe appearances and then show how they are transformed into symptoms.

Humans feel like they make observations of the world around them. To a neo-Kantian this misses out important details. As described in Chap. 3, we bring perceptions under concepts to construct objects which causally interact with one another (Kant 1998, A26/B43 & A34/B51 & A80/B106). We cannot access the basic perceptions we receive from the world. Rather, at the most basic level, we can access a world constituted as having objects which causally interact with one another. This is what we

typically call the observable world (even though that 'observable' world has already been conceptually constituted).

Within this observable world we observe other people and we observe them exhibiting behaviour. This typically occurs subconsciously though it can be influenced by conscious thought. Humans are also typically predisposed to notice certain behaviour as salient. Which behaviour is salient can depend upon a variety of factors, such as underlying psychology of the observer, mood of the observer, the social arrangement between the observable and the observed, the circumstances which the behaviour is exhibited in and nature of the behaviour in question. This may be driven by underlying structures of our mind whereby fixed or malleable biological or psychological structures force or impel us to focus on certain behaviours. Also, social and cultural factors likely play a role in making some behaviour salient, such as cultural expectations about how women should behave may lead to behaviour which defies those expectations being salient. Finally, an individual's past experience will also likely play a role in which behaviour is salient.

Humans are also typically predisposed to group behaviour together. Rather than just seeing singular instances of behaviour, which are unconnected to one another, they instead group behaviours together. For example, some behaviours are considered to be anger and other behaviours are considered to be sadness. The behaviours are grouped together as instances of a wider class. These groups will typically be quite implicit since this process is rarely driven by a conscious decision making process. Also, this grouping process is likely to be ambiguous and imprecise. This process is also likely driven by structures of our mind, social and cultural factors and our past experience. This grouping process produces what I call the observable world. Although the product of multiple contributions from us, humans typically think that they see people being angry or feeling sad in the world around us. Such generalised behaviour is typically considered part of the observable world.

The results of these two steps, of noticing behaviour as salient and grouping behaviour together produce an observable world of behaviour but it is not scientific knowledge. This observable world is too imprecise to be scientific knowledge. As I will show, scientific knowledge is a transformation of that observable world. Scientists do not simply use that

observable world but rather they construct an idealised version of it which results in groupings and demarcations which are distant from that observable world. I shall highlight this with a quick example before giving a detailed account and justification of this process. Consider the symptoms of low social skills which commonly occur among autistic individuals. The way in which we think of low social skills is quite different to our initial observation of behaviour in the observable world. We might observe someone being rude or we might observe someone who is hostile in social situations but these two behaviours are quite different to autistic low social skills. Autistic individuals can exhibit low social skills despite not being rude or hostile. As a second example, we might observe someone being confused whereas we might assign them the symptom of literal understanding, delusions or low intellect. In both these examples a transformation takes place between what we observe and what symptoms are employed. Our initial observations provide one account but by doing science a different account is produced. This is not simply a case of renaming behaviour, rather, I shall show how, firstly, the symptoms group together behaviour in ways that are distant from the groupings we make in observation and, secondly, the way the behaviour is understood is changed by bringing it under additional concepts. That being, whilst concepts were used to conceptually constitute an observable world, further concepts are used to turn that observable world into scientific knowledge. I will now proceed to outline this process in greater detail and show how neo-Kantians understand the justification for it.

To set up a neo-Kantian account of symptoms I first draw upon an important late twentieth-century development in philosophy of science. Traditionally, philosophy of science demarcates sharply between data and theory. Traditionally, data is seen as the product of experiments and theories explain that data. An influential 1988 paper by Bogen and Woodward argued this picture missed out an important intermediate step. Though not neo-Kantians, Bogen and Woodward's claims contain important concepts for a neo-Kantian position. Bogen and Woodward suggest the traditional picture of data and theories needs to be supplemented by phenomena, an additional step lying between them. The basis for this was the immense number of causal factors present when scientists perform experiments. They describe how

> the outcome of any given application of a thermometer to a lead sample depends not only on the melting point of lead, but also on its purity, on the workings of the thermometer, on the way it was applied and read, on interactions between the initial temperature of the thermometer and that of the sample, and a variety of other background conditions. (Bogen and Woodward 1988, p. 309)

Therefore, the reading of any specific phenomena will be the product of many different causes, ones which can change from experiment to experiment. Bogen writes that "[d]ata typically result from complex, loosely connected, short-lived assemblies of causal factors… The causes that produce data sets are never exactly the same from one trial to another" (Bogen 2010, p. 789). Specific experimental set-ups produce specific results due to different set-ups involving some different causes. Many causes are not relevant to the phenomena, since "in typical cases data are the result of many causal factors and at most [only] some of these will have to do with the phenomena of interest" (Woodward 2010, p. 167). When building theories scientists desire a more general figure than is provided by specific experiments, one not tied to specific causal factors present in specific instances of the experiment. Scientists are not interested in insignificant factors like variations in air temperature which have a minimal influence on results, perhaps only resulting in changes of a few decimal places. Such causal factors are not considered the cause which determines the melting point of lead, rather, they are incidental factors which interfere with finding the correct figure. The true cause is not taken to be such incidental factors, rather, the true cause is the molecular structure of lead. Consequently, Bogen and Woodward posited a notion of phenomena which is a more idealised, more abstracted, more generalised notion than is given by individual experiments. Thus whilst different experiments might provide results of 328°C, 327.1°C, 327.3°C, etc., scientists might agree upon phenomena as 327.5°C, a general figure abstracted away from specific experiments. Thus data is yielded by experiments produced by specific causal factors, but for scientific purposes we require a more generalisable notion.

Bogen and Woodward have not fully outlined exactly what phenomena are. Bogen describes them as "ontological furniture" (Bogen 2011,

p. 8) employed in science. Does this mean they are mind-independent parts of the world or imposed by scientists? Bogen and Woodward's classic paper takes a realist approach to phenomena (1988, p. 337). This, however, seems to discount the possibility that choices need to be made over which causes to abstract away. Woodward more recently argues phenomena are relative to purpose (Woodward 2010, p. 174). This would leave phenomena sitting between two clear positions of being real and being constructed. I believe that Bogen and Woodward need further philosophical work to argue phenomena fit both categories.

Massimi has produced a detailed neo-Kantian account which accommodates phenomena as both real and being constructed. Massimi sees "conceptually determined appearance or *phenomena* as the proper object of scientific knowledge" (Massimi 2008, p. 14, emphasis original). Scientists bring appearances under concepts. Parts of appearances are removed and abstracted away in this process, whereby scientists only focus upon parts of appearances, considering other parts as noise. In doing so, scientists give properties to appearances which they otherwise do not have. Massimi gives the example of Galileo's experiments, describing how

> the goal of the inclined plane experiment was to extract from the appearance (motion of a bronze ball along an inclined plane) the property of uniform acceleration…. we should not think that what we observe, say, a free-falling object, is the rough-and-ready observable phenomena… If we stick to the level of observable[s]… Galileo may seem no more right than Aristotle. (Massimi 2008, p. 25)

On this account what Galileo observed was brought under concepts, using a series of theoretical positions. Parts of the appearance are removed, such as the actual motion of the ball which moves at a non-uniform rate due to the varying surface of the wood. The results of this process transformed what was observed, giving the observations of the property of uniform acceleration which it otherwise lacks. There are important differences between the appearance of the ball and the property of uniform acceleration; they are not the same thing. Consequently, phenomena are something which we construct through our concepts. Most people would

consider the melting point of lead to be determined by the mind-independent world, that is, mind-independent causes determine that lead has a particular melting point. However, we actually have to decide which causes to idealise and abstract away and which to consider relevant when formulating the melting point of lead. If we abstracted away different causes then we might formulate a different figure for the melting point of lead. As Massimi writes, "phenomena we infer depends on the way we have carved and 'massaged' those data" (Massimi 2011, p. 104). This means that "phenomena are something that… we *make*, rather than something that comes to us as ready-made in nature" (Massimi 2008, p. 8, emphasis original).

Related ideas can be found in other neo-Kantian thinkers. Rickert writes that "every leaf on a tree appears different from the leaves next to it, or that no fragment of a chemical substance in a retort [a distillation device] is exactly like any other fragment of the 'same' substance" (1986/1902, p. 38). If studied closely enough there are specific details of particular leaves which are different to any other leaf, such as the exact shape or texture. Science removes such details when creating a more abstract and idealised image of the world (Rickert 1986/1902, p. 217).

Natorp outlines how the entities posited by science do not resemble what we observe around us. Natorp writes that "we should think of each thing [the entities posited by science], without confusion and in distinct separation, as the precise thing it is and not as something else… In sense perception, by contrast, one and the same thing may appear as one or as two, as larger or as smaller, etc., without perception by itself being capable of disentangling this confusion of contradictory determinations" (2004/1921, p. 202). In the observable world we might see that phenomena as being produced by two different objects or as varying in size in different manifestations, whereas scientists might posit a single entity with particular characteristics as responsible for a particular phenomenon.

Similarly, Cassirer writes that when doing science there is an "inexactness of reproduction, which never retains the whole of the earlier impression but merely its hazy outline, renders possible this unification of elements that are in themselves dissimilar" (1953/1910, p. 18). Cassirer means that science inexactly reproduces an image of the world, removing many parts of what we perceive through sensation, but by doing so it can

join together elements of the world which have similarities, such as considering two leaves as one type of objects despite both leaves having differences from one another. This means that, for Cassirer, science corrects what we observe, resulting in a "transformation from observation to science" (1953/1910, p. 143). On these accounts, phenomena cannot be mind-independent, rather, they are mind-dependent because we construct them by removing parts of what comes to us in appearances and give properties to the parts which remains (see also Butts 1984, p. 187; Windleband 2015/1883, p. 272; Vaihinger 1935/1911, p. 189).

All this shows how neo-Kantians see phenomena as mind-dependent. Following Bogen and Woodward's distinction between data and phenomena, scientists need to develop more general notions which are not tied to specific, localised factors. This is done through scientists bringing appearances under concepts. In doing so appearances are transformed into phenomena and thus phenomena do not reflect a mind-independent reality but rather are constructed.

In doing the transformation from observation to science new connections are made (Cassirer 1953/1910, p. 224; Natorp 2004/1921, p. 353; Rickert 1986/1902, p. 110; Vaihinger 1935/1911, p. 50). Building an idealised image of the world which does not resemble the observable world can help us make more accurate predictions. By removing factors that are deemed irrelevant we can build an idealised image of interactions that can help us make more precise predictions of the observable world even when those irrelevant factors actually are present. Similarly, by idealising away factors deemed irrelevant we can gain a better understanding of the interaction of a limited range of causes, which then helps us better understand the observable world even though it contains those irrelevant factors. For example, by ignoring air resistance when considering the effect of gravity, we can produce very general descriptions which can help make more accurate predictions once we remodel the effects of air resistance back in. Similarly, by removing all unwanted causal influences on the melting point of lead we can gain a better understanding of how molecules causally interact, which then helps understand causal interactions when we observe lead even when those unwanted causal influences actually are present.

4.4 Psychiatric Symptoms

Applying these notions to psychiatry, I see behaviour as data and symptoms as phenomena. By behaviour I mean our observations of other people. We make observations of others and this forms basic units of data. Such observations could occur in everyday settings (e.g. chance encounters in the street or regular encounters with family members), in care settings (e.g. a quick assessment by a psychiatrist or long-term care by a psychiatric nurse) and in experimental settings (a single experiment by a psychologist or a long-term follow-up study). We might also observe ourselves, such as having awareness of our feelings and perceptions. These observations of behaviour might be literally of bodily behaviour but will also include observations of the feelings, thoughts and perceptions of others (such observations of course actually involves inferences and concepts but neo-Kantians see all observations, including of bodily behaviour, as being conceptually constituted). By observation I mean enough content to obtain a basic unit of data. This might simply consist of a sufficient length of time for a behaviour to appear salient or might be whatever length of time is required for a test to produce a particular result. Such basic units of data are not the same as a symptom. Behaviour is too tied to localised situations, potentially dependent upon specific causal factors which are not present in many manifestations of a symptom. Symptoms need to be formulated with an increased degree of generality. They need to be abstracted away from specific instances to create a notion which is usable across many different notions. I thus see symptoms as phenomena, a generalised notion which is not tied to specific situations and can manifest in multiple ways. Phenomena are behaviour which have been brought under concepts and have undergone a modelling process. I now explore the ways in which this can occur, highlighting how symptoms are constructed and highlighting how there can be choices over how symptoms are created.

In bringing behaviour under concepts a transformation takes place. This is similar to how Galileo's notion of uniform acceleration involves transforming observations of the ball rolling down the inclined plane (since resistance means the ball does not actually accelerate uniformly).

The concept involves applying properties which are not present in what is observed. This is also the case with psychiatry since observations of someone being rude are transformed into the symptoms of low social skills. This works in multiple ways. Firstly, to say someone is being rude is to take them as either purposely exhibiting that behaviour or as making insufficient effort to avoid exhibiting that behaviour. In contrast, the low social skills of autism are taken as being something which is outside of the control of the individual, or at least significant effort would be needed to make even a limited difference. Secondly, the symptom of low social skills covers a variety of behaviour, such as being rude, but also low social skill links together much more behaviour than is covered by rudeness, such as missing social cues. Thirdly, rudeness has aspects which are not covered under notions of autistic low social skills. Purposefully provoking others might be considered to be rude but such behaviour is not covered by low social skills. All this shows that symptoms are not the same as what we observe. A transformation takes place when we bring observations under concepts.

In doing so we can make more connections in the observable world around us. For example, by taking some people who seem rude and instead considering them as having autistic social skills, we can notice more specific ways in which the behaviour is being manifested. We might notice that those who have autistic social skills have a much greater inclination to monologue compared to those who are considered rude, whereas those who are considered rude are more likely to make jokes that are considered obscene than those who exhibit autistic social skills. Similarly, new causal connections can be spotted. We might notice that rude behaviour does not change significantly even when in highly structured social situations where everyone is given a set length of time and topic for speaking, whereas autistic social skills do change in that highly structured social situation. As such, by transforming the observable world into science we can then also gain a better understanding of the observable world.

4.5 Epistemic Standards for Psychiatric Symptoms

I have argued that symptoms are not part of the mind-independent world. Rather, they are constructed by psychiatrists. This may raise epistemological concerns. If symptoms were part of the mind-independent world then we would have a standard to judge posited symptoms on. The true symptoms would be those which reflect the mind-independent world. However, if symptoms are constructed then alternative standards are needed than reflecting the mind-independent world when epistemologically assessing symptoms. Neo-Kantians have standards upon which to formulate phenomena. This means that there are constraints upon what constitutes a good symptom, since symptoms which have been formulated to respect these standards should be considered good science. Although these standards constrain what counts as a good symptom I will also show that within these standards there is still a level of choice over how to formulate symptoms.

Neo-Kantians assess scientific concepts by considering the place of a scientific concept within a system. They appeal to various principles when assessing this. We can consider the wider system and these principles when considering how to bring behaviour under concepts to formulate knowledge.

Neo-Kantians emphasise embedding scientific concepts within other scientific concepts. When we posit a symptom we need to consider its place within the wider system. Neo-Kantians typically understand this in terms of causation and explanation whereby one part of the system is considered to cause or explain another part of the system. To posit a notion of autistic social skills is to embed this symptom within a wider causal and explanatory network. We consequently might embed autistic low social skills within notions of autism and within notions of how people are brought up to follow social rules within particular societies. Our theorising about these should play a role in deciding what factors to disregard and which to incorporate when formulating autistic social skills. We take autism as a relevant cause or explanation of autistic social skills whereas, for example, the ambient air temperature is not taken as

causally or explanatory relevant even though this might influence the specific words said by an autistic individual in a particular situation. A more complicated example might be an individual who had a bad upbringing. Whatever behaviour they manifest some might not be considered an instance of the symptom of low social skills because the causal influence is so different from most causes of low social skills. Modern psychiatrists would consider low social skills arising from a bad upbringing to not be autistic low social skills because it does not stem from the causes of autism, such as theory of mind-deficits (see Evans 2013, p. 23; Evans 2017, p. 419 for analysis).

What constitutes social skills and social relations involves theoretical assumptions, be them explicit or theoretical. What exactly constitutes social understanding rather than emotional or cognitive understanding involves many theoretical claims about the functioning of the mind and what influences it. Additionally, theoretical standpoints are required to demarcate different types of social skills. For instance, one instance of low social skills might be abnormal social relationships, such as an autistic person wishing for a social relationship to take an abnormal form. Other instances might be abnormal social understanding, such as wishing for a typical social relationship but not knowing how to socialise in the manner which non-autistic people do. This subdivision requires various theoretical claims relating to desires and understanding.

Neo-Kantians also value simplicity through constructing scientific phenomena with generality. The number of phenomena posited should, where possible, be limited. Neo-Kantians seek general notions which remove much of what we see in the world around us. Consequently, we should posit a much smaller number of scientific concepts rather than simply (if this was somehow possible) replicating what we observe in the world around us. As such, we desire symptoms to have a level of generality. For example, social skills can manifest in many different ways. Low social skills can take place in many different locations (home, school, work), can be exhibited to many different people (complete strangers, lifelong family members, someone higher or lower in a social hierarchy), can take place for many different purposes (being friendly, requesting information, expressing views) and each person exhibiting low social skills will have many other specific parts of their personality. The

behavior exhibited by individuals in these cases will be far from identical. Consequently, symptoms need formulating with a significant degree of generality. They need to be abstracted and idealised away from specific instances to create a notion which is usable across many different possible manifestations.

When respecting the principle of simplicity, we still need to respect the data itself. Whilst it is perfectly legitimate and desirable that scientists do not fully convey all data when formulating scientific concepts, we still need the scientific concepts to be constrained by the data. As such, symptoms do not need to fully describe behaviour but they do need to be constrained by behaviour. Symptoms need not fully describe behaviour because the behaviour will be heavily abstracted and idealised when the symptom is formulated. For example, a symptom might be derived from drawing a graph curve through a range of scattered data points representing instances of behaviour, idealising away to a greater degree the data points which are further from the drawn curve. Despite this, the behaviour needs to still constrain the symptoms because scientists cannot draw the curve in any manner they feel like. It needs to be constrained by various sophisticated statistical, modelling and mathematical techniques.

An important consideration when scientific concepts are constrained by data is the frequency of any particular data point. Behaviour needs have a level of frequency to constitute an instance of a symptom. This is typically done through statistical analysis but could also be done through a general feeling obtained through multiple clinical observations. In doing this some behaviour may be relatively common and other behaviour may be relatively infrequent. Behaviour which occurs with insufficient frequency may not be considered part of the symptom. By being so rare the behaviour is not then formulated as part of a symptom, it might effectively be disregarded by psychiatry or only spoken about in case studies rather than be formulated into a symptom which is part of the DSM. Alternatively, it might be considered a manifestation of a symptom but one which occurs with insufficient frequency to be considered a core manifestation. For example, imagine a study asks if autistic individuals exhibit a particular manifestation of low social skills. Such manifestations are present in a study but with insufficient frequency to meet statistical relevancy, such as low social skills resulting from emotional

causes rather than cognitive causes. They would be considered localised manifestations of the symptom low social skills rather than core manifestations.

Another important neo-Kantian goal is coverage. A system of science would ideally cover everything that is considered scientifically relevant. For example, if scientists make new discoveries then ideally those new discoveries would be incorporated into the system of science. Scientists might then formulate a new phenomenon to cover these new discoveries. Alternatively, they might decide that these new discoveries are already covered by an existing phenomenon, perhaps being manifestations of that existing phenomenon in a new domain or more precise measurements of that existing phenomenon. Additionally, they might think currently formulated phenomena might cover those new discoveries after we expand the phenomena by giving them new properties. We can judge a system of science partly upon the degree that all relevant data are being incorporated into the system. This is then relevant for assessing particular phenomena within a system whereby we might need to make changes to the phenomena to increase coverage. If a new behaviour is discovered that is exhibited by a subset of autistic people but does not fit any currently formulated symptom then we might formulate a new symptom of autism, extend a current symptom or consider the behaviour to not be an instance of autism.

When formulating symptoms, we may need to find a balance between these different considerations. For example, to increase coverage or accuracy we might desire to formulate more symptoms but this might reduce simplicity and may make it harder to link all the symptoms of a diagnosis together using a smaller number of theories. Similarly, to increase coverage we might need to expand a symptom but this might make the symptom too general which then reduces accuracy to the data. Also, to increase the accuracy we might need to alter theories in ways that then make them not cohere so well with other theories. An important consideration when assessing symptoms is the diagnoses and theories themselves. If a particular behaviour does not readily fit the theories associated with a diagnosis then perhaps that is a problem with the theories or the diagnosis. Perhaps theories or diagnoses should be changed to increase accuracy or coverage.

As such, assessing symptoms also requires assessing diagnoses and causes, which is the topic of the next two chapters.

I will highlight alternative ways of formulating symptoms with two examples. The first one is changing notions of what constitutes low social skills in autism. During the era in which psychoanalysis held considerable influence there was much focus on the inner psychic life of the individual. Many symptoms were seen as being influenced by dynamic, inner emotional processes (see Evans 2017, Chap. 2; Hollin 2014, p. 104; Nadesan 2005, Chap. 5; Silverman 2012, Chap. 3). Following this, the low social skills of autism were seen to arise from emotional processes. In the 1980s autism was reformulated as a disorder of social understanding. Additionally, the notion of social understanding employed was cognitive, relating to the abnormal way in which autistic people perceive the world and process information (see Evans 2017, Chap. 5; Eyal et al. 2010, Chap. 9; Hollin 2014, p. 102; Verhoeff 2013, p. 451). Thus, emotional factors were no longer considered responsible for the low social skills of autistic people, rather, cognitive factors drive this. I now consider the role of emotions in low social skills in terms of data and phenomena. It is uncontroversial that someone being anxious or more generally feeling emotionally distressed can influence how they socialise. How should scientists accommodate this? I discuss three possible options (see Fellowes 2021 for more details on this example). As a second example, I also relate the issues I discuss to Bogen and Woodward's classic example of melting point of lead.

The first option is that the influence of emotions, or the influence of cognition, should be abstracted away. I shall highlight this by showing how emotions could be abstracted away. Many factors can influence whether low social skills are expressed and how they are expressed, such as who the autistic person is talking to and why. It makes sense that many causal factors should be abstracted away. Emotions would be one such factor. Emotions can influence whether someone exhibits low social skills and the specific words said but this does not mean emotions are part of the symptom low social skills, much like whether you are talking to your boss or your friend can influence how low social skills manifest but are not themselves part of what constitutes low social skills. Consequently, they are abstracted away. Anyone sufficiently emotionally distressed

might be excluded from a study. Alternatively, such an individual might be included within a data set but if the data set were sufficiently large and most instances were caused by cognition then the effects of emotion would become a statistical outlier, being levelled out by all those who are not affected by emotions. This is much like how the ambient air temperature is abstracted when determining the phenomena which lead melts at. An individual piece of lead may melt at 325.1 degrees or 324.8 degrees due to ambient air temperature but the melting point is actually 325 degrees.

The second option is that notions of low social skills could be expanded to incorporate emotions. Under this approach many different factors are still abstracted away, such as whether someone is talking to their boss or whether they are talking to their friends, but emotions are factors which are not abstracted away. When an individual exhibits low social skills it could be caused by emotions, it could be caused by cognition or it could be caused by both. Both count as low social skills. Also, low social skills can manifest differently depending on whether they are caused by cognition, caused by emotion or caused by both. All these count as instances of low social skills. Under this approach what behaviour constitutes low social skills may need to be expanded, considering additional behaviour as constituting low social skills than currently occurs. Some scholars believe notions of autistic low social skills should be expanded in this manner (Hobson 2002, p. 134; Hodge 2004, p. 53; Maiese 2013, p. 181). This would be like expanding the notion of what constitutes the melting point of lead to incorporate ambient air temperature. There would no longer be a notion of an abstract melting point of lead independent of the environment, rather, the melting point of lead must always include the ambient air temperature and thus the melting point of lead would constitute a larger range of figures than it currently does.

The third option is that two types of low social skills could be formulated, the first being a cogitative low social skills and the other an emotional low social skills. Sometimes low social skills are caused primarily by cognition and take a particular form due to cognition. Other times low social skills are caused by emotions and take a particular form due to emotions. Consequently, two types of low social skills can be formulated. It is possible that both cognition and emotions can simultaneously be

causal factors. However, if one factor were more dominant then the other factor would be incidental and would be abstracted away. This would be like formulating two different melting points of lead, one for high ambient air temperature and one for low ambient air temperature. There is no longer one melting point of lead, rather, there is now phenomena of lead melting at high ambient air temperature and the phenomena of lead melting at low ambient air temperature.

We can appeal to neo-Kantian standards when aiming to assess this. Whether the effects of emotions should be abstracted away could depend on accuracy to instances of behaviour. The symptom could accurately account for whether a particular instance of behaviour is cognitive or emotional in nature, be it entirely one of those or a combination of the two. The frequency that particular instances of behaviour are primarily cognitive or emotional is also relevant for symptom formation. If one is significantly more frequent than the other then the less frequent factor may be abstracted away. It can also depend on causal factors. If low social skills are primarily caused by cognition then there is reason to consider a causal contribution from emotion as incidental whereas if sufficiently frequent then emotions might not be considered an incidental factor. It may also depend upon wider theories which can be used to understand how and under what circumstances emotions and cognition could play a significant role in processing social information. If theory suggested that social understanding was primarily reliant upon the cognitive or emotional parts of our mind then there would be reason to abstract away the other. If theory suggests both simultaneously play a role then there would be reasons to formulate one notion of low social skills which covered both whereas if theory suggests both could play a role but operated independently of one another then there would be reason to formulate two different types of low social skills. Thus, which factors are considered causally relevant and which should be abstracted away can be influenced by frequency of data and theories. Finally, symptoms can be formulated with varying degrees of generality depending upon the desires of psychiatrists. They could be formulated just as low social skills, as cognitive low social skills and emotional low social skills or formulated into multiple different types of cognitive low social skills and multiple different types of emotional low social skills.

I have highlighted how neo-Kantianism provides standards for constructing phenomena. Those standards do not rest upon mind-independent phenomena or notions that the mind-independent world causes mind-independent phenomena. Though neo-Kantians provide standards for formulating phenomena there can still be significant levels of choice in phenomena formulation. This mainly relates to how scientists go about modelling and idealising, facing significant choices over which factors to include and which to ignore.

4.6 Conclusion

Symptoms are typically seen as mind-independent parts of the world. Just as animals have objective characteristics, so too humans have objective characteristics. So too psychiatric symptoms are typically seen as mind-independent parts of the world. People exhibit symptoms and these are the objective characteristics which psychiatrists aim to determine. However, a neo-Kantian would reject notions of mind-independent symptoms. Rather, neo-Kantians see symptoms as being constructed. They are systematised appearances. I have described how appearances are brought under concepts and are thus transformed by abstracting away some parts whilst giving other parts new properties. Appearances are transformed and therefore symptoms are not mind-independent. I have outlined principles that neo-Kantians would employ to construct scientifically legitimate symptoms.

References

Beebee, Helen, and Nigel Sabbarton-Leary. 2010. Are psychiatric kinds 'real'? *The European Journal of Analytical Philosophy* 6 (1): 11–27. https://doi.org/10.31820/ejap.

Bentall, Richard. 2014. What are we to believe about how we believe. *Philosophy, Psychiatry, and Psychology* 21 (1): 39–41. https://doi.org/10.1353/ppp.2014.0004.

Bogen, James. 2010. Noise in the world. *Philosophy of Science* 77 (5): 788–791. https://doi.org/10.1086/656006.

Bogen, J. 2011. 'Saving the phenomena' and saving the phenomena. *Synthese* 182: 7–22. https://doi.org/10.1007/s11229-009-9619-4.

Bogen, James, and James Woodward. 1988. Saving the Phenomena. *The Philosophical Review* 97 (3): 303–352. https://doi.org/10.2307/2185445.

Boorse, Christopher. 1977. Health as a theoretical concept. *Philosophy of Science* 44 (4): 542–573. https://doi.org/10.1086/288768.

Butts, Robert. 1984. *Kant and the double government methodology: Supersensibility and method in Kant's philosophy of science*. Dordrecht: D. Reidel Publishing Company.

Cassirer, Ernst. 1953. *Substance and function and Einstein's theory of relativity*. Dover Publications.

Cooper, Rachel. 2005. *Classifying madness: A philosophical examination of the diagnostic and Statistical Manual of Mental Disorders*. Dordrecht: Springer.

Evans, Bonnie. 2013. How autism became autism – The radical transformation of a central concept of child development in Britain. *History of the Human Sciences* 26 (3): 3–31. https://doi.org/10.1177/09526951134843.

———. 2017. *The metamorphosis of autism*. Manchester: Manchester University Press.

Eyal, Gil, Brendan Hart, Emine Onculer, Neta Oren, and Natasha Rossi. 2010. *The autism matrix*. Cambridge: Polity Press.

Fellowes, Sam. 2021. How autism shows that symptoms, like psychiatric diagnoses, are 'constructed': Methodological and epistemic consequences. *Synthese* 199 (1): 4499–4522. https://doi.org/10.1007/s11229-020-02988-3.

Hobson, Peter. 2002. *The cradle of thought*. London: Macmillan.

Hodge, Sally. 2004. A psychological perspective on theories of Asperger's syndrome. In *The many faces of Asperger's syndrome*, ed. Maria Rhode and Trudy Klauber, 39–53. London: Karnac.

Hollin, Gregory. 2014. Constructing a subject: Autism and human sociality in the 1980s. *History of the Human Sciences* 27 (4): 1–18. https://doi.org/10.1177/0952695114528189.

Kant, Immanuel, Paul Guyer, and Allen Wood. 1998. *Critique of pure reason*, The Cambridge edition of the works of Immanuel Kant. Cambridge: Cambridge University Press.

Maiese, Michelle. 2013. Autism, empathy, and affective framing. In *The philosophy of autism*, ed. Jami L. Anderson and Simon Cushing, 167–186. Plymouth: Rowman & Littlefield.

Massimi, Michela. 2008. Why there are no ready-made phenomena: What philosophers of science should learn from Kant. In *Kant and philosophy of science today*, ed. Michela Massimi, 1–35. Cambridge: Cambridge University Press.

———. 2011. From data to phenomena: A Kantian Stance. *Synthese* 182: 101–116. https://doi.org/10.1007/s11229-009-9611-z.

Murphy, Dominic. 2006. *Psychiatry in the scientific image*. Massachusetts: Massachusetts Institute of Technology.

Nadesan, Majia Holmer. 2005. *Constructing autism*. London: Routledge.

Natorp, Paul. 2004. *Plato's theory of ideas: An introduction to idealism*. Academia Verlag Sankt Augustin.

Rashed, M. 2010. Religious experience and psychiatry: Analysis of the conflict and proposal for a way forward. *Philosophy, Psychiatry, and Psychology* 17 (3): 185–204.

Rickert, Heinrich. 1986. *The limits of concept formation in natural science: A logical introduction to the historical sciences*. Cambridge University Press.

Silverman, Chloe. 2012. *Understanding autism: Parents, doctors, and the history of a disorder*. New Jersey: Princeton University Press.

Vaihinger, Hans. 1935. *The philosophy of 'As If': A system of the theoretical, practical and religious fictions of mankind*. 2nd ed. Routledge & K. Paul.

Verhoeff, Berend. 2013. Autism in flux: A history of the concept from Leo Kanner to DSM-5. *History of Psychiatry* 24: 442–458. https://doi.org/10.1177/0957154X13500584.

Windleband, William. 2015. Critical or genetic method? In *The neo-Kantian reader*, ed. Sebastian Luft, 271–286. London: Routledge.

Woodward, James. 2010. Data, phenomena, signal, and noise. *Philosophy of Science* 77: 792–803. https://doi.org/10.1086/656554.

5

Neo-Kantianism and Scientific Entities

5.1 Introduction

As discussed in Chap. 2, DSM and ICD psychiatric diagnoses have been criticised on multiple grounds. This has led many critics to call for currently employed psychiatric diagnoses to be abandoned and either replaced with new superior psychiatric diagnoses or replaced with an alternative that are not categorical psychiatric diagnoses. One common motive for such claims is that currently employed psychiatric diagnoses are not entities, things, objects or natural kinds. This implies that, if there actually are entities that psychiatric diagnoses could describe, then we should produce new psychiatric diagnoses which actually do describe those entities. Alternatively, if such entities actually do not exist, that is, there are no demarcated entities, then there exists nothing for those categorical psychiatric diagnoses to describe so categorical psychiatric diagnoses should be completely abandoned.

In this chapter I shall consider how neo-Kantians understand scientific entities. Neo-Kantians do not see science as describing mind-independent entities. Rather, they see scientists as constructing entities. This means that scientific theories should not be judged on whether they reflect

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_5

mind-independent entities. The scientific legitimacy of any particular scientific entity needs judging on some alternative grounds. I will outline criteria which neo-Kantians think should be used when constructing scientific entities. Therefore, that psychiatric diagnoses do not reflect mind-independent entities is not a reason to reject psychiatric diagnoses. What constitutes a good psychiatric diagnosis needs determining on the basis of some other criteria. Psychiatric diagnoses constructed on criteria which neo-Kantians employ for constructing scientific entities would not be arbitrary or merely invented.

In this chapter I shall outline common reasons why psychiatric diagnoses are considered not to be entities. I then develop a neo-Kantian account of psychiatric diagnoses and show what standards psychiatric diagnoses should be constructed on. I then show that constructing psychiatric diagnoses on those standards means that common criticisms of psychiatric diagnoses do not leave them scientifically illegitimate. Note that this chapter is exclusively concerned with the worry that psychiatric diagnoses might not be properly considered 'entities'. A related criticism, that people receiving a particular diagnosis may not be homogenous at the level of underlying causes, will be discussed in Chap. 6.

5.2 Concerns over Psychiatric Diagnoses

As outlined in Chaps. 1 and 2, psychiatric diagnoses have been described as arbitrary, constructed, an invention and made-up. They are also believed to not describe anything real and are believed to be scientifically illegitimate. Psychiatric diagnoses are not considered to be properly scientific, or are considered significantly inferior when compared to the scientific entities postulated in physics, chemistry and biology. People typically consider electrons, magnesium and zebras to be real and scientifically legitimate in a way ADHD, schizophrenia and autism are not considered to be real and scientifically legitimate. I will now outline common concerns over psychiatric diagnoses whilst leaving the concern that they lack homogeneous causes to Chap. 6.

Firstly, there is the problem of being polythetic, that psychiatric diagnostic criteria often only require someone to exhibit some symptoms

from a long list of possible symptoms (Kinderman et al. 2013, p. 3; Johnstone 2018, p. 39; Vanheule 2017, p. 85). To receive a psychiatric diagnosis an individual needs to only exhibit some, rather than all, the symptoms on the diagnostic checklist. Different individuals, who exhibit different combinations of symptoms, can thus receive the same diagnosis (the diagnosis is polythetic). This means that people who exhibit different symptoms are receiving the same diagnosis. Why are these people given the same diagnosis even though they do not have the same symptoms? Secondly, there is the problem of co-morbidity, that is, most people who receive one diagnosis typically qualify for more than one (Hyman 2010, p. 167; Kirk et al. 2015, p. 69). For example, around 45% of people diagnosed with autism meet the diagnostic criteria for major depressive disorder (Sterling et al. 2008, p. 1013) and around 40% meet the diagnostic criteria for anxiety (van Steensel et al. 2011, p. 309). Despite this, depression and anxiety are not parts of the diagnostic criteria for autism. Why is this? Thirdly, there is the problem of subthreshold individuals (Jablensky 2016, p. 27; Okasha 2009, p. 130). These are people who meet some but not all of the diagnostic criteria. Why are these people not diagnosed whereas others are? Fourthly, knowing that someone has a psychiatric diagnosis does not then reveal to what degree they meet the diagnostic criteria (Helzer et al. 2006, p. 1673; Hudziak et al. 2007, p. 21). A categorical system of psychiatric diagnoses registers people as either having or lacking the psychiatric diagnosis; it does not describe the extent to which they meet the diagnostic criteria. This means that someone who has every symptom on the diagnostic criteria and someone who only just meets the diagnostic criteria are not demarcated between. Fifthly, psychiatric diagnoses miss out important aspects of people (Johnstone 2014, p. 275; Vanheule 2017, p. 275). A psychiatric diagnosis primarily lists the symptoms an individual may have. It does not cover other important aspects like past life history, life goals, socio-economic situation, personality, etc. These are important for understanding the psychic development of people.

A related concern is the question of who decides, and how they decide, which psychiatric diagnoses there are. This is currently done by a committee of psychiatrists. For each successive edition of the DSM a committee decides, after reviewing evidence from various subcommittees, which

diagnoses should be retained, which should be replaced and which should be modified (Kendlar et al. 2011, p. 1149; Kincaid 2014, p. 151; Jablensky 2008, p. 90). This seems to be a decision making process, which seems at odds with the idea that scientists should discover which scientific entities exist. Scientists should discover rather than decide which scientific entities exist. A related concern is that psychiatrists cannot seem to make up their mind about which diagnoses are the correct ones given that they often modify the diagnostic criteria for various diagnoses, remove diagnoses and create new diagnoses. Both these points seem especially concerning given that they often have disagreements. For example, some members of the committee deciding whether to retain Asperger's Syndrome when producing the DSM 5 ended up quitting in protest at plans to remove it. All this raises concerns that psychiatrists are making these decisions based upon their own preferences rather than the solid evidence found in other sciences. If psychiatric diagnoses are the product of opinion rather than evidence then this can be taken to suggest that psychiatric diagnoses are not describing anything real and that they are not scientifically legitimate (Bentall 2004, p. 39; Boyle 1990, p. 169; Cushing 2013, p. 38; Horwitz 2002, p. 5; Read 2004, p. 21; Watson 2019, p. 2, see for discussion Raskin and Lewandowski 2000, p. 16).

There is a further concern, and perhaps the most worrying one, that psychiatric diagnoses are causally heterogeneous but I deal with this in Chap. 6. I shall now respond to these concerns by showing how neo-Kantians believe that scientific entities are constructed and that the opinions of scientists can legitimately play a role when constructing scientific entities. I shall then apply this to psychiatric diagnoses.

5.3 Neo-Kantians and Constructing Scientific Entities

Neo-Kantians believe that scientific entities are not ready-made objects in the external world but are constructed by us. The exact commitments of this claim need carefully outlining. Firstly, it is not the claim that nothing exists beyond human thought. Neo-Kantians typically believe that there

is something external to us. More controversial is whether whatever is independent of us consists of entities. Kant explicitly says that objects are mind-dependent but on rarer occasions he seems to suggest that there are mind-independent objects. Also, Kant talks of things in themselves and these are often understood as being mind-independent objects.[1] However, should they exist, such mind-independent objects are inaccessible to us. For neo-Kantians, we cannot know what they are like. Kant says that "they are nothing to us" (Kant 2002, 4:336). Rather, we can only know how they appear to us. As such, since they are inaccessible to us, we cannot judge scientific claims by comparing what scientists postulate with those mind-independent entities. For Kant it is appearances which are the basis of science, whilst the claim that mind-independent objects cause those appearances is a claim relating to something that is outside the realm of science and knowledge, relating instead to the transcendental. For neo-Kantians science does not consist of ready-made objects. Rather, neo-Kantians believe that we need to construct scientific objects. I will now outline this process of constructing scientific objects.

To neo-Kantians scientific objects are the result of a series of transformations. The starting point of those transformations is the sensations we receive from the world. To a neo-Kantian this is a chaotic, indeterminate stream, an "indeterminate manifold" (Rickert 1986 [1902], p. 36), a "pure flux" (Cassirer 1955 [1925], p. 32; see also Natorp 2004 [1921], p. 161; Vaihinger 1935 [1911], p. 53). This stream is inaccessible to us, we can cognise nothing from it, and were we to somehow have access to it we would only perceive sensations rather than objects. As discussed in Chap. 3, for Kant we only start having conscious awareness of sensation when we impose space and time onto that manifold of sensation. Doing this means that we start to see clumps of sensation as occupying the same area (they are located spatially) and they follow one another (they are

[1] However, some Kantians take what is known as a phenomenalist position whereby things in themselves are also mind-dependent (Jankowiak 2017). This is closer to an idealist reading and was how Kantians typically positioned themselves in the years following Kant but is a less common position among Kantians today. Also, note that there is a stereotypical notion that Kant associated noumena and things in themselves but one reason why they are different is that things in themselves can be seen as mind-independent whilst noumena can be simultaneously seen as mind-dependent (and this is the position held by one of the most realist, and least idealist, readings of Kant (Allison 1983)).

located temporally). This still does not give us objects. To start to see objects we need to employ concepts, forming spatio and temporally located clumps of sensations into objects. We start positing that some clumps of sensation take the form of objects, a spatio-temporally located substance which cannot ever be in the same spatio-temporal position as another substance. At this point we move from spatio-temporally located sensations to spatio-temporally located objects with properties like shape, duration and mass. We bring those sensations under concepts in a manner that makes them cognisable and measurable, and doing this gives them characteristics. As outlined in Chap. 4, this produces what I call the observable world, whereby what we observe is conceptually constituted. This, however, is not science. To make scientific claims we need to apply further concepts to that conceptually constituted observable world to provide more precise properties, and in doing so we might start moving away from the observable world. Kant details how we bring our sensations under mathematical principles, such as constituting objects with measurable characteristics like mass, length and velocity. All this gives form to the sensation. The key claim is that those objects do not have those properties outside of us constituting them with those properties. Doing this creates something that must adhere to the parameters set by the *a priori* concepts being applied.[2]

In general, Kant thinks the faculty of the understanding supplies concepts which set the form that objects must take, and so creates a world of experience, but the faculty of the understanding and the concepts it supplies are not enough. The concepts we apply specify that objects must have shape, size and colour but do not specify types of objects. We might perceive a round object of a certain size and colour but this does not establish which type of object it is, such as whether it is an apple or a ball. Having applied concepts which specify the form that objects must take, we are still left with an "immeasurable multiplicity of things" (2000,

[2] This first step relates to the faculty of the understanding which provides the categories. Technically, for Kant there is another step to this transformation. Firstly, we employ the faculty of intuition to structure sensation into space and time, then we employ a different faculty, the faculty of the understanding which employs the categories, to structure those spatio-temporal sensations into objects. Kantians after Kant have questioned this notion of intuition and because I am not discussing space and time I shall not mention it again.

20:215–216) and "a great heterogeneity of forms of nature" (Kant 2000: 9:203). Kant thinks that having constituted sensation with form through applying concepts to it from the faculty of the understanding we also then need principles of reason to constitute what I shall call thought objects.[3] A thought object does not correspond to anything we can experience. Thought objects abstract and idealise away from that world of experience. To specify which specific types of objects there are we need to use principles of reason that come from the faculty of reason. This is Kant's general approach though he takes some slightly different stances in different texts.[4]

Kant provides two examples of thought objects. Firstly, he outlines what he calls pure elements which are "pure earth, pure water, pure air" (Kant 1998, A646/B674). These are thought objects because any chemical we come across will be impure because it consists of more than one element. We can only access a pure chemical, one which solely consists of one element, in thought. As such, it is a thought object, an intelligible object, a mental creation. It is an idealisation which we can never experience. Secondly, Kant gives the example of the orbit of the planet and comets. Kant describes how we might see celestial objects, such as planets, moving in a circle (Kant 1998, A663/B691). However, as we see inconsistencies and variation in their orbit we hypothesise that they are not a circle, that is, their movement is not fully circular. To move beyond

[3] Meer (2018, p. 343) outlines various terms used such as intelligible objects, thought objects and object of ideas. The key issue is that they are the product of the faculty of reason. They contrast with the product of the faculty of the understanding, which can be translated as 'sensible object' or 'object of sensibility'.

[4] Kant makes some different claims about this in different texts. In the *Critique of Pure Reason* Kant seems to see the faculty of the understanding as sufficient for formulating knowledge of the world meanwhile the faculty of reason can only order the claims produced by the understanding rather than contribute to actually formulating empirical claims. This position seems stronger in the *Metaphysical Foundations of Science* where it potentially looks like reason is not required to make empirical claims. However, in the *Critique of the Power of Judgement* Kant seems to suggest that the faculty of reason can contribute to empirical claims above and beyond those made by the faculty of the understanding. He even seems to suggest that all empirical claims need input from the faculty of reason. Whether Kant changed his mind or not is of some dispute, as is whether these claims can be successfully integrated. In this book I take Kant's position in the *Critique of the Power of Judgement* to either be compatible with or to overrule his position in earlier claims. However, were priority placed upon his earlier texts then this makes no difference to my argument since empirical claims which only require the faculty of the understanding, and do not need the faculty of reason, appear to only be those found in physics.

observations we hypothesise that they are adopting a different path, such as an ellipse or a parabola. We cannot observe these or experience them as sensible bodies, or how sensible bodies are acting. As such, they are a thought object, or an idealisation, which we cannot experience. When constructing ellipses and parabola, we move from our observation of numerous bodies and then correct what we observe, moving from our observation of the circle to the corrected thought entity of an ellipse or parabola.[5]

Neo-Kantians typically do not adhere to a sharp separation between a faculty of the understanding and a faculty of reason. Neo-Kantians, unlike Kant, believe that reason can play a role in determining what concepts to employ. They still heavily endorsed the notion that science deals with idealisations and abstractions which do not resemble what we see in the world around us. Cassirer describes how

> [w]e investigate the impact of bodies by regarding the masses, which affect other, as perfectly elastic or inelastic; we establish the law of propagation of pressure in fluids by grasping the concept of a condition of perfect fluidity; we investigate the relations between the pressure, temperature and volume of gas and comparing a hypothetically evolved model to the direct data of sensation. (1953a [1910], p. 130)

Also, Kantian philosopher Butts outlines how triangles and numbers are not sense data or properties of things in themselves, rather, they are idealisation (1984, p. 187). I outline in Chap. 4 how Massimi's account of Galileo's experiments involved idealisations and abstractions. Natorp also outlines how Galileo used abstractions and idealisations when doing astronomy (2004 [1921], p. 210). Finally, Rickert says that "science is possible only by means of the reshaping [of what we observe] by the subject" (1986 [1902], p. 217–218). Whilst using different terminology, they also thought science involved thought objects.

Since we cannot obtain thought objects in observation we need to employ something beyond observation to construct thought objects. Kantian philosophers of science have claimed that principles that come

[5] Another way to put this is that we subsume the circular motion under a higher thought entity of an ellipse, which is in turn subsumed under the even higher thought entity of the parabola.

from us should be used to construct objects (Buchdahl 1969, p. 658; Butts 1984, p. 267; Cassirer 1953b [1923], p. 301; Morrison 1989, p. 162; Natorp 2004 [1921], p. 161; Rickert 1986 [1902], p. 114; Windleband 2015a [1910], p. 323). Neo-Kantians believe that we need to employ principles to form thought objects. I shall outline the principles which Kant and neo-Kantians think we should employ. As outlined in Chap. 3, Kant outlines three principles of reason, those of specificity, homogeneity and affinity.[6] These principles of reason are used to refine sensible objects, correcting them and improving them. Those principles transform the world of sensibility by abstracting away some aspects of sensibility and adding connections which are not present in sensibility. I will now describe how these principles can be employed to form scientific entities.[7]

Kant believes we should use the principles of homogeneity to reduce the complexity of empirical objects. It simplifies the observable world by abstracting away from particular features of particular entities by subsuming them under general classes. Thus, one red spherical object has its particular features abstracted away to become an instance of the general class 'ball' whilst a different red spherical object has its particular features

[6] Kant talks of the principles of homogeneity, specificity and affinity in the *Critique of Pure Reason* but talks of genera and species in the *Critique of the Power of Judgment*. I take homogeneity to be equivalent to genera, and specificity to be equivalent to species.

[7] There is a long running debate over whether principles of reason are purely regulative or whether they are also constitutive. They can be easily understood as regulative, which means that they sort empirical claims into a system of science, rather than being constitutive, which is where empirical claims are created. On Kant's position it is easy to see how the faculty of the understanding constitutes empirical claims whereas the faculty of reason plays only a regulative role of ordering empirical claims produced by the faculty of the understanding. However, Kant simultaneously seems to suggest that principles of reason play both a regulative and a constitutive role. He also seems to suggest regulative principles are required to formulate empirical concepts (these claims seem present in the second edition of the *Critique of Pure Reason* and are stronger in the *Critique of the Power of Judgement*). There is much dispute over how this should be understood. This is a critical matter because Kant sees constitutive as being in some sense objective whilst regulative are described as both subjective and objective. So on one reading, some empirical claims are formulated on constitutive principles so are objective meanwhile other empirical claims are formulated on regulative principles so are subjective. On another reading, all empirical claims require regulative principles and, since regulative principles are subjective, all empirical claims are subjective. On a third reading, all empirical claims require regulative principles but, since regulative principles are objective in some manner not fully articulated by Kant, empirical claims are objective. For further discussion see Brittan 1991; Buchdahl 1967; Butts 1984; Friedman 1992; Grier 1997; Guyer 1990; McLaughlin 2014; Morrison 1989; Ostaric 2009.

abstracted away to become an instance of the general class 'apple'. We can thus order sensible objects into thought objects through the principle of homogeneity.

Kant believes we should use the principle of specificity to specify the way in which things are different. The idea of specificity is that it is always possible to differentiate generic concepts into two or more specific ones. We can always produce more fine-grained descriptions or concepts to apply to what we observe. For example, we can move from employing a generic concept like planet to a more specific concept like 'gas giant' or 'icy moon'. We can further specify the type of 'gas giant' depending upon, for instance, the type of gas which is predominant, and further specify gas giants of a particular type of gas by, for instance, level of gravity or number of moons. These are also idealisations since we do not experience planets as particular types, just like how we do not experience spherical objects with a red colour as a type of fruit or type of apple. The sensation we get when looking at these things significantly constrains how we should think about them but it does not determine what instances of entities they are. Gas giants and icy moons are thought objects rather than sensible objects.

Kant argues that we should use the principle of affinity to order the thought entities produced by the principles of homogeneity and specificity. The thought objects produced by the principle of homogeneity should be seen as more generic versions of the thought objects produced by the principle of specificity (similarly, the thought entities produced by the principle of specificity are more specific versions of the thought entities produced by the principle of homogeneity). Doing this produces a systematically ordered science whereby each thought entity, whether a more specific one or a more generic one, has a specified relationship to other thought entities.

Kitcher employs similar principles to those of homogeneity and specificity but provides more detail on how they should be integrated together into a system of science. He describes how an ideal system of science should cover as much phenomena as possible in as much detail as possible whilst using as few theories as possible. These two goals pull in the opposite direction. It is easier to describe a smaller number of phenomena when only using a small number of theories. Similarly, many theories

make it easier to describe a large number of phenomena. This has some similarity to Kant's notion of homogeneity and specificity since homogeneity abstracts away detail, leading to a smaller number of entities being posited, whilst specificity increases detail, leading to a larger number of entities being posited. This could lead to a tension, with scientists facing the conflicting aim of positing both more and less entities. Kitcher argues that the correct entities are those which find a place in a system of science which best balances those conflicting principles. Since those principles come from us, and since it is us who decide how best to balance them, the correct entities are mind-dependent. Cassirer also seems to endorse this general approach when he says that science should be based upon the principles of generality and simplicity. Generality demands that as much phenomena are incorporated into the system as possible (both amount of data and closeness to the data) whilst simplicity demands as few entities are employed as possible (Cassirer 1953a [1910], p. 308).

The principles of homogeneity and specificity, or similar principles, need not be endorsed to take a neo-Kantian position. For example, Massimi describes natural kinds as "functionally relevant clusters of properties" (2014, p. 428). On such an account we formulate natural kinds based partly upon them playing a functional role in explaining a range of phenomena, the natural kind being created to posit an entity which causally explains the result of scientific experiments, and partly upon applying theoretical virtues such as simplicity, coherence and scope. Also, natural kinds need to be formulated to fit within, and made consistent with, a wider system of science. Additionally, natural kinds should be formulated to create a tractable system of science, whereby scientists can approach the complexity of the natural world in a manner which suits our cognition. As such, the entities that a system of science forms should respect the principles of simplicity, consistency and embedding.

This means that objects need constructing on a Kantian approach. Reichenbach writes that "[i]t was Kant's great discovery that the object of knowledge is not immediately given but constructed" (1965 [1920], p. 49). Cassirer writes that "objects are not 'given' to consciousness in a rigid, finished state, in their naked 'as suchness', but that the relation of representation to object presuppose an independent, spontaneous act of consciousness. The object does not exist prior to and outside of synthetic

unity but is constituted only by this synthetic unity" (1955 [1925], p. 29). Natorp talks of "the task of putting together the object out of its components" (Natorp 2004 [1921], p. 343). Rickert writes that "[w]e must abstract from its unique distinctiveness and particularity in order to find our way in the real [by real Rickert means sensation]. Otherwise the infinite manifold of its content eliminates any possibility of orientation" (Rickert 1986 [1902], p. 42). Neo-Kantianism does not see mind-independent objects as scientific objects. Rather, scientific objects are those we construct. We do this by taking what comes to us from the world and bringing it under concepts, transforming what comes to us from the world into sensible objects. We then apply further principles to transform those sensible objects into thought objects. Also, as outlined in Chap. 3 neo-Kantians typically see these principles as being, or as related to, a standard of truth because they are based on reason. This means that the correct scientific objects are not mind-independent ones but are mind-dependent ones which we have created. I shall now apply this approach to psychiatric diagnoses.

5.4 Psychiatric Diagnoses as Kantian Thought Objects

The starting point of constructing psychiatric diagnoses is observations of people. As outlined in Chap. 4, we observe people around us. These observations are the product of combining what comes to us from the world with concepts. This means our observations are conceptually constituted. We see people interacting with the world and with one another through applying various concepts. Additionally, as outlined in Chap. 4, we see people as having various characteristics, such as being angry or sad. I then outlined how symptoms are more refined characteristics, produced by bringing observations under more specific concepts than are unconsciously employed in everyday observations. In this chapter I will outline an alternative way of thinking about people. Whereas in Chap. 4 I described the process of assigning people characteristics in this chapter I will outline the process of assigning people to different types of person.

Psychiatric diagnoses group together symptoms. This means psychiatric diagnoses are already some distance from what we observe. As described in Chap. 4, symptoms are produced by bringing observations under concepts. This transforms what we observe, moving from imprecise descriptions to more abstract descriptions which removes some of what we observe and joins together parts in a manner which is distant from what we observe. Psychiatric diagnoses then produce a further transformation, taking the symptoms which are present in particular individuals and rearranging them to produce psychiatric diagnoses. As outlined in Chap. 2, psychiatric diagnoses are polythetic whereby there are multiple combinations of symptoms which can result in an individual meeting the diagnostic criteria. This means an individual can receive the diagnosis despite not exhibiting all the symptoms. Thus the psychiatric diagnosis is a thought entity which is distant from particular individuals who receive the diagnosis. The symptoms associated with the psychiatric diagnosis are different to the symptoms exhibited by people who are assigned the diagnosis because the psychiatric diagnosis is polythetic, that is, someone need not exhibit all the symptoms of the diagnostic criteria to receive the diagnosis. This means a transformation has taken place, resulting in a thought object which is distant from what we observe. I now describe how the symptoms of particular individuals are transformed into psychiatric diagnoses.

In everyday observations we may observe others as being particular types of person. After observing others for a length of time we might start seeing others as unfriendly types of person, sly types of person, reliable types of person, dangerous types of person, intelligent person, etc. The basis for this will be the characteristics we believe others to exhibit. This could be based upon assigning an individual a single characteristic, such as considering someone who exhibits the characteristic of anger to be an angry person. Often, someone is considered to be a type of person if they are believed to exhibit a number of characteristics. For example, someone might be considered to need to exhibit the characteristics of integrity, honesty and consistency to be a trustworthy person. This process of assigning people into different types will typically be imprecise and will often be unconscious though it can involve conscious thought. These judgements can be formed very quickly or over a period of time. Also,

they will typically involve many assumptions which will have been influenced by their culture, their social environment and their upbringing. These assumptions might also change over time and the characteristics of people we observe can change over time.

Psychiatric diagnoses also place people into types of people. However, psychiatric diagnoses make this process much more precise compared to everyday placing of people into types. They bring our observations of others under principles, turning sensible objects (everyday observations of people) into thought objects (psychiatric diagnoses). They transform our observations into scientific objects. This is because the psychiatric diagnosis is distant from the individual. As described, psychiatric diagnoses are polythetic, meaning the diagnostic criteria can be met in multiple ways. This means the psychiatric diagnosis groups together symptoms in a manner which is different to the symptoms individuals actually exhibit. This heavily fits the neo-Kantian notion outlined in Chap. 3 that we need to transform what we observe when doing science.

I now outline this process in more detail, describing it in terms of the two aspects of scientific knowledge I outlined in Chap. 3 and in this chapter. The first is to conceptually constitute the world through applying a set of concepts to data. Scientists employ a relativise *a priori* to constitute a world of objects. The second is applying principles to determine which specific objects there are. I start by outlining this first step.

5.5 Categories and Dimensions

The *a priori* provides a form to objects. This is the notion of the *synthetic a priori* (something independent of experience is applied to experience). The object itself lacks the form and only gains it through the *a priori*. Whether we think of objects in terms of things in themselves or what we observe in the world around us, our knowledge of objects is restricted to applying a form through the *a priori*, providing objects with properties that they otherwise do not have. In doing so, this makes knowledge possible. Through constituting objects with various characteristics, we are able to then learn about those characteristics and thus gain scientific knowledge. Cassirer describes such *a priori* principles as "a means of

measurement… [that] serves to make possible this very order" (Cassier 1953a [1910], pp. 140–141; see also Rickert 1986 [1902], p. 144; Reichenbach 1965 [1920], p. 53).

I will emphasise three different ways of giving people form. These are the three different ways of thinking about psychiatric diagnoses which I outlined in Chap. 2, namely categorical, dimensional and person centred. Each of these should be considered different sets of constitutive principles whereby they act as measurement systems that provide properties to objects.

For a categorical approach I will consider DSM 5 autism [known as Autistic Spectrum Disorders]. I do this because it is well known rather than because I think DSM 5 notions of autism are the best approach to categorical notions of autism. Anyone who is correctly diagnosed will exhibit at least a certain number of symptoms on one list ("Persistent deficits in social communication and social interaction across multiple contexts" (APA 2013, p. 50)) and at least a certain number of symptoms on a second list ("Restricted, repetitive patterns of behavior, interests, or activities" (APA 2013, p. 50)). According to the DSM, if they exhibit a sufficient number of symptoms upon one list but not the other then they are not autistic. Additionally, the symptoms need be present to at least a certain degree. In contrast, on a dimensional approach autism extends far beyond the categorical description in the DSM. Autism consists not just of DSM descriptions but manifestations of the symptoms and causes of autism in the entire population. Note that categorical approaches only cover people who pass a cut off whereas dimensional approaches cover people regardless of if they meet any categorical cut off, meaning dimensional notions of autism include anyone covered under the categorical notions of autism. In contrast to categorical and dimensional diagnoses, on a person centred approach each individual is not part of a general class of a psychiatric diagnosis. Each person is simply unique and not placed into a general class which other people are also members of.

These are thus three different ways of constituting people (firstly, either in or not in a general class, secondly, in a general class to a particular degree, thirdly, no one is placed into a general class). The way in which we understand people to have psychiatric diagnoses will differ based upon which constitutive principle is employed. This is not intended to describe

all possible ways that psychiatric diagnoses can take a constitutive form, rather, I outline currently popular ways of thinking of psychiatric diagnoses and I argue these should be understood as constitutive principles. Additionally, this is not to say that one of these is the only constitutive principle needed, since we also need more basic constitutive principles such as those relating to forming basic notions of objects and causes.

Neo-Kantians typically see *a priori* principles as stemming from reason. This provides basic forms for dividing up what we experience (Cassirer 1953a [1910], p. 269; Natorp 2004 [1921], p. 348; Reichenbach 1965 [1920], p. 87). These forms are logical principles that can be used to divide up data and arrange data. In psychiatry we can bring data under logical structures to make measurements (1) relating to present or absent (categorical), (2) relating to present to a particular degree (dimensional) or (3) not asking questions of presence or absence or present to a degree (person centred). The form of categorical, dimensional and person centred we apply will make the data attribute different properties, as I outline below. In this sense the form makes knowledge possible, whereby the form employed means different properties are entailed.

We can take any set of data and apply these principles to them. Also, having applied one of these principles to constitute a form, we can always apply further principles to produce a more specific grouping. If we have a dimensional grouping whereby members of the group are rated as being present to a degree we can always put upon this a further categorical divide whereby we demarcate between those who score above a particular figure and those who score below that figure. Similarly, once we have a categorical divide we can further dimensionally rate those who meet the categorical criteria. We can then add categories onto this to create subdivisions and then dimensionally rate the degree that the subdivisions are present. Additionally, at any point throughout this process we can stop considering groups of individuals and instead focus upon a particular individual. We can then rate what characteristics the individual has without then considering how the individual relates to the broader grouping. Having then focused upon an individual without considering the wider group we can still rate the particular attributes that the individual exhibits as categorically present or dimensionally present to a degree. Here I

am making a logical point that it is always possible to add categorical and dimensional divides.[8]

It is certainly true that when applying a particular measurement we might find a categorical or dimensional approach fits much better. For example, if we take Huntingdon's as a measurement we find that everyone either does or does not have the gene which causes Huntingdon's. This fits a categorical approach. Similarly, if we take the vast majority of current DSM diagnoses we find that we can measure the degree they are present far beyond the categorical boundaries imposed by the DSM (that being, lots of people score above zero without meeting the categorical criteria). This fits a dimensional approach. Also, we might conduct statistical analysis of a particular domain using methods like factor analysis and find that our statistical and mathematical models produce categories (discontinuous groupings) or dimensions (continuous results without demarcated groupings). However, my point is that we can always add a further category or dimension onto these. Huntingdon's might seem like a clear example of a category given that everyone with the gene develops Huntingdon's and no one without the gene develops Huntingdon's. However, we can further demarcate Huntingdon's by giving a more specific picture of the exact causal details of how the gene interacts with other parts of the brain and body to produce the symptoms, and a more specific picture of which symptoms manifest and exactly how they manifest (I will provide detailed explanations of this idea below and in Chap. 6). Logically, it is possible to take Huntingdon's and add further categories and dimensions onto it. The gene for Huntingdon's provides a major constraint upon what modifications we can make, making Huntingdon's much more constrained than most psychiatric diagnoses, but this constraint does not prevent further categorical or dimensional distinctions being made. This is not to say that it is scientifically worth making such further distinctions, rather, determining when we should or should not add categories and dimensions depends upon a standard. Specifically, I

[8] The only ultimate stopping point, in classical physics at least, is that no two things can occupy the same place in both space and time. Until this point is reached we can always impose a categorical divide upon whether something is or is not present at a particular location in space and time, and always impose dimensional rating based upon distance in space and time from a particular point in space and time.

appeal to notions of systematicity which I outlined in Chap. 3 and will consider again later in this chapter and in Chap. 6.

On this understanding, reality understood as a mind-independent world independent of what we contribute is not categorical, dimensional or person centred in nature. Categorical, dimensional and person centred approaches are ways of constituting the world. We cannot see reality as it is, only as it appears to us. Categorical, dimensional and person centred are ways of making the world appear to us, they are ways of giving form to the world around us. I do not think it makes sense to say that autism is in reality categorical, dimensional or person centred. When saying autism exists in reality categorically then we mean that people either do or do not exist as an instance of autism. This, however, faces the problem that we can say that any person exhibits only degrees of autism (dimensional approaches) rather than as present or absent. We can also say the person exists in many other ways than is covered by the diagnosis of autism and those ways cannot be neatly demarcated from the symptoms of autism (person centred approaches). When saying that autism exists in reality dimensionally then we mean that people exhibit autism to a particular degree. This faces the problem that we can always place further more specific groups onto the dimensions (categorical approaches) and the above mentioned problem of person centred. If we say people exist in a person centred manner which is not part of a general class then we face the problem that the person seems to have commonalities with a general grouping of people we call autistic, whether we think of that general grouping as categorical or dimensional. When it comes to questions of reality I struggle to make sense of the question of whether people are in reality one of these forms but not the others.

Similarly, on a realist picture we might think there really is an entity we call Huntingdon's which exists in reality in a categorical form. However, we can add further categorical subdivisions whereby we demarcate between those who have the gene which causes Huntingdon's and have a particular symptom and those who have the gene but do not have that symptom. Similarly, we could dimensionally rate people with Huntingdon's based upon number of symptoms associated with Huntingdon's they exhibit. Additionally, though I can only show this in Chap. 6, we can make further causal categorical and dimensional

subdivisions of Huntingdon's whereby we demarcate based upon the specific causal interactions which the gene that causes Huntingdon's takes place in. Which then is the real Huntingdon's: the current categorical diagnosis or one of these modifications? The key Kantian idea is that because whatever we posit can be further broken down categorically and dimensionally, the mind-independent world cannot determine what is the correct answer. It may certainly be true that, given principles that we apply to the world, that there is very good reason to link together the gene and the symptoms we currently associate with Huntingdon's in the manner that we currently do rather than make more specific demarcations. However, it is us rather than the world which rules out making further categorical or dimensional divides.

As such, we cannot say that, for example, the nature of reality which is studied by psychiatry is dimensional rather than categorical in reality. In this regard I reject a scientific realist understanding of autism, if we assume scientific realism to mean describing how things are independent of our concepts. On a scientific realist position, we might have to use concepts when working out what properties objects have but those objects have those properties independently of us applying those concepts. In contrast, on my position, objects only gain those properties which we can know about through applying a form to constitute them with properties that it otherwise would not have. As such, autism only gains the properties it has through our act of constituting it with those properties by bringing data under forms. However, a typical scientific realist type of thinking seems prevalent within the history of psychiatry and modern psychiatry. Whether people have thought that our psychiatric diagnoses succeed in describing reality or not, mainstream psychiatrists have effectively thought that, depending upon the era and using different terminology, that reality took a categorical, dimensional or person centred form. I outlined in Chap. 2 how Kraepelin effectively used categorical approaches, psychoanalysts primarily used dimensional approaches and the neo-Kraepelians primarily used categorical approaches. Meanwhile, psychiatry seems to be moving into a dimensional approach with the advent of RDoC and HiTOP. Each of these groups has effectively seen reality as being categorical or dimensional and thus favoured one approach over the other. I argue this position is

untenable. On a neo-Kantian position they are constitutive of scientific data, giving scientific data form from the chaos of sensation we receive from the world. Therefore, we cannot say the scientific data shows the world to be categorical, dimensional or person centred in reality independent of what we apply.

On a Kantian position, what we apply to the world makes knowledge possible. Following this, categorical, dimensional and person centred in some sense make knowledge possible. These make different knowledge possible by entailing different groupings. To show this general point I will highlight it by comparing categorical and dimensional notions of autism. Doing this allows me to address notions that we should make the DSM dimensional. I suggest that there would still be good reason to put categories upon dimensional DSM diagnoses since they make possible different knowledge. Assuming this argument works, it would likely not impress advocates of RDoC and HiTOP who think dimensionally rating DSM diagnoses would be still too broad and that more homogeneous dimensions need to be employed. I will also consider how categories could be employed in relation to RDoC and HiTOP. I start by considering inductive knowledge made possible and then consider causal knowledge made possible.

5.5.1 Inductions Associated with the Entire Unit

The basic idea of inductions is that knowledge of what happened in the past and what happens in the present gives us probabilities about what is expected to happen in the future. For example, we have boiled water in the past and can boil it in the present. We can measure at what temperature water boils in those instances and use this to build probabilities about what temperature it will boil at in the future. We have many measurements of water boiling and if those measurements are always 100 degrees then we can assign a very high probability to water boiling at 100 degrees in the future (more realistically, we might assign different probabilities based upon differences in altitude and difference in purity of the water, and we might get measurements quite different to 100 degrees if the measurement device is broken). Many inductions in science produce

much lower probabilities and have probabilities relating to multiple outcomes, such as how a particular gene might have a probabilistic relationship of 1% to one symptom, and 2% to a different symptom. That being, one percentage of people studies so far have exhibited the first symptom and 2% have exhibited the second symptom. To associate autism with a set of inductions is to say that autistic people are likely to manifest certain symptoms and certain causes to a certain degree and in particular ways (also, potentially in certain circumstances). We can say that 80% of autistic people studies so far have exhibited a particular symptom, 45% have exhibited another particular symptom and 3% have exhibited another particular symptom. We can do the same with causes. I now outline how different inductive knowledge is produced by categorical and dimensional approaches to autism.

Each way of making knowledge possible produces a different group which then makes possible a different set of inductions. Here I am considering the whole unit, that being, everyone who is categorically autistic or everyone who is dimensionally autistic (I consider subsets in the next section). On a categorical approach, autism consists of everyone who has been properly diagnosed according to the categorical diagnostic criteria. Everyone accurately diagnosed will exhibit at least a certain number of symptoms and exhibit those symptoms to a certain degree. In contrast, on a dimensional approach which measures the traits in the population, autism consists of manifestations of the symptoms of autism in the general population regardless of if they meet categorical notions of autism. This means the dimensional grouping covers everyone within categorical notions and anyone else exhibiting any symptoms of autism. In dimensional autism, these symptoms will occur in many individuals who exhibit a smaller number of symptoms and exhibit them to lower degrees than is required to meet the categorical diagnostic criteria. This means that the total number, manner of manifestation and degree of manifestation of symptoms exhibited by the combined population of categorical notions of autism is different to that exhibited by the combined population of dimensional notions of autism. The same is true in relation to causes whereby someone who is categorically autistic is more likely to exhibit causes associated with autism than someone who is dimensionally autistic. As such, each way of constituting autism produces different inductive

units, whereby the inductions associated with categorical notions of autism are different to that associated with dimensional notions of autism. If we ask how common a particular symptom or cause is within autism then we get very different answers when considering categorical or dimensional notions of autism.

This is a very broad way to think about dimensions and most psychiatric researchers will recognise this. However, this broad approach does occur. It is sometimes claimed that everyone is a bit autistic, a bit OCD or a bit ADHD (for discussion see Koi 2021; Spencer and Carel 2021). Under a dimensional approach, this claim is largely true. Everyone who scores above zero for a symptom or cause of autism would be considered dimensionally autistic. In contrast, under a categorical approach, this statement is very much false. Vast portions of the population do not meet the categorical criteria for these diagnoses. Everyone is not a bit autistic, OCD or ADHD on categorical approaches. Relatedly, some criticise the diagnosis of schizophrenia because lots of people in the population hear voices in ways that seem to be non-pathological. We can put everyone on a schizophrenia spectrum and in doing so we will find aspects of schizophrenia in the wider population but there is a clear difference in number of symptoms between someone who meets the categorical diagnostic criteria and many people who occasionally hear voices. To say everyone is a bit autistic is true or false is not a question that we can establish by simply studying reality, rather, it depends upon if we categorically or dimensionally constitute autism.

5.5.2 Inductions Associated with Specific Parts of the Unit

We can focus on particular parts rather than the entire grouping. Rather than consider everyone who meets the categorical diagnosis or consider everyone who is dimensionally autistic, we can instead narrow the inductive unit by being more specific. We can consider anyone who exhibits a particular symptom or particular cause, or exhibits a particular symptom or cause above a particular threshold. Similarly, we might consider anyone who exhibits a particular combination of symptoms, combination of

causes or combination of both. Doing this allows inductive claims about the likelihood of particular symptoms being exhibited given the presence of any other particular symptoms and particular causes. For example, we can measure the probability that someone will exhibit symptom X based upon knowing they exhibit symptoms Y and Z. This then provides inductive knowledge whereby knowing that someone exhibits some specific symptoms means they have an inductive probability of exhibiting other specific symptoms. We can also establish probabilities between presence of specific symptoms and presence of specific causes, and similarly we can establish inductive probabilities between specific causes being present and specific symptoms or specific causes being present.

By specifying a narrower grouping rather than the entire grouping of categorical or dimensional autism it might seem that we can consider a group that is identical regardless of whether categorical or dimensional notions are employed. For example, we could study literally everyone who exhibits, say, repetitive thoughts and executive dysfunctions. However, the total symptoms exhibited by people within that group differ depending upon whether autism is constituted categorically or dimensionally. If someone exhibits repetitive thoughts and executive dysfunctions and is categorically autistic then they will exhibit many other symptoms of autism. In contrast, if people exhibit repetitive thoughts and executive dysfunctions but are dimensionally autistic then some may exhibit many other symptoms of autism but some may only exhibit few other symptoms of autism. As such, even when looking at more narrow groupings, statistical associations between symptoms will change depending upon whether autism is constituted categorically or dimensionally. This means different inductions are made possible by those forms. Rather than impose a further categorical divide, we might simply say that the higher someone scores upon the dimensions the more likely they are to exhibit more symptoms and causes of autism. However, two people who score, say, 80 on a dimension can still differ from one another in ways that a further categorical divide might bring out. For example, two people might both exhibit eight out of the nine symptoms on the diagnostic criteria for major depressive disorder but they differ in which specific symptom they miss out. A categorical divide upon the

presence or absence of that symptom might helpfully convey important inductive and causal probabilities.

5.5.3 Symptoms Causally Interacting with Other Symptoms

I now consider the role that categorical and dimensional psychiatric diagnoses play in making causal knowledge possible. I shall discuss here how categorical and dimensional diagnoses specify different groupings which then results in different statistical probabilities of how symptoms interact with causes (in Chap. 6 I will consider how categorical and dimensional diagnoses specify different statistical probabilities of causes interacting with other causes). Also, whilst here I focus upon causal rather than inductive knowledge, it is worth noting that causal changes can then entail new inductions, which in turn can then produce new inductive knowledge.

Symptoms can causally interact with one another and this can then alter symptoms. This works in two different ways. Firstly, one symptom causes another symptom to occur. The first symptom is present independent of a second symptom whereas the second symptom is brought about by the first. For example, people diagnosed with major depressive disorder can exhibit the symptoms of low mood and difficulty sleeping (Borsboom and Cramer 2013, p. 115). Rather than seeing difficulties with sleeping as caused by biological or environmental causes, people who exhibit low mood can often worry and worrying can cause difficulties sleeping. Alternatively, biological or environmental causes might predispose someone to rarely manifesting difficulties with sleeping but the additional presence of low mood results in regular manifestation of difficulty sleeping. Secondly, two independently present symptoms can interact in a manner which changes both of them. For example, evidence suggests that the simultaneous presence of symptoms associated with autism and ADHD does not simply result in both sets of symptoms, rather, some of the symptoms start expressing in a qualitatively different manner (Petrolini and Vicente 2022 p.14). This shows how symptoms can causally interact with one another.

Categorical notions of autism and dimensional notions of autism have different probabilities of symptoms being present which then relate to different probabilities of causal interactions between symptoms. Someone who is categorically autistic has a much higher probability of exhibiting many symptoms of autism and exhibiting them to a higher degree compared to someone who is dimensionally autistic. As such, given the greater probability of more symptoms being present to a higher degree, there is much higher likelihood of symptoms interacting with one another in categorical rather than dimensional notions of autism. This then alters which causal interactions take place within each way of constituting autism, constituting different inductive and causal probabilities.

5.5.4 Symptoms Interacting with the Environment

Symptoms can interact with the environment which can then change the symptoms. Whilst this is true in general, here I aim to show that different symptom and environmental interactions can occur depending upon whether autism is constituted categorically or dimensionally.

Environments can impact symptoms in many different ways. Firstly, there are statistical correlations between early interventions and later exhibitions of autism. The earlier a child engages in support service the lower the number or degree of symptoms exhibited (Clark et al. 2018, p. 93; Dawson et al. 2012, p. 1159). Secondly, support can give someone the tools which they can potentially use to influence how and when they exhibit their symptoms, and may also alter their desire to start or stop exhibiting symptoms. Walsh et al. write that "environmental inputs may help them [autistic individuals] to learn social communication behaviours and to adjust to their surroundings and, more generally, may minimize (or exacerbate) their symptoms" (2011, p. 606; see also Kite et al. 2013, p. 1698; Johnstone, 2014, p. 280; Vanheule 2017, p. 97). For example, an autistic person might learn a coping strategy that prevents an autistic meltdown, and they might decide to try and avoid monologuing in conversations when they realise that other people do not enjoy this. Thirdly, symptoms can be reactions to environments. For example, autistic people can struggle with literal language so are much less likely to

exhibit this in environments where people do not use literal language, stimming can be a reaction to stress so this is less likely to be exhibited in stress free environments and whether an autistic person exhibits difficulties with planning depends upon the need for planning and the nature of the required planning. Fourthly, the neurodiversity movement has heavily emphasised that some environments make symptom expression a positive experience and some make it a negative experience, due to either the environment itself taking a form which fits or does not fit symptom expression, or due to the tolerant or hostile views of individuals in the environment (Chapman 2023, p. 117). All this shows how symptoms can interact with the environment.

Categorical notions of autism and dimensional notions of autism can be influenced by the environment in different ways. For example, someone who meets the criteria for categorical notions of autism is typically more likely to receive early intervention, receive therapy and receive support compared to someone who is dimensionally autistic. Additionally, even if hypothetically everyone who did exhibit any trait of autism to any degree above zero did receive early interventions, therapy and support, the effect of these would likely be very different given the lower number of traits being exhibited and them being exhibited to a lesser degree. In relation to symptoms being a reaction to the environment and neurodiversity, society is typically set up for people who exhibit a low level and a low number of traits associated with psychiatric symptoms. As such, the environment and wider society are more likely to have a causal impact on categorical notions of autism than dimensional notions of autism. Finally, people who exhibit more traits of autism or are diagnosed under categorical notions are more likely to end up in less common environments, such as being the outsider in a school class room, attending special education, and being removed from an education system by leaving school early or being home-schooled. Similarly, they are more likely to be isolated and unemployed in adulthood. Additionally, someone who is categorically autistic is much more likely to exhibit more symptoms of autism compared to someone who is dimensionally autistic. As such, not only are they more likely to be in different environments or in environments which do not suit them, people who are categorically autistic will have more symptoms for the environment to interact with. For these reasons,

the constituted inductive and causal probabilities between the environment and categorical autism are different to the constituted inductive and causal probabilities between the environment and dimensional autism.

5.5.5 Causal Interaction from Stance Towards Symptoms

An individual can take various stances towards their diagnosis. They might embrace it or suppress it. By embracing it they might aim to not prevent themselves manifesting the symptoms of the diagnosis. For example, autistic people sometimes stim (short for stimulating) whereby they make repetitive hand movements such as hand flapping, because it feels pleasant or reduces anxiety (Doan and Fenton 2013, p. 61). This behaviour can look strange to an outside observer but the autistic individual might not seek to prevent into because it harms no one else and benefits themselves. In contrast, a diagnosed individual might seek to repress manifestations of symptoms of their diagnosis. They might aim to mask whereby they aim to pass as normal (Hull et al. 2017, p. 2525). This might be simply because they find it more practically beneficial to do so even though they have no negative views of their diagnosis or because they actively dislike the symptoms themselves above and beyond any negative reaction others have to them. Also, an individual might take different stances towards particular symptoms they exhibit, their stance may change over time and might change depending upon who they are around. All this shows that the stance someone has to their diagnosis can influence which symptoms manifest and how they manifest.

Under a categorical approach, people are given a psychiatric diagnosis of autism. In contrast, on a dimensional approach, most people who exhibit attributes of autism do not meet the DSM criteria. Those who have the categorical diagnosis of autism can take a stance towards their diagnosis whereas most people who are dimensionally autistic cannot. As such, the percent of individuals who reflect about the way they are autistic will be much higher on categorical notions of autism than dimensional notions of autism. Additionally, imagine if everyone was actually given a dimensional diagnosis of autism whereby everyone was told they

exhibit autism to a certain degree or exhibit traits of autism to a particular degree. It seems highly plausible that many people who exhibit low or moderate degrees of a smaller number of traits of autism will hold a different stance towards this compared to someone who exhibits a sufficient number of traits of autism to a sufficiently high degree to meet current categorical diagnostic criteria. As such, the constituted inductive and causal probabilities between the stance people take towards their diagnosis and categorical autism are different to the constituted inductive and causal probabilities between the stance taken and dimensional autism.

5.5.6 RDoC and HiTOP

I have argued that DSM dimensional diagnoses like autism, ADHD, schizophrenia, major depressive disorder, schizoid personality disorder all produce one set of knowledge and their categorical counterparts would produce a different set of knowledge. On this basis, if we decided to make DSM psychiatric diagnoses dimensional then we would have good reason to still retain categorical diagnoses.

However, there are more radical critiques of the DSM like RDoC and HiTOP which do not base their research on DSM diagnoses. I now consider how categories might relate to these. As previously described, RDoC and HiTOP wish to move to a dimensional system. However, it is not clear what diagnoses RDoC and HiTOP will produce. In relation to RDoC, whilst it was never intended as a diagnostic manual a recent article by the current director of RDoC outlines what clinical diagnoses based upon the RDoC framework might look like. It mentions "the use of computational techniques to identify data-driven phenotypes not dependent upon traditional diagnoses" (Cuthbert 2020, p. 4). It provides an example of "a recent study [which] employed a wide variety of measures to analyze the heterogeneity in a large sample of patients diagnosed with schizophrenia, schizoaffective disorder, or psychotic bipolar disorder; a multiple step analysis revealed three transdiagnostic clusters ('biotypes') of patients" (Cuthbert 2020, p. 4). It looks like the RDoC project thinks that clinical diagnoses will consist of phenotypes (i.e. sets of observed characteristics) or biotypes which will be more homogenous in

symptoms covered and underlying causes, and may have biomarkers that would assist diagnosis. In relation to HiTOP, it is primarily used in academic research but it has also been used in clinical practise (see Balling et al. 2023). HiTOP involves rating people upon a variety of dimensions such as "Internalising", "Externalising", "Fear", "Distress", "Antagonism" and "Disinhibition". These are linked to syndromes, that is, groups of symptoms. Interestingly, in earlier HiTOP publications these syndromes actually are current DSM diagnoses. HiTOP shows that these DSM diagnoses can be linked to various dimensions but those dimensions are transdiagnostic, that being, they do not line up with the DSM diagnoses. Given the general lack of statistical evidence in favour of syndromic entities such as "major depression" and "generalized anxiety", it is unclear whether HiTOP will retain a syndrome level in future iterations and if it is retained, it is unclear what will be included in it (Awais Aftab, personal communication (a member of the HiTOP consortium)). In relation to autism, the diagnosis I have mainly focused upon, a recent study aimed to statistically analyse the symptoms described in the vast majority of DSM diagnoses (Forbes et al. 2024). The statistical analysis did not find an autism grouping. Rather, it found two domains, the first one composed of 'Social Communication Difficulties', 'Altered Sensation and Attentional Control' and 'Ritualized Behaviour' constructs, and the second one composed of 'Neurocognitive Impairment', 'Difficulties with Organization' and 'Forgetfulness' constructs. People currently diagnosed with autism, ADHD and OCD would broadly instead primarily fit in the first domain but could also fit within the second domain. Perhaps being rated upon those constructs is what an HiTOP clinical diagnosis would consist of for someone who is currently diagnosed as autistic. Alternatively, perhaps different constructs will emerge given further study or perhaps further subfactors will be added to those constructs.

Whilst there is uncertainty as to what clinical diagnoses resulting from RDoC and HiTOP will be, I now suggest that there is likely value in adding categorical diagnoses to the sort of dimensional diagnoses that both projects appear to desire. Firstly, are these dimensional diagnoses going to be polythetic, that being, is there more than one way to get onto the dimension? A dimension might cover a single symptom and rate it as present to a degree or cover multiple symptoms, all of which must be

present, and rate them to a degree. This would not be polythetic. Alternatively, whilst less heterogeneous than DSM diagnoses, these dimensional diagnoses might rate the presence of multiple symptoms but not all symptoms are always present, so these dimensional diagnoses would be polythetic whereby there are multiple ways to get onto the dimension. If they are polythetic then adding categorical divides may produce more specific inductions and probabilities.

Secondly, whilst more homogenous than DSM diagnoses, will these dimensions extend into the general population? Whilst we can always score people dimensionally, will these dimensions produce clusters with gaps between them or will they be distributed quite evenly across the population. For example, a dimensional measure might find people scoring literally every figure between 0 and 100 with a quite even distribution across those figures. This situation resembles my previous discussion of the autism spectrum and I outlined the role categories can play here. In contrast, a dimensional measure might find virtually everyone scores between 100 and 70, or score 0, whilst almost no one scores between 1 and 69. In this situation there might still be value in adding categories. Whilst such a dimensional diagnosis where everyone scores highly or 0 would cover a much more homogenous population than DSM diagnoses, the inductive and causal probabilities might differ between those who score 90 and those who score 75 and this could be captured by adding categories to the dimension.

Thirdly, even if RDoC and HiTOP produce homogenous dimensions, how might people on those dimensions vary in ways not covered by those dimensions? If a dimension covers depressive phenomena, people on that dimension may vary in ways unrelated to depression. This seems a realistic possibility given that RDoC and HiTOP aim to study transdiagnostic dimensions whereby the same dimensions can be present despite manifesting in quite different symptom profiles. Adding categorical divides may produce useful subdivisions which cover different inductive and causal probabilities.

Fourthly, I outlined in Chap. 4 how symptoms are themselves idealisations and abstractions which are distant from behaviour. In this sense symptoms are themselves heterogeneous whereby two people can exhibit the same symptom yet exhibit significant difference in behaviour covered

by that symptom. As such, two people might score the same figure on a dimensional measure which covered the exact same symptom yet still exhibit differences in behaviour. These differences might result in different inductive and causal probabilities which might be brought out by adding a category to that dimensional diagnosis.

5.6 Systematicity

The second step to formulating scientific knowledge is to employ principles to determine what types of things there are. Applying constitutive principles to the world only tells us what form the world must take but does not establish what types of things there are which take that form. Deciding to employ categorical constitutive concepts means that psychiatric diagnoses must take a constitutive form but this does not then tell us which psychiatric diagnoses there are. Employing a constitutive form does not determine if we should employ, say, autism and schizoid personality disorder or some alternative diagnoses. We need to employ principles to a constituted form to decide which type of psychiatric diagnoses there are.

Kantians often talk of science as being the product of unifying according to principles (Cassirer, 1953a [1910], p. 307; Kitcher 1989, p. 499; Natorp 2004 [1921], p. 386; Rickert 1986 [1902], p. 110). The basic idea is that we take the chaos and flux of sensation and rearrange it according to principles. Doing this removes the chaos and flux, leaving us with an ordered image of the world which is distant from what we see in the world around us. We take seemingly disparate and unconnected elements and unify them together by positing theoretical entities. In relation to psychiatry, we see one person exhibiting one set of symptoms, a second person and a third person exhibiting different overlapping sets of symptoms, and unify these together by positing the psychiatric diagnosis of autism. We take different people and abstract those differences by unifying them under the thought object of autism. The legitimacy of unifying these different people under autism depends upon applying principles to the data. These principles are what convey scientific legitimacy to the

thought objects (assuming that the data we apply the principles to is reliable) because the principles themselves are based in reason.

I will now consider how we can construct psychiatric diagnoses by employing the principles Kant employs, the principles of homogeneity and specificity. The principle of homogeneity produces generalised classes that are quite distinct from particular cases. Greater emphasising of homogeneity results in a psychiatric diagnosis which covers a larger number of symptoms, or symptoms that cluster together with a lower level of statistical frequency. That is, a larger number of combinations of symptoms could result in this diagnosis compared to emphasising the principles of specificity to formulate a psychiatric diagnosis. To highlight this with an example, consider the diagnoses of autism and schizoid personality disorder. Currently, both diagnoses have different diagnostic criteria. However, both diagnoses share some symptoms, such as social withdrawal, unusual social communication and elements of repetitive and restricted interest. Psychiatrists sometimes find differentiating between each diagnosis difficult (APA, [DSM-5] 2013, p. 655). If we emphasise the principle of homogeneity then we could merge both diagnoses together. This combined diagnoses would have the diagnostic criteria of those shared symptoms and potentially some of the symptoms that are specific to autism and some that are specific to schizoid personality disorder. Some diagnosed individuals have the shared symptoms and some currently associated with autism, whereas other diagnosed individuals would have the shared symptoms and those currently associated with schizoid personality disorder. This diagnosis would be more abstract than both autism and schizoid personality disorder are, in that it will typically cover a greater number of symptoms above what particular individuals exhibit whereas the diagnoses of autism and schizoid personality disorder are less abstract because they cover a smaller number of symptoms above what particular individuals with one of those diagnoses has. This combined diagnosis is a greater transformation of what we observe compared to autism or schizoid personality disorder. It would be a thought entity which is more distant from what we observe, compared to autism and

schizoid personality disorder, much like how a parabola is more distant from observation than an ellipse is.[9]

Psychiatric diagnoses can also be formulated on the principle of specificity. This is where a diagnosis is narrower, allowing a smaller number of combinations of symptoms to meet the diagnostic criteria. This can also be highlighted by the diagnosis of autism. This diagnosis is polythetic, allowing a diagnosis to be given if an individual exhibits one of a number of combinations of symptoms. By emphasising the principle of specificity we could produce more narrowly defined psychiatric diagnoses whereby a smaller number of combinations of symptoms would result in the diagnosis. For example, we could take the modern diagnosis of autism and produce a new diagnosis that covered a smaller number of symptoms, such as no longer covering individuals who have sensory issues. This means that the alternative diagnosis is now closer to people who are diagnosed with it compared to autism since there is more resemblance between the diagnostic criteria and people who are diagnosed. To show this consider two individuals who qualify for a diagnosis of autism but who do not have every symptom listed on the diagnostic criteria. Imagine one just lacks the symptom of stereotyped movements whilst the other individual lacks the symptoms of stereotyped movements and the symptom of sensory issues. Both these individuals would be eligible for the current diagnosis of autism, whereas only one would be eligible for the altered diagnosis. This reduces the gap between the symptoms covered by the diagnosis and the symptoms particular people with that diagnosis actually exhibit. This means the diagnosis provides a more precise description of individuals who receive the diagnosis. These approaches will make psychiatric diagnoses more specific, narrowing the gap between the psychiatric diagnosis and diagnosed individuals.

Kant also describes the principle of affinity whereby the relationship between different objects is specified to create a system of science. This is

[9] It is possible, arguably likely, that this new diagnosis would actually fit some specific people better than autism or schizoid personality disorder does. However, when I claim that such a diagnosis is more distant from what we observe I mean when we collectively observe many individuals, such as collating the evidence from many different experiments and studies. As such, we can consider this diagnosis which joins together autism and schizoid personality disorder to be more distant from what we observe, when measured by many observations of people, compared to the separate diagnoses of autism or schizoid personality disorder.

where thought objects are placed in a hierarchical relationship whereby thought objects produced on the principle of specificity are subsumed under those produced on the principle of homogeneity. For example, the more homogenous parabola is at the top of the hierarchy, with the more specific ellipse placed under it, and the even more specific circle placed under that. Similarly, a psychiatric diagnosis which combined autism and schizophrenia would be at the top of the hierarchy, and under it would be subsumed the separate diagnoses of autism and schizophrenia, and under each of them would be subsumed more specific versions of each diagnosis such as subtypes, and these subtypes would then subsume specific people under them.

There is an obvious tension between the principles of homogeneity and specificity. Homogeneity lends towards more abstract groupings whereas specificity lends towards less abstract groupings. Kantians have slightly different views on how these two approaches should be combined. Kant criticises those who focus excessively on homogeneity without considering how things differ from one another, and he criticises those who focus upon specificity without focusing on how things are similar (1998, A688/B696).[10] This suggests that Kant believes that some sort of balance is needed between these two principles but otherwise details are scarce.

More detail on how to employ conflicting principles is outlined in Philip Kitcher's neo-Kantian approach.[11] He argues that science should

[10] Kant also contrasts what he calls determinate judgements, whereby we have no control over how we apply those principles of reason, and reflective judgements, whereby we do have control over how we apply those principles. He suggests in some texts that all science depends on reflective, and not determinate, judgements, though in other texts he seems to suggest that physics is actually based on determinate judgements.

[11] From around the 1960s there have been three major ways in which to interpret Kant's philosophy of science (see Breitenbach 2018 and McNulty 2015 for discussion). The key issue is how Kant can retain a notion of objectivity and necessity (which then delivers truth) despite believing that we cannot access things in themselves. One branch could be called systematisers. Leading advocates are Philip Kitcher, Buchdahl, Brittain and Allison. On this account, scientific truths depend upon how we build a system of science, and therefore depend upon the principles we use to build that system. I adopt and endorse this systematiser approach. It is briefly worth considering how this view relates to the other two major interpretations. There is the derivation account primarily espoused by Micheal Friedman where the principles of reason play a limited role. Rather, it is the categories, which stem from the faculty of the understanding, which provide the basis for truth. However, it is recognised that this view only works in relation to physics and more resources are

cover as much phenomena as possible whilst describing that phenomena in as much detail as possible but doing so whilst employing as few theories as possible. This would mean that more scientific knowledge is covered but there is still a level of simplicity because as few theories as possible are being employed (Kitcher 1989, p. 432; Kitcher 1993, p. 172). As Kitcher points out, there is a tension between, on one hand, covering as much phenomena in as much detail as possible and employing as few theories as possible (Kitcher 1999, p. 343).

The best set of psychiatric diagnoses would balance these two aims by covering as many symptoms as possible whilst employing as few psychiatric diagnoses as possible. To do this, each diagnosis would have to cover a greater number of symptoms, meaning that more symptoms are covered by the system. However, the ideal set of psychiatric diagnoses must also give detailed descriptions, meaning that each psychiatric diagnosis should be as close as possible to describing actual people. This produces a tension between employing more psychiatric diagnoses, to reduce the gap between each psychiatric diagnosis and actual people, and employing less psychiatric diagnoses, to respect a principle of parsimony.

I now consider what constitutes a good DSM diagnosis on this approach and later consider how it relates to RDoC and HiTOP. Particular psychiatric diagnoses vary in how much they contribute to a system which covers as much behaviour as possible whilst employing as few psychiatric diagnoses as possible. Many symptoms are covered by the diagnosis of autism. Mentioned by DSM-5 are social-emotional reciprocity, low eye contact, deficient body language, abnormal facial expressions, lack of shared imaginative play, stereotyped movements, idiosyncratic phrases, adherence to routines, excessively circumscribed interests, disliking unexpected changes, rigid thinking, extremely variable intelligence levels, sensory issues, clumsiness and difficulty with planning (APA 2013, p50–55). Other symptoms which are not part of the diagnostic criteria of

required to apply this interpretation to sciences other than physics. Thirdly, more recently there has developed the necessitation account view. This is where there is necessity in the world via the causal powers of things in themselves but those causal powers, being produced by things in themselves, are inaccessible to us. This means that we still rely upon principles of reason to formulate a science which is truthful to us. Though most emphasised on the systematiser view, on all three accounts importance is placed on the principles of reason and there is an outstanding question about how to employ those principles.

autism but are associated with autism include obsessive interests, peculiar visual perceptions and good memory. In contrast, the diagnosis of social anxiety disorder only covers two main symptoms, namely being anxious and fearing social situations. Autism covers many more symptoms than social anxiety disorder does. This means autism makes a greater contribution, compared to social anxiety disorder, to building a system which covers as many symptoms as possible. Additionally, psychiatric diagnoses vary in how closely they resemble particular individuals. For example, the diagnosis of major depressive disorder can be diagnosed provided five symptoms of the nine of the diagnostic criteria are present in an individual (though one of those must be depressed mood or loss of interest or pleasure (APA 2013, p. 160)). This means that there is significant distance between the diagnosis of depression and people who receive the diagnosis. Two people can receive the diagnosis of major depressive disorder despite only having one symptom in common (see Horwitz 2014, p. 219, also p.222). By contrast, a diagnosis of autism requires multiple symptoms from two different categories, making diagnostic criteria for autism much more stringent than major depressive disorder. An individual with ten random symptoms of autism will not meet the diagnostic criteria of autism unless some of those symptoms are particular ones listed in the social-communication category and some others are particular ones listed in the repetitive and restrictive behaviours category. Both diagnoses differ in how large the gap is between the symptoms described by the diagnosis and actual people. More symptoms, and more specific symptoms, of the diagnostic criteria of autism are needed for a diagnosis compared to the symptoms required for a diagnosis of major depressive disorder, meaning the gap between the diagnostic criteria and the people with the diagnosis is closer for autism. This means that autism both covers a large number of symptoms and groups them together in a manner which reduces the gap between the diagnosis and the diagnosed individual. Autism thus makes a helpful contribution to balancing those aims. Since we judge thought entities based upon how well they contribute to forming a system of thought entities, we should consider autism to have a significant level of epistemic strength (we also need to consider causation, which I do in the next chapter). As such, this provides a means of

epistemologically assessing psychiatric diagnoses without appealing to correspondence with the mind-independent world.

At the same time, I think it would be better if nuanced subtypes of autism were employed. From DSM-III-TR (1987) to DSM-IV-TR (2000) there effectively were subtypes to autism which, though in different forms and names across different editions, conveyed a distinction between high functioning and low functioning autism. This was then subsumed into a single diagnosis of Autism Spectrum Disorder (ASD) in DSM 5 (2013). Since then, some psychiatrists and psychologists have argued that ASD is too heterogeneous and amorphous and that subtypes should return. Many autistic advocates on social media are heavily against subtyping (whether this is true of autistic individuals who avoid social media is something I cannot judge). They typically argue that the distinction between high functioning and low functioning lacks nuance because people can function differently in different areas and functioning can be influenced by the environment. Also, they argue such subtypes are unethical because they stereotype people and because they are based upon a comparison with normally functioning people. Autistic advocates typically reject normality as a good standard for judging people. I completely agree that high functioning vs low functioning are bad subtypes. However, much more nuanced subtypes which might convey a nuanced picture of how autism can manifest might be possible (prior to the 1980s there were multiple subtypes employed which were not simply based upon high vs low functioning). I have argued for subtyping elsewhere on various grounds (Fellowes 2017, p. 57; Fellowes 2020, p. 45; Fellowes 2021, p. 20; Fellowes 2024, p. 18); here I shall only consider the issues from the neo-Kantian position I develop.

I have argued that which psychiatric diagnoses we employ partly depends upon how we employ various principles. I then argued that we should find a balance between these. We should employ a smaller number of concepts but we also should desire concepts to accurately describe what we study which then typically requires employing a larger number of concepts. I have specifically employed the example of autism to argue that it is a good psychiatric diagnosis because it makes a good contribution to a system of psychiatric diagnoses which covers lots of relevant phenomena whilst limiting the number of diagnoses employed. As such,

I think that currently employed notions of autism are scientifically legitimate. At the same time, there is no denying that the concept of autism is currently very heterogeneous. The level of simplicity that autism adds to the diagnostic system is very high but this comes at the cost of making autism extremely distant from how symptoms manifest in actual people. An alternative would be to subtype autism, which adds psychiatric diagnoses, but those psychiatric diagnoses would be less heterogeneous, better accommodating how symptoms manifest in actual people. If we accept that we should aim to balance the principles neo-Kantians endorse then I think that we should aim to increase accuracy at the cost of simplicity by formulating a set of nuanced subtypes for autism. The diagnosis of autism is scientific but I think adding a set of nuanced subtypes not based upon functioning level could be even more scientific by better balancing those principles.

Having outlined DSM approaches I now consider the more radical approach of RDoC and HiTOP. Having analysed the domain of psychiatry dimensionally the hope is that more homogenous groupings would be found. The general approach of RDoC and HiTOP of employing factor analysis to delineate the smallest number of groupings to account for the data heavily fits a neo-Kantian notion of science whereby data is brought under principles of reason to build a structure of science. Relationships between data is conveyed by formulating hypothetical entities through employing statistical techniques, producing a system of science that best fits the data whilst employing the smallest number of entities to account for variance in the data. Following Kitcher's account, we should construct a set of hypothetical entities which covers as much phenomena as possible whilst positing as few entities as possible. Exactly what would constitute an ideal set of entities will depend upon what data RDoC and HiTOP produce, and value decisions when statistically analysing that data, but I believe their general approach is a good one. As such, neo-Kantianism provides a justification for the general approach of RDoC and HiTOP of constructing a system of hypothetical entities through statistical analysis.

My question here is whether this still leaves room for putting categories upon those dimensional constructs. Potentially, there might be no need to employ categories if RDoC or HiTOP produced psychiatric

dimensional diagnoses which were sufficiently homogenous in symptoms covered and causes covered. Perhaps placing categorical divides upon RDoC or HiTOP produce diagnoses would not result in significantly different inductive knowledge or causal probabilities. This all depends upon what these diagnoses are like but I outlined four reasons earlier to believe categories could still play a role here. It could be responded that any role for categories in producing knowledge simply means that any dimensional diagnoses produced by RDoC and HiTOP are still too heterogeneous and that producing even more homogenous dimensional diagnoses would eliminate all benefit of placing categorical divides upon those dimensional diagnoses. Whilst this in itself might be true, this would go against a neo-Kantian standard of employing as few entities as possible when formulating a system of science. In this regard, it would be preferable to not have many very fine-grained dimensions if a smaller number of more homogenous dimensions could also account for the data. That smaller number of homogenous dimensions then gives room to put categorical divides upon those dimensions to produce different inductive knowledge and causal probabilities. In this regard whether to employ categorical or dimensional diagnoses, or a particular combination of the two, partly depends upon principles we employ when systematising. We might do a factor or cluster analytical study of a particular domain and find the smallest number of hypothetical entities that best account for variance within the domain. These hypothetical entities might be categorical or dimensional but we can always add further categorical or dimensional demarcations to them to constitute new inductive and causal probabilities. This might then increase the accuracy of the system but in doing so we need to consider the constraint of employing as few entities as possible. As such, whether to employ categorical and dimensional diagnoses and in what combination partly depends upon the constraints of systematicity.

I have shown that categorical psychiatric diagnoses fit the neo-Kantian notion of scientific objects. They should be seen as thought objects that result from a series of transformations. The sensations we receive from the world are transformed into sensible objects and then we use principles to transform sensible objects into thought objects. These thought objects do not reflect some set of mind independent objects because they have been

produced by those principles. They are also very distant from actual people because psychiatric diagnoses are the product of multiple transformations. My position contrasts with a realist position since diagnoses are something we create rather than being based upon describing mind-independent entities. Additionally, my position contrasts with sceptical positions because I have supplied standards for judging psychiatric diagnoses and suggest those standards are ultimately rooted in reason. We need make a choice over how we apply those standards but they also provide a constraint that allows us to demarcate between better and worse psychiatric diagnoses. I will now respond to concerns over psychiatric diagnoses and see how concerning they are once psychiatric diagnoses are understood as neo-Kantian thought objects.

5.7 Concerns over Psychiatric Diagnoses

Neo-Kantian understanding of psychiatric diagnoses mitigates many of the concerns about DSM psychiatric diagnoses. I shall respond to six concerns here, whilst leaving the concern that psychiatric diagnoses lack homogeneous causes to Chap. 6 (since this is the main concern about psychiatric diagnoses and since I need draw upon additional resources, I shall devote Chap. 6 to responding to this concern).

Firstly, there is the concern that psychiatric diagnoses cover heterogeneous groupings of symptoms (Kinderman et al. 2013, p. 3; Johnstone 2018, p. 39; Vanheule 2017, p. 85). That is, two individuals with the same diagnosis can have quite different symptoms. This suggests that both those psychiatric diagnoses are not describing the same thing. However, I have described how neo-Kantians believe that scientific objects are the result of a series of transformations. The sensation we receive from the world is transformed into sensible objects and these are then transformed into thought objects. Both these transformations involve abstraction. It involves moving from the specific to the general, placing particular individuals into general classes which removes variability. Thus particular instances of the generalised class will have characteristics that the generalised class will not. As such, it is in principle acceptable that psychiatric diagnoses cover heterogeneous groupings of symptoms.

The psychiatric diagnosis is a generalised class which covers particulars which vary from that generalised class. However, there is a limit on the level of heterogeneity which is acceptable. I have argued that psychiatric diagnoses should form a system which covers as many symptoms as possible, whilst describing that phenomena in as much detail as possible, whilst employing as few psychiatric diagnoses as possible. Such a system would reduce the heterogeneity of particular diagnoses because they would describe symptoms in as much detail as possible. That is, there should be a small gap between the symptoms covered by the psychiatric diagnosis and the symptoms exhibited by people with that diagnosis. This gap can be smaller when there are less combinations of symptoms which can meet the diagnostic criteria of the psychiatric diagnosis. That gap can be reduced to zero if each possible combination of symptoms received its own diagnosis. Doing this would radically increase the number of psychiatric diagnoses that need to be employed, conflicting with the goal of employing as few psychiatric diagnoses as possible. Employing as few psychiatric diagnoses as possible means that each psychiatric diagnosis will have to cover multiple combinations of symptoms rather than just one combination. As such, it is preferable to have psychiatric diagnoses which are polythetic, that there are multiple combinations of symptoms which could result in the diagnosis, if the alternative would require radically increasing the number of psychiatric diagnoses employed.

Secondly, there is the concern over co-morbidity (Hyman 2010, p. 167; Kendell and Jablensky 2003, p. 7; Kirk, Cohen & Gomery, 2015 p.69). This is where an individual receives more than one psychiatric diagnosis. If psychiatric diagnoses are natural kinds then people can simultaneously be two natural kinds. Whether this is considered problematic depends upon how natural kinds are understood, but it certainly goes against any notion that natural kinds are naturally demarcated entities which are intrinsically separate from one another. For example, something cannot simultaneously be a cat and a dog. However, this concern loses its force when we stop thinking that psychiatric diagnoses should reflect mind-independent entities and instead see them as thought entities. A thought entity is generalised, abstracting away from particular instances. It is not meant to describe particular instances with one to one accuracy. Rather, it is a generalised class which can cover much variety

and also will not cover aspects of particular instances. This leaves open the possibility and the legitimacy of a second thought entity simultaneously being applicable to that individual. The second thought entity could cover aspects of them which is not covered by the first, and it is legitimate that some aspects are covered by both given that thought entities are abstracted away from what we experience and are distant from the specifics of what we experience. If psychiatric diagnoses were formulated to avoid co-morbidity then many more diagnoses would be required. It would require that psychiatric diagnoses cover more specific combinations of symptoms so that one, rather than many, diagnosis accounts for all the symptoms of particular individuals. Doing this would increase the number of diagnoses being employed. Alternatively, co-morbidity can be reduced by lowering the number of diagnoses employed but making those psychiatric diagnoses cover a greater number of symptoms. This would mean that there is more chance that the symptoms of a particular individual are covered by one diagnosis. This would, however, reduce the fit between the diagnosis and the symptoms exhibited by people with that diagnosis.

Thirdly, there is the problem of subthreshold individuals (Jablensky 2016, p. 27; Okasha 2009, p. 130). Someone is subthreshold when they exhibit a significant number of symptoms on a diagnostic criteria but not enough to actually qualify for the diagnosis. For instance, someone who exhibits only four of the symptoms of depression when five symptoms are required for a diagnosis is subthreshold. It seems strange to consider individuals with five symptoms as a natural kind, and thus a legitimate target of scientific investigation, whereas those with only four do not form a natural kind, and would thus be an illegitimate object for scientific research. However, a neo-Kantian approach gives good reason to consider subthreshold individuals as worthy of scientific investigation in their own right. Kant describes the principle of homogeneity and specificity and he describes how thought entities formulated on the principle of specificity can be more specific versions of thought entities formulated on the principle of homogeneity. We can also see subthreshold individuals as more specific versions, with additional detail, of thought entities which have been founded on the principle of homogeneity.

Fourthly, there is the concern that psychiatric diagnoses fail to demarcate the degree to which individuals meet or fail to meet the diagnostic

criteria (Helzer et al. 2006, p. 1673; Hudziak et al. 2007, p. 21). An individual who has all the symptoms of a diagnostic criteria is given the same diagnosis as someone who only just meets the diagnostic criteria. Similarly, an individual who fails to meet the diagnostic criteria by missing one symptom and an individual who has no symptoms of the diagnostic criteria are not demarcated between by being registered as not having the diagnosis (if subthreshold diagnoses are being employed then this point is applicable when comparing individuals who have no or very few symptoms of the diagnostic criteria and those who fall into the subthreshold diagnosis). This would be problematic if psychiatric diagnoses were understood as mind-independent entities which were naturally similar. On that understanding, to know that someone is an instance of a naturally forming entity means that you entirely or largely know what they are like, but this is not true of psychiatric diagnoses. When understood as neo-Kantian thought entities, psychiatric diagnoses should not be expected to give a one to one description of diagnosed individuals. Rather, variability is to be expected which then gives reason to investigate how instances of the diagnosis have differences from the thought entity that they are instances of.

Fifthly, there is the concern that psychiatric diagnoses leave out important information about people (Johnstone 2014, p. 275; Vanheule 2017, p. 275). They primarily describe which symptoms a diagnosed individual may have but do not cover other aspects of individuals such as their current life situations, their life goals, their life history and many aspects of their personality. However, psychiatric diagnoses should be understood as abstract thought entities which do not cover all the specifics of individual instances. Psychiatric diagnoses which have been established on the principle of homogeneity abstract away details, focusing upon a smaller number of characteristics which are distant from the characteristics of individuals which it covers. There is consequently good reason to investigate many other aspects of the individual. They should be considered scientifically legitimate and important, regardless of whether they are covered by the psychiatric diagnosis or not.

Sixthly, there is the concern that psychiatric diagnoses are the product of a decision making process. If psychiatric diagnoses were naturally forming entities then we should discover what they are rather than decide

what they are. They would be out there to be discovered and either we do discover them or we do not discover them (or discover some aspects of them). We might need to decide *if* we have discovered them but we should not be *deciding* what they are. However, there is a decision making process to constructing thought entities. Which entities there are depends upon how we apply the principles that come from us. Whilst it would be possible to argue that there is a maximally rational way to apply those principles, thus entailing one set of correct entities, I do not believe this is actually implementable. Kant criticised those who focus excessively on the principle of specificity whilst disregarding the principle of homogeneity, and similarly criticises those who focus excessively on the principle of homogeneity at the cost of the principle of specificity, but generally gives no indication of how to balance them (1998, A668/B696). Kant likely thought the principle of affinity would create a necessary link which entails one correct set of entities but Kant seems to have understood this as an unattainable regulative goal (what he calls a focus imaginarius) rather than something which actually could be achieved. Some philosophers have argued that Kant sees the principles of homogeneity and principles of specificity as two different modes of thought which produce different types of scientific knowledge. Which type of knowledge is desired seems partly dependent upon human interests (see Grier 1997, p. 4; Kitcher 1999, p. 343; Morrison 1989, p. 162). We should employ both principles but how we weigh them is dependent upon our interests (this seems a plausible reading of Kant 1998, A688/B696). Kitcher seems to suggest that there is a maximally rational way to balance these principles in some of his classic papers (Kitcher 1989, p. 498), however, more recently he has rejected this view (Kitcher 1999, p. 343). The Marburg and Southwestern neo-Kantians typically emphasised rationality as creating necessary relationships between different parts of science, whereby we should seek an ultimately unattainable one correct set of entities by applying rationality to what comes to us from the world. However, I outlined in Chap. 3 how some neo-Kantian accounts of their contemporary science suggested that science had to develop in ways that would necessarily fit together. I suggested this was an untenable stance and that I was not endorsing notions of rationality entailing necessity in science. Finally, much of Massimi's work is very much against the notion that

science will produce one final outcome. We need decide how those factors should be weighed. We cannot establish scientific entities outside of this process. Whilst a statistical analysis of the symptoms covered by the DSM might produce dimensions rather than categories, we need make further decisions about whether to then add further categories or dimensions. Rather than seeing the neo-Kantian approach I have developed as being able to produce the correct psychiatric diagnoses, it instead outlines constraints upon what is a legitimate psychiatric diagnosis. Thus, we need to make decisions over how to weigh those factors to establish psychiatric diagnoses.

5.8 Conclusion

There has been much concern over categorical psychiatric diagnoses. They are often considered to be arbitrary impositions based upon the judgements of psychiatrists rather than secure scientific evidence. On this understanding currently employed psychiatric diagnoses are not mind-independent entities. Therefore psychiatry should instead formulate new diagnoses, ones which reflect mind-independent entities or should abandon categorical psychiatric diagnoses altogether.

In response to these concerns I have drawn upon a neo-Kantian approach to portray psychiatric diagnoses as thought entities. Rather than being naturally forming entities these are constructs which result from a series of transformations. The sensations we receive are transformed into people with symptoms (as outlined in Chap. 4). A further transformation takes place when people with symptoms are transformed into psychiatric diagnoses. This transformation takes place by employing constitutive principles, such as categorical or dimensional principles. A further transformation takes place by employing principles of reason. Doing this results in thought entities, and these are not sensible objects, nor the sensation we receive from the world, nor things in themselves. They are mental creations, entities that only exist within our thoughts. It is something we have constructed through a series of abstractions, resulting in something generalised which is distant from instances of the thought entity.

Many of the concerns over psychiatric diagnoses relate to them being distant from actual individuals, such as the concern that psychiatric diagnoses cover heterogeneous clusters of symptoms, that they do not demarcate between the degree to which someone meets the diagnostic criteria, that they do not cover many aspects of individuals which are not symptoms. By portraying psychiatric diagnoses as thought entities I show that psychiatric diagnoses can be legitimate scientific entities despite being distant from actual individuals. Some currently employed psychiatric diagnoses could be good science despite being polythetic (and therefore heterogeneous).

References

Allison, Henry. 1983. *Kant's transcendental idealism: An interpretation and defense*. New Haven: Yale University Press.

APA. 2013. *Diagnostic and statistical manual of mental disorders*. 5th ed. Washington, DC: American Psychiatric Association.

Balling, Caroline, Susan C. South, Donald R. Lynam, and Douglas B. Samuel. 2023. Clinician perception of the clinical utility of the Hierarchical Taxonomy of Psychopathology (HiTOP) system. *Clinical Psychological Science* 11 (6): 1108–1121. https://doi.org/10.1177/21677026221138818.

Bentall, Richard. 2004. *Madness explained: Psychosis and human nature*. London: Penguin.

Borsboom, Denny, and Angélique O.J. Cramer. 2013. Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology* 9: 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608.

Boyle, Mary. 1990. *Schizophrenia: A scientific delusion?* London: Routledge.

Breitenbach, Angela. 2018. Laws and ideal unity. In *Laws of nature*, ed. Walter Ott and Lydia Patton, 109–121. Oxford: Oxford University Press.

Brittan, Gordon. 1991. Systematicity and objectivity in the third critique. *Southern Journal of Philosophy* 30 (S1): 167–186. https://doi.org/10.1111/j.2041-6962.1992.tb00664.x.

Buchdahl, Gerd. 1967. The relation between 'understanding' and 'reason' in the architectonic of Kant's philosophy. *Proceedings of the Aristotelian Society* 67: 209–226. https://doi.org/10.1093/aristotelian/67.1.209.

———. 1969. *Metaphysics and the philosophy of science; The classical origins: Descartes to Kant*. Oxford: Basil Blackwell.

Butts, Robert E. 1984. *Kant and the double government methodology: Supersensibility and method in Kant's philosophy of science*. Dordrecht: D. Reidel Publishing Company.

Cassirer, Ernst. 1953a. *Substance and function and Einstein's theory of relativity*. Dover Publications.

———. 1953b. *The philosophy of symbolic forms volume one: Language*. Yale University Press.

———. 1955. *The philosophy of symbolic forms volume two: Mythical thought*. Yale University Press.

Chapman, Robert. 2023. *Empire of normality: Neurodiversity and capitalism*. Pluto Press.

Clark, Megan, Erin Louise, Zoe Vinen, Josephine Barbaro, and Cheryl Dissanayake. 2018. School age outcomes of children diagnosed early and later with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 48: 92–102. https://doi.org/10.1007/s10803-017-3279-x.

Cushing, Simon. 2013. Autism: The very idea. In *The philosophy of autism*, ed. Jami Anderson and Simon Cushing, 17–46. Plymouth: Rowman & Littlefield.

Cuthbert, Brunce. 2020. The role of RDoC in future classification of mental disorders. *Dialogues in Clinical Neuroscience* 22 (1): 81–85. https://doi.org/10.31887/DCNS.2020.22.1/bcuthbert.

Dawson, Geraline, Emily J.H. Jones, Kristen Merkle, Venema Kaitlin, Rachel Lowy, Susan Faja, Dana Kamara, Michael Muras, Jessica Greenson, Jamie Winter, Milani Smith, Sally J. Rogers, and Sara J. Webb. 2012. Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child & Adolescent Psychiatry* 51 (11): 1150–1159. https://doi.org/10.1016/j.jaac.2012.08.018.

Doan, M., and A. Fenton. 2013. Embodying autistic cognition: Towards reconceiving certain "autism-related" behavioural atypicalities as functions. In *The philosophy of autism*, ed. J. Anderson and S. Cushing, 47–72. Plymouth: Rowman & Littlefield.

Fellowes, Sam. 2017. Putting the present in the history of autism. *Studies in History and Philosophy of Biological and Biomedical Sciences* 61: 54–58. https://doi.org/10.1016/j.shpsc.2016.11.002.

———. 2020. Additional challenges to fair representation in autistic advocacy. *American Journal of Bioethics* 20 (4): 44–45. https://doi.org/10.1080/15265161.2020.1730514.

———. 2021. Scientific perspectivism and psychiatric diagnoses: Respecting history and constraining relativism. *European Journal of Philosophy of Science* 11 (8). https://doi.org/10.1007/s13194-020-00320-x.

———. 2024. A history of childhood schizophrenia and consequences for autism. *History and Philosophy of the Life Sciences* 46 (29). https://doi.org/10.1007/s40656-024-00627-5.

Forbes, Miriam K., et al. 2024. Reconstructing psychopathology: A data-driven reorganization of the symptoms in DSM-5. *Clinical Psychological Science*. https://doi.org/10.1177/21677026241268345

Friedman, Michael. 1992. Regulative and constitutive. *The Southern Journal of Philosophy* 30: 73–102. https://doi.org/10.1111/j.2041-6962.1992.tb00658.x.

Grier, Michelle. 1997. Kant on the illusion of a systematic unity of knowledge. *History of Philosophy Quarterly* 14 (1): 1–28.

Guyer, Paul. 1990. Reason and reflective judgment: Kant on the significance of systematicity. *Noûs*: 17–43. https://doi.org/10.2307/2215611.

Helzer, John, Helena C. Kraemer, and Robert F. Krueger. 2006. The feasibility and need for dimensional psychiatric diagnoses. *Psychological Medicine* 32: 1671–1680. https://doi.org/10.1017/S003329170600821X.

Horwitz, Allan V. 2002. *Creating mental illness*. University of Chicago Press.

Horwitz, Alan. 2014. The social function of natural kinds: The case of major depression. In *Classifying Psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 209–226. Massachusetts: MIT Press.

Hudziak, James, Thomas M. Achenbach, Robert R. Althoff, and Daniel S. Pine. 2007. A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research* 16 (S1): 16–23. https://doi.org/10.1002/mpr.217.

Hull, Laura, K.V. Petrides, Carrie Allison, Paula Smith, Simon Baron-Cohen, Meng-Chuan Lai, and William Mandy. 2017. "Putting on my best normal": Social camouflaging in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders* 47 (8): 2519–2534. https://doi.org/10.1007/s10803-017-3166-5.

Hyman, Steven E. 2010. The diagnosis of mental disorder: The problem of reification. *Annual Review of Clinical Psychology* 6: 155–179. https://doi.org/10.1146/annurev.clinpsy.3.022806.091532.

Jablensky, Assen. 2008. The nosological entity in psychiatry: A historical illusion or moving target. In *Philosophical issues in Psychiatry II: Nosology*, ed. Kenneth S. Kendler and Josef Parnas, 77–94. Oxford: Oxford University Press.

———. 2016. Psychiatric classifications: Validity and utility. *World Psychiatry* 15 (1): 26–31. https://doi.org/10.1002/wps.20284.

Jankowiak, Tim. 2017. Kantian phenomenalism without Berkeleyan idealism. *Kantian Review* 22 (2): 205–231. https://doi.org/10.1017/S1369415417000024.

Johnstone, Lucy. 2014. Controversies and debates about formulation. In *Formulation in psychology and psychotherapy: Making sense of people's problems*, ed. Lucy Johnstone and Rudi Dallos, 2nd ed., 260–289. Hove: Routledge.

———. 2018. Psychological formulation as an alternative to psychiatric diagnosis. *Journal of Humanistic Psychology* 58 (1): 30–46. https://doi.org/10.1177/00221678177222.

Kant, Immanuel. 1998. *Critique of pure reason*, The Cambridge edition of the works of Immanuel Kant. Ed. and Trans. Paul Guyer and Allen Wood. Cambridge: Cambridge University Press.

———. 2000. *Critique of the power of judgment*, The Cambridge edition of the works of Immanuel Kant. Ed. and Trans. Paul Guyer. New York: Cambridge University Press.

———. 2002. *Theoretical Philosophy After 1781* (Kant, Immanuel, 1724–1804. Works. English 1992). Ed. and Trans. Henry Allison and Peter Heath, Trans. Gary Hatfield and Michael Friedman. Cambridge: Cambridge University Press.

Kendell, Robert, and Assen Jablensky. 2003. Distinguishing between the validity and utility of psychiatric diagnosis. *American Journal of Psychiatry* 35: 139–144. https://doi.org/10.1176/appi.ajp.160.1.4.

Kendlar, Kenneth, Peter Zachar, and Carl Craver. 2011. What kinds of things are psychiatric disorders? *Psychological Medicine* 41: 1143–1150. https://doi.org/10.1017/S0033291710001844.

Kincaid, Harold. 2014. Defensible natural kinds in the study of psychopathology. In *Classifying psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 145–174. Massachusetts: MIT Press.

Kinderman, Peter, John Read, Joanna Moncrieff, and Richard P. Bentall. 2013. Drop the language of disorder. *Evidence-Based Mental Health* 16 (1): 1–3. https://doi.org/10.1136/eb-2012-100987.

Kirk, Stuart, David Cohen, and Tomi Gomory. 2015. DSM-5: The delayed demise of descriptive diagnosis. In *The DSM-5 in perspective: Philosophical reflections on the psychiatric Babel*, 63–81. Springer: Netherlands.

Kitcher, Philip. 1989. Explanatory unification and the causal structure of the world. In *Scientific explanations*, ed. Philiphp Kitcher and Wesley C. Salmon, 410–506. Minnesota: Minneapolis.

———. 1993. *The advancement of science*. New York: Oxford University Press.
———. 1999. Unification as regulative ideal. *Perspectives on Science* 7 (3): 337–348.
Kite, Donna, Judith Gullifer, and Graham A. Tyson. 2013. Views on the diagnostic labels of autism and Asperger's disorder and the proposed changes in the DSM. *Journal of Autism and Developmental Disorders* 43: 1692–1700. https://doi.org/10.1007/s10803-012-1718-2.
Koi, Polaris. 2021. Genetics on the neurodiversity spectrum: Genetic, phenotypic and endophenotypic continua in autism and ASD. *Studies in History and Philosophy of Science* 89: 52–62.
Massimi, Michela. 2014. Natural kinds and naturalised Kantianism. *Noûs* 48 (3): 416–449. https://doi.org/10.1111/j.1468-0068.2012.00862.x.
McLaughlin, Peter. 2014. Transcendental presuppositions and ideas of reason. *Kant-Studien* 105 (4): 554–572. https://doi.org/10.1515/kant-2014-0026.
McNulty, Michael Bennett. 2015. Rehabilitating the regulative use of reason: Kant on empirical and chemical laws. *Studies in History and Philosophy of Science Part A* 54: 1–10. https://doi.org/10.1016/j.shpsa.2015.06.002.
Meer, Rudolf. 2018. Immanuel Kant's theory of objects and its inherent link to natural science. *Open Philosophy* 1: 342–359. https://doi.org/10.1515/opphil-2018-0025.
Morrison, Margret. 1989. Methodological rules in Kant's philosophy of science. *Kant-Studien* 80 (2): 155–172. https://doi.org/10.1515/kant.1989.80.1-4.155.
Natorp, Paul. 2004. *Plato's theory of ideas: An introduction to idealism*. Academia Verlag Sankt Augustin.
Okasha, Ahmed. 2009. Would the use of dimensions instead of categories remove problems related to subthreshold disorders? *European Archives of Psychiatry and Clinical Neuroscience* 259: 129–133. https://doi.org/10.1007/s00406-009-0052-y.
Ostaric, Lara. 2009. Kant's account of nature's systematicity and the unity of theoretical and practical reason. *Inquiry* 52 (2): 155–178. https://doi.org/10.1080/00201740902790235.
Petrolini, Valentina, and Agustin Vicente. 2022. The challenges raised by comorbidity in psychiatric research: The case of autism. *Philosophical Psychology* 35 (8): 1234–1263. https://doi.org/10.1080/09515089.2022.2052829.
Raskin, Jonathan, and Adam M. Lewandowski. 2000. The construction of disorder as human enterprise. In *Constructions of disorder: Meaning making framework for psychotherapy*, ed. Robert A. Neimeyer and Jonathan D. Raskin, 15–40. American Psychological Association.

Read, John. 2004. The invention of schizophrenia. In *Models of madness, psychological, social and biological approaches to schizophrenia*, ed. John Read, Richard Bentall, Loren Mosher, and Jacqui Dillion, 21–34. Hove: Brunner-Routledge.

Reichenbach, Hans. 1965. *The theory of relativity and apriori knowledge*. Berkeley: University of California Press.

Rickert, Heinrich. 1986. *The limits of concept formation in natural science: A logical introduction to the historical sciences*. Cambridge University Press.

Spencer, Lucienne, and Havi Carel. 2021. 'Isn't Everyone a Little OCD?' The epistemic harms of wrongful depathologization. *Philosophy of Medicine* 2 (1): 1–18. https://doi.org/10.5195/pom.2021.19.

Sterling, Lindsey, Geraldine Dawson, Annette Estes, and Jessica Greenson. 2008. Characteristics associated with presence of depressive symptoms in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 38: 1011–1018. https://doi.org/10.1007/s10803-007-0477-y.

Vaihinger, Hans. 1935. *The philosophy of 'As If': A system of the theoretical, practical and religious fictions of mankind*. 2nd ed. Routledge & K. Paul.

van Steensel, Francisca, Susan M. Bögels, and Sean Perrin. 2011. Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis. *Clinical Child and Family Psychology Review* 14: 302–317. https://doi.org/10.1007/s10567-011-0097-0.

Vanheule, Stijn. 2017. *Psychiatric diagnosis revisited: From DSM to clinical case formulation*. Cham: Palgrave Macmillan.

Walsh, Pat, Mayada Elsabbagh, Patrick Bolton, and Ilina Singh. 2011. In search of biomarkers for autism: Scientific, social and ethical challenges. *Nature Reviews Neuroscience* 12: 603–612. https://doi.org/10.1038/nrn3113.

Watson, Jo. 2019. Introduction. In J. Watson. (Ed.) Drop the Disorder: Challenging the Culture of Psychiatric Diagnosis (pp. 1–7) Monmouth: PCCS Books.

Windleband, William. 2015a. Philosophy of Culture and Transcendental Idealism. In The Neo-Kantian Reader, ed. Sebastian Luft, 317–324. London: Routledge.

6

Neo-Kantianism and Causes

6.1 Introduction

Psychiatrists and philosophers typically place great importance on causes when evaluating psychiatric diagnoses. Correctly identifying the causes of a psychiatric diagnosis is taken to convey scientific legitimacy to the diagnosis and it makes the psychiatric diagnosis worthy of belief. As philosopher of psychiatry Dominic Murphy writes, "nosology [psychiatric diagnoses] must be based on causal explanations of what is being classified" (Murphy 2006, p. 5; see also Beebee and Sabbarton-Leary 2010, p. 22; Cooper 2005, p. 53; Meehl 1995, p. 273; Tsou 2016, p. 417). A related approach is where we do not need to identify a cause or small number of causes, rather, there need to only be reasons to believe the psychiatric diagnosis has an underlying cause or small number of causes even though they have not yet been identified. In contrast, a psychiatric diagnosis without an underlying cause or without a small number of underlying causes would be considered to not merit belief. In this case the problem is not that the cause has not yet been identified, rather, it is that the psychiatric diagnosis does not track underlying causes. Specifically, rather than there being a single or a limited number of causes which

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_6

produce the diagnosis, many different causes produce the psychiatric diagnosis. Anyone with the psychiatric diagnosis typically only has a proportion of those causes and those causes are present in people without the psychiatric diagnosis. In varying ways and in varying forms these general ideas have been endorsed by psychiatrists and philosophers of psychiatry. The main motive for this is that causes are seen as being mind-independent. Basing psychiatric diagnoses on such mind-independent causes means that they are based on something real, rather than simply being based upon the views of psychiatrists.

Neo-Kantians also heavily value causes as an epistemic standard for judging science. However, I will argue that, following a neo-Kantian approach, causation in science should not be seen as mind-independent. Rather, scientific causes are constructed by us. This means that we cannot use mind-independent causes as a standard to judge psychiatric diagnoses. Rather, alternative standards need employing when using causes to formulate psychiatric diagnoses. I shall outline the alternative standards which neo-Kantians employ when formulating causes, how a neo-Kantian would use mind-dependent causes as a basis for formulating psychiatric diagnoses and also how psychiatric diagnoses play a role in formulating causes.

6.2 Causes as an Epistemic Standard of Psychiatric Diagnoses

A cause is typically considered something which results in something else. A particular state of affairs is the outcome of a cause or multiple causes, thus that cause or those causes brought it about. Examples of causes in psychiatry could be genes, neurological structures, thoughts, emotions, early upbringing, social environments and pollutants. The presence or absence of these sometimes makes a difference to what occurs. Typically, all these causes have effects upon humans. The genes of an individual can influence their attributes. Their neurology and psychology can influence how they act. Upbringing and social environment also influence people, even if these might act as a cause in a different way to

genes and brain structure. Psychiatry has often held divergent views upon what type of thing can be a cause. For example, some more extreme views during the psychoanalytical era focused solely on early upbringing whilst genes and neurology were the sole focus of more extreme views during the biogenetic era. Middle grounds between various more extreme views are possible and psychiatry might currently be closer than usual to such a middle ground given the recognition of heterogeneity of causes.

Psychiatrists typically put significant epistemic worth on causes. This is typically thought of in terms of validity. "The validity question is how 'good' are these symptoms at representing the hypothesised underlying illness" (Goodyer 2012, p. 335). Issues of validity relate to real entities and real diseases (Pies 2008, p. 49), "actual neural or genetic mechanisms" (Andreasen 1995, p. 162). Validity is typically conceptualised in terms of sensitive and specific causes, that is, everyone with the diagnosis has the cause whilst everyone without the diagnosis lacks the cause. Sensitivity is where the cause is always present in individuals with the psychiatric diagnosis whilst specificity means that the cause is restricted to the psychiatric diagnosis. Additionally, validity typically focuses on biological factors rather than environmental or social factors. These points seem related since it seems possible that a particular biological factor might be present in a very specific population whereas it seems less likely that social relationships are always linked to one and only one specific set of behaviours.

Why then should such sensitive and specific causes be considered epistemologically significant for psychiatric diagnoses? The notion that sensitive and specific causes have epistemic importance seems based in an implicit notion of natural kinds. As outlined in Chap. 2, philosophers often see natural kinds as entities in the world which are internally the same or have significant degrees of internal similarity. This is the notion which psychiatrists discussing validity seem to implicitly have in mind. A psychiatric diagnosis with sensitive and specific causes is taken to be valid because everyone with the diagnosis has the same cause. This thus means such validated psychiatric diagnoses are natural kinds. Natural kinds are sometimes seen as being determined by the mind-independent world (as I outline below, philosophers of psychiatry have developed more nuanced notions of natural kinds). The real, actual causes determine the attributes

of the natural kind. Thus the natural kind and its causes are out there waiting to be found, they are not constructed by humans. This fits a realist picture whereby real causes create real entities and the correct psychiatric diagnoses would be those that describe those real entities. Seemingly following this, validated psychiatric diagnoses are said to be determined by the world rather than by the arbitrary views of psychiatrists. Kendell and Jablensky say this very clearly when describing invalid diagnoses as "merely arbitrary" (Kendell and Jablensky 2003, p. 3).

This approach is problematic given, as outlined in Chap. 2, the vast levels of heterogeneity in psychiatry. Very few DSM diagnoses have been validated on sensitive and specific causes (Cuthbert and Insel 2009, p. 3; Kozak and Cuthbert 2016, p. 287; Sanislow et al. 2010, p. 632). There is thus reason to believe not only that currently employed psychiatric diagnoses are not natural kinds but also that there might be very few such natural kinds in psychiatry for us to discover, at least if natural kinds are understood to have sensitive and specific causes.

As outlined in Chap. 2, validation is also possible without sensitive and specific causes. This seems to be the direction psychiatry is heading in given Hyman's call for alternative approaches (Hyman 2010) and the development of RDoC (Cuthbert and Insel 2009) and HiTOP (Kotov et al. 2017). On non-sensitive and non-specific approaches the general idea is that the causal basis of a diagnosis can be established even though it is not sensitive and specific. By being not sensitive the cause produces the diagnosis even if that cause can produce other diagnoses as well. By being not specific that cause is not required since other causes can produce the diagnosis. Despite this, it is still a fact that the cause produces the diagnosis in a particular instance of the diagnosis even if, in other instances, the diagnosis can be produced without the cause being present and even if the cause can produce other diagnoses as well. Therefore, the cause validates the diagnosis.

However, some of the intuitive reasons why sensitive and specific causes convey epistemic significance to psychiatric diagnoses do not easily fit non-sensitive and non-specific causes. On a typical realist position, the epistemic significance of sensitive and specific causes lies in them being natural kinds formed by the mind-independent world. If a cause is not sensitive or specific to a psychiatric diagnosis then it could be argued

that the psychiatric diagnosis is not part of the mind-independent world. The concern is that the world does not determine what the cause results in since the cause does not clearly link to the specific effect of that psychiatric diagnosis, rather, it is determined by psychiatrists. Also, it could be argued that without a sensitive and specific cause the psychiatric diagnosis cannot be a natural kind. If natural kinds necessarily have members that are internally the same or similar then the concern is that sensitive and specific causes are required for something to be a natural kind. Therefore, psychiatric diagnoses with non-sensitive and non-specific causes are not natural kinds which have been determined by the world, rather, they have been grouped together by psychiatrists. If the absence of sensitivity and specificity means that psychiatric diagnoses are no longer mind-independent natural kinds then it is not intuitively clear why non-sensitive and non-specific causes should have epistemic significance for psychiatric diagnoses. This means we should not understand psychiatric diagnoses as real things that are produced by the world, which then raises constructivist concerns that psychiatric diagnoses are something we have invented. These problems are not necessarily insurmountable but further explication is required. I now turn to how philosophers of psychiatry have addressed these issues.

As outlined in the introduction, philosophers of psychiatry often conceptualise psychiatric diagnoses as natural kinds. Whilst natural kinds have often been associated with essentialism there arose in the 1990s within philosophy of science notions of non-essentialist natural kinds (Boyd 1991; Dupré 1993, p. 5; Hacking 1991; Psillos 1999, p. 289).[1] On this account there are mechanisms in the world which bring about properties. Those properties cluster together, in that one property will typically co-occur with certain other properties and not others. Thus the mechanism causes both the properties and also the manner in which they co-occur. This is a non-essentialist position for two reasons. Firstly, each instance of the mechanism need not be identical. Take for example a particular mechanism of evolution which influences how a species relates

[1] In Chap. 1, I outlined Haslam's four notions of natural kinds. He only calls essentialist natural kinds as 'natural kinds' whereas non-essentialist kinds are given other names. In this section I follow typical convention by calling both essentialist and non-essentialist kinds 'natural kinds'.

to their environment. The mechanism in one particular instance involves a particular species and a particular environment whereas another instance involves a different species and a different environment. The causal details are different in both cases since difference in species and environment means that different interactions take place. Thus each instance of a mechanism is not identical and therefore not essentialist. Secondly, the mechanism produces clusters of co-occurring properties rather than one set of identical properties. This is due to variations in the specific causal process resulting in different outcomes. However, Boyd takes these mechanisms as being homeostatic, that is, they are effectively self-correcting by typically ending up back at an equilibrium. Consequently, they typically, though not always, produce a relatively narrow set of results.

Boyd's homeostatic mechanisms provide potential solutions to how non-sensitive and non-specific causes can provide epistemic legitimacy to psychiatric diagnoses. Firstly, instances of a particular mechanism can vary from one another and yet still produce the same rough range of results. They vary from one another in the causal interactions they cover yet all are instances of the same homeostatic mechanism. Secondly, homoeostatic mechanisms are not necessarily sensitive and specific. The mechanism may produce a set of co-occurring properties but other mechanisms might also be able to produce a similar set of co-occurring properties. Additionally, the mechanism may be present and yet the particular characteristics are not present. The mechanism might not be activated or perhaps it activates in non-typical ways which result in an abnormal output. Homeostatic mechanisms also mean that psychiatric diagnoses can be natural kinds despite lacking sensitive and specific causes. Natural kinds are typically understood as internally the same but homeostatic mechanisms show how clusters of co-occurring properties can arise without sensitive and specific causes. Such a natural kind is not identical internally or externally, rather, it is an entity because significant levels of internal similarity, the homeostatic mechanism, results in significant levels of external similarity, the co-occurring properties.

Philosophers of psychiatry often appeal to Boyd's account of homeostatic cluster kinds when outlining notions of psychiatric diagnoses as natural kinds (Beebee and Sabbarton-Leary 2010, p. 23; Kendlar et al. 2011, p. 1146; Murphy 2014, p. 120; Tsou 2016, p. 411). Tabb writes

that "[p]hilosophers have defended the pursuit of these underlying mechanisms as the right way to validate diagnostic kinds [psychiatric diagnoses], that is, to show that they represent real entities rather than social constructions" (2019, p. 2178). On this account psychiatric diagnoses with homeostatic mechanisms are natural kinds whereas psychiatric diagnoses without homeostatic mechanisms are not natural kinds. The homeostatic mechanism results in multiple symptoms co-occurring together into an entity. The same basic idea is true of philosophers who do not endorse Boyd's account of natural kinds such as Cooper (2007) who endorses Dupre's slightly different account of natural kinds and Murphy (2006) who does not endorse notions of natural kinds but instead focuses on causal mechanisms.[2]

All this shows how psychiatric diagnoses can be natural kinds despite not having sensitive and specific causes. Why, however, should a homeostatic mechanism convey epistemic legitimacy to psychiatric diagnoses? What makes a psychiatric diagnosis any more real or worthy of belief just because it has a homeostatic mechanism and therefore is a natural kind? The general idea seems to be that causal mechanisms are real parts of the objective, mind-independent world. The objective mechanisms in the world determine the objective natural kinds. Some quotes highlight this approach. In their discussion of how psychiatric diagnoses can be natural kinds Beebee and Sabbarton-Leary talk of "real causal mechanisms" and "genuine causal mechanisms" (2010, p. 22). Cooper talks of "real differences in the causal powers of entities" (2007, p. 46). Murphy talks of "causes of mental illness [that] are genuinely out there waiting to be

[2] Dupre's notion of natural kinds is generally understood to vary from Boyd's by seeing less of a link between mechanisms and specific clusters. Boyd sees mechanisms as being linked quite closely to constrained clusters of properties. Dupre sees a link between mechanisms and individual properties but then believes there is more freedom in how clusters of symptoms can be formed into entities. He thinks they are promiscuously pre-grouped. Recall the varying approaches of Haslam and Cooper as described in Chap. 1. Haslam seems to see fuzzy kinds as being determined by the world (even though there is ambiguity over who is a fuzzy kind). Cooper, in contrast, seems to believe that clusters of properties are not pre-grouped into one type of grouping, rather, psychiatrists must choose, given all the ways in which the properties are pre-grouped by the world, which of the pre-groupings they take the entity to be. This seems to follow from Haslam endorsing Boyd whilst Cooper endorses Dupre. Note also that Murphy thinks whether something is or is not a natural kind to be relatively unimportant in his 2006 book whereas in a 2014 article he seems to endorse Boyd's notion of natural kinds.

found" (Murphy 2014, p. 62). Mechanisms and causes are taken as real. Consequently, the properties caused by those causal mechanisms are real. So too are groupings of properties on most accounts. The reality of causal mechanisms conveys epistemic legitimacy to psychiatric diagnoses with causal mechanisms. This could rescue realism. There are real causes in the world and these convey a level of realism to diagnoses, even if those real causes do not have a sensitive and specific relationship with psychiatric diagnoses. This is not to say that currently employed DSM diagnoses are real measured by this standard, merely that this is what form a real categorical psychiatric diagnosis would take.

There is a potential gap in this notion that mechanisms make psychiatric diagnoses real. A concern is the lack of one to one correspondence between mechanisms and the diagnoses. One instance of the psychiatric diagnosis is produced by a set of mechanisms whereas a second instance of the diagnosis might be produced by a different, if overlapping, set of mechanisms. In what sense do the mechanisms join together that psychiatric diagnosis? The key issue here is in what manner we can say the mechanisms group together and make real the psychiatric diagnosis. For example, rather than thinking a set of overlapping mechanisms which need not always be present produce a real particular psychiatric diagnosis we might think there are multiple psychiatric diagnoses based upon the specific combination of mechanisms present in a particular instance. To a neo-Kantian, what joins together the psychiatric diagnosis out of the mechanisms is not reality but our minds. Some philosophers who appeal to natural kinds or causal mechanisms to legitimate psychiatric diagnoses recognise that decision making is required when formulating real psychiatric diagnoses out of mechanisms (Cooper 2007, p. 51, Kendlar et al. 2011, p. 1149). As such, some realists think decision making is compatible with realism whereas a neo-Kantian does not. This is a metaphysical debate between positions which I outlined in Chap. 3. Rather than repeat this, I will now point to a potential deeper problem, namely in what manner mechanisms themselves are real.

6.2.1 Are Mechanisms Real?

In what sense are mechanisms real things out there waiting to be found? Whilst it might seem intuitively obvious that mechanisms exist, this notion can be challenged if we consider the ongoing discussion about the status of mechanisms by philosophers of science. In 2000 Machamer, Darden and Craver published a seminal paper named 'Thinking about mechanisms'. This paper renewed interest in causal mechanisms among philosophers of science. These three philosophers, alongside philosophers holding a similar position, became known as the new mechanists.

Machamer, Darden and Craver see mechanisms as organisations of things which regularly produce outcomes. The definition given is that "[m]echanisms are entities and activities organised such that they are productive of regular changes from start or set-up to finish or termination… Mechanisms are composed of both *entities* (with their properties) and *activities*. Activities are the producers of change. Entities are the things that engage in activities" (2000, p. 3 italics original). When an organisation of things is present then an outcome will occur (or a series of steps will occur which eventually results in an outcome). Consequently, the organisation of those things caused the outcome. The literature on mechanisms is replete with examples which can be drawn upon to characterise mechanisms. For example,

> [t]he retina is a mechanism for transducing light into electrical impulses. The brain's medulla is a mechanism for triggering the gag reflex… The electrosensors that line the globlin shark's snout are part of a mechanism for detecting prey; this mechanism, in turn, is part of a larger mechanism for capturing prey. (Garson 2013, p. 317)

Thus a particular organisation, the electrosensors being in a certain proximity to prey, regularly brings about a particular result, the detection of prey.

Machamer, Darden and Craver see mechanisms as ontologically distinct features of the world. They argue that mechanisms are not reducible down to something more fundamental and that they cannot be replaced by some other non-mechanistic thing (2000, p. 4). Mechanisms are not

simply just entities and not simply just activities, rather, they are the combination of these and thus are ontologically distinct from either one (2000, p. 8). Since they are irreducible to either entities or activities, to posit a mechanism is not merely to gloss over a lack of knowledge about more specific causes. Rather, mechanisms are ontologically distinct features. This would seemingly support notions of real mechanisms which philosophers of psychiatry have appealed to.

Since the publication of Machamer, Darden and Craver's seminal paper there has been a debate between the new mechanists and their critics over the status of mechanisms. Also, Machamer, Darden and Craver have further developed their position. There are multiple challenges to notions of mechanisms as simply things out there waiting to be found and I will now outline some of them. Notions of mechanisms are sometimes used in quite different ways and different potential criticisms are applicable to different usages. The following issues are applicable to the types of mechanisms which are seen as potentially grounding psychiatric diagnoses.

Causal mechanisms are often quite general. Two different things are taken as being instances of the same mechanism even though they have significant differences in the specific causal interactions they cover. To use an example from Boone and Piccinini, a hippocampi is a causal mechanism which allows, among other things, a creature to navigate its environment. Now imagine two specific rats, both with a working hippocampi. Both rats have the same causal mechanism, the hippocampi, yet there are differences in the specific causes present. One rat's brain consists of a specific configuration of neurons whereas the other rat's brain will have a different specific configuration. Additionally, the first rat will have one set of neurons which interact with a second set of neurons whereas the second rat will have a third set of neurons which interact with a fourth set. Each of these sets of neurons will be different. These differences are true given any two particular rats but they become even stronger given that all rats with working hippocampi have differences to any other rat with a working hippocampi. The differences are further emphasised given that humans and whales also have hippocampi and these will involve radically different configurations of neurons compared to any specific rat. The specific configuration of neurons and their specific interactions in a

particular creature is tied to just that specific creature, whereas the causal mechanism of a hippocampi is not dependent upon a specific configuration of neurons or specific interaction between neurons. The causal mechanism of a hippocampi covers multiple instances of specific configurations and interactions of neurons. The basic idea is that causal mechanisms abstract away much of the specific causal detail (Boone and Puccinini 2016, p. 693; Glennan 2002, p. 345; Overton 2011, p. 943).

The key issue is how to move from diverse causal happenings to all those diverse causal happenings being instances of the same mechanism. I outline two challenges. Firstly, causal mechanisms appear to have underdetermined boundaries. A mechanism could be posited which covered a particular set of causal interactions. An alternative mechanism could also be posited which covered a largely similar, though not identical, set of causal interactions. This alternative mechanism might not cover all the causal interactions which the first mechanism covers or might cover additional causal interactions which the first mechanism does not cover (or a combination of these). This makes it unclear which mechanism should be posited (Overton 2011, p. 953). Secondly, we could just say that the specific causal interactions are real but the mechanism is not. For example, the mechanisms by which a globlin shark captures prey is made up of a mechanism by which it detects prey. The detection mechanism is made up of, among other things, electrosensors. Those electrosensors will be made up of other mechanisms, such as cells. Those mechanisms will be made up of atoms and so on. We might posit some or all of these components of the detection mechanism as being real but deny that the detection mechanism is real above and beyond the components that make it up (Baetu 2015, p. 778; Rosenberg 2018, p. 15; Theurer 2013, p. 911). All this raises the question over whether and in what manner we should see mechanisms as real and ontologically distinct from their components.

One of the originators of the new mechanist movement admits that the new mechanists "provide scant basis for saying when two mechanisms are mechanisms of different types and for saying where one mechanism ends and another begins" (Craver 2009, p. 576). Also, Craver outlines how we could characterise a "mechanism very abstractly [but this] potentially glosses over sub-kinds of mechanisms. [In contrast, c]haracterizing the mechanism in maximal detail threatens to make each particular

mechanism a kind unto itself" (Craver 2009, p. 587). We can formulate mechanisms that are very general but quite distant from the specific causal happenings. When doing this, each mechanism can have many instances in the world. Alternatively, we can formulate mechanisms which are extremely close to the specific causal happenings but this means each mechanism loses generality, applying to very few instances in the world. There are many intermediate possible levels of generality between these positions. The key issue is that we seem to have a choice over how to formulate mechanisms. Indeed, Craver writes that, "[t]he world does not come pre-chunked into mechanisms; it takes considerable effort to carve mechanisms out of the busy and buzzing confusion that constitute the causal structure of the world [the specific causal happenings]" (2013, p. 10). Notions that there are not pre-chunked mechanisms in the world certainly suggest that we need more than the intuitive notion that mechanisms are real parts of the world out there waiting to be found.

There is an ongoing debate about the status of mechanisms and a variety of views can be held; for my purposes I have aimed to show that current debates about mechanisms make it plausible to question notions of mind-independent mechanisms. This suggests that the way in which causal mechanisms could provide an epistemological basis for psychiatric diagnoses is philosophically complicated. Not only do we have a choice over what psychiatric diagnoses to formulate out of causal mechanisms, we also might have a choice over what causal mechanisms there are. I am not aware of any work applying neo-Kantianism to the current debate among philosophers of science about the status of mechanisms. The neo-Kantian stance is that mechanisms are formulated by applying something that comes from us. Any mind-independent mechanisms existing independent of our knowledge would be inaccessible to us so the mechanisms we can know of are those which we construct. I will now outline a neo-Kantian notion of causation.

6.3 Neo-Kantians and Causation

Neo-Kantians are typically understood to believe that causes are not mind-independent. They are typically taken as believing that we impose causes on the world. Such a position would clearly be at odds with how most philosophers of psychiatry think of causation. However, exactly what a neo-Kantian believes about causation is complicated. I will here outline a Kantian approach to causation and then later apply it to psychiatry.

Whether Kant believed that there are causes that are independent of us is a complicated and controversial question that need not trouble us. For all of Kant's statements that we impose causation onto the world there are very rare suggestions that things in themselves might cause sensations in us (Kant 1998, A496/B524). The meaning of these claims is of some dispute but such issues are not relevant here because, should such mind-independent causes exist, they are inaccessible to us.[3] To a Kantian, the correct scientific theories are not those we believe to describe any such mind-independent causes since we cannot access such causes. Rather, when we consider causation in assessing scientific theories it is the causation that we impose upon the world which is relevant.

There are multiple steps when Kantians see us constituting the world as having causes. This is similar to the transformation I described in Chap. 5 in relation to scientific entities, that of taking sensation we receive from the external world and turning it into scientific objects. As described in Chap. 5, we move from observable objects to thought objects, the latter being scientific objects. This process also occurs in relation to causation. Firstly, Kantians think that we cannot see or perceive causes in what comes to us from the world prior to conceptualisation. Rather, what comes to us from the world is an unfiltered, unstructured

[3] Kantians can believe that there are mind-independent causes through seeing things in themselves as mind-independent entities which have causal powers. If things in themselves are mind-independent and if they have causal powers then there are mind-independent causes (even if these are inaccessible to us). However, a Kantian can believe that there are mind-independent entities without then believing they have causal powers, leaving causes as mind-dependent. Also, some Kantians believe that things in themselves are actually mind-dependent. Things in themselves might be mental creations or taken as something which we can logically imagine rather than mind-independent entities. In this case they would not involve mind-independent causes.

and unformed mass of sensation. This is inaccessible to us and were we to somehow access it we would not see causes, we would just see a constantly changing manifold of sensation. For Kantians, we only start having conscious awareness of sensation when we impose space and time onto that manifold of sensation. Doing this means that we start to see clumps of sensation as occupying the same area (they are located spatially) and they follow one another (they are located temporally). This is still not causation because we do not see anything causing the clumps of sensation or see one clump cause another. We start seeing causation when we impose objects and causes onto those clumps. Having constituted those sensations as objects (as outlined in Chap. 5), we posit that such objects causally interact, in that when there are multiple objects present they are followed by a particular occurrence. To say that we see objects cause one another is only to say that we see a succession of events but it is us who connect those events together to demarcate causal claims. We see one object in spatio proximity to another object and then temporally see a different state of affairs but it is us who demarcate this as an instance of causation. This means that we constitute the sensation we receive from the external world as causal.

This means that for Kantians causation is mind-dependent. Care, however, is needed to understand what this means. Kantians are not claiming that we simply decide what causes there are. We constitute sensation with causal attributes. We cannot help but do it and we have no control over whether we make such connections. In this regard causes are mind-dependent but are not simply made up in any manner we wish. When we impose causation we see a billiard ball knock into another one, or we might see an apple turn into mush after falling from a tree, or we might see colour change and bubbles when chemicals are mixed together. These connections are imposed but scientists cannot make up the causal connections that we impose here.[4]

[4] This is Kant's reply to Hume. Scottish philosopher David Hume believed that causation had no objective foundation and that we formulate causal claims based on simple convention. Kant responded that formulating causes is not a simple matter of convention since we automatically make causal connections. Additionally, Kant thought that this conveys objectivity because he links objectivity with necessity.

However, when it comes to producing scientific claims about causation, we need to move beyond the causes we see in the conceptually constituted world. Although we must apply causation, there are so many different ways in which we can constitute sensation as causal. We might see objects causally interacting in the observable world but we do not see *types* of objects or *types* of causes (see Kant 2000, 20:203). We need to employ additional principles to order the observable world and constitute thought entities as causal. That a reaction occurs when one substance interacts with another substance is visible in the observable world but to say that one element has an oxidation reaction with another element is not visible in that observable world. We need to move beyond that observable world to say that one type of thing undergoes an oxidation reaction with another type of thing rather than simply saying one thing causes a colour change in another thing. We also need to move beyond that observable world to say that one substance will interact with other substances, with particular results, in a range of circumstances other than those observed in the observable world at a particular time. To move beyond the observable world we need to add something to that observable world, reorganise it to conceptualise different causal relations. We need to be able to say an observable object which we see causally interact with other observable objects is a type of object which engages in a type of causation in particular situations.

This works in two different ways. Firstly, whilst causation is itself an *a priori* principle used to structure experience, we may have to pick from a particular type of causal *a priori* principle to generate more specific causal claims (as described in Chap. 3, Kant wrongly thought there only was one set of *a priori* principles). As outlined in Chap. 3, scientists employ a relativised *a priori* to constitute a world of causal connections as taking a particular form. Secondly, we need place conceptualised causal connections into a system of science through applying principles. Doing this refines the causal claims, correcting them and improving them.

Despite seeing causes as mind-dependent neo-Kantians still believe that causes can convey epistemic legitimacy. The typical reason why philosophers and scientists see causes as providing scientific legitimacy is that those causes are mind-independent parts of the external world. As such, neo-Kantians need take an alternative approach. As discussed in Chap. 3,

neo-Kantians value systematicity whereby scientific concepts are placed within a broader framework of science. Causes are important to neo-Kantians because they can be used to connect different parts of science together, producing a system of science (Cassirer 1957 [1927], p. 33; Kant 1998 A663/B691; 2000 5:183–184; Kitcher 1989, p. 436; Massimi 2008, p. 33).[5] Rather than having any two parts of science being unconnected they can be connected together by specifying one part as the cause of another part. Relating different parts of a science together in a causal manner significantly increases the epistemic goal of building a system of science. It shows interconnections between different areas and it embeds claims within one another. Thus the more causal connections included within a system the greater the interconnections the system has. Therefore, causes are important to neo-Kantians even though causes are not mind-independent.

6.4 Causation in Psychiatry

In this chapter I focus upon the relationship between diagnoses and causes, giving a neo-Kantian account of how they relate to one another. Before doing so, it is briefly worth considering how a neo-Kantian would consider causes in psychiatry in general, rather than considering how they relate to psychiatric diagnoses.

We observe causal connections in the conceptually constituted world in many different areas of psychiatry. We can do experimental research in psychology where we watch the reactions of people in various setups. We can observe quick or slow reactions to stimuli, we can use eye trackers to observe which general area eyes are looking at and we can observe how someone reacts to others in a social situation. Similarly, we can observe areas of a screen light up on a scan when doing neuroscience and we can observe colour changes when doing genetic research. In a sense these are all observations of causation in the world around us providing we understand the observable world to be conceptually constituted. That is, the visual field that we have access to is conceptually constituted, and within

[5] For analysis of Kant's position see Butts 1990, p. 9 Guyer 1990, p. 41; McNulty 2015, p. 4.

that visual field we see one thing as following another and take this as causation. In this sense we are seeing causes occur in psychological, neuroscientific and genetic research, providing we understand what we observe in these experiments to be conceptually constituted, including by the concept of causation. The concept is prior to the observation. We cannot simply make up what causes we see in psychiatric research even though, as I have argued, we should not think of causes as being out there in the mind-independent world waiting to be discovered.

The causation we observe in the conceptually constituted world is typically too narrow for most scientific purposes. We need to bring it under further concepts and principles to idealise it and abstract it. This is in line with my examples from Chap. 4. We get variable results when melting lead but we say lead melts at 325 degrees due to the molecular structure of lead causally determining the melting point to be 325 degrees. We say things fall with uniform acceleration because gravity makes acceleration uniform even though things actually fall at a non-uniform rate due to factors like air resistance and wind speed. Similarly, in psychiatry we typically posit broader causes which abstract and idealise away from the specifics of the results we observe.

To highlight this I shall consider theory of mind-deficits. I choose this because it is well known and well studied rather than because it is the best causal explanation of autism. Typically, false belief tests are used to study theory of mind-deficits. Participants in experiments are required to understand that someone fictional lacks knowledge which the participant has. For example, a child watches a scene where two dolls are interacting in a room. Doll 1 and Doll 2 both see a marble being placed in a particular location in the room. Doll 1 leaves the room. Doll 2 then moves the marble to a different location in the room. Doll 1, who has not seen the marble being moved, returns to the room. The participant is asked where Doll 1 thinks the marble is. Most three-year-olds will say that Doll 1 thinks it is in the location it actually is, whereas most four-year-olds will realise that Doll 1 did not see the marble being moved the second time and so report that Doll 1 thinks it is in the location it was before Doll 2 moved it. Autistic people are more likely to fail this test. This then formed an evidence basis for the notion that autistic people have theory of

mind-deficits, whereby they struggle to see others as having mental states and unique perspectives.

I now show how the posited cause of theory of mind-deficits is very different to the causes we see in the world around us. Firstly, within the above described experiment we observe different results in different participants. We then produce a statistical generalisation from that experiment, showing that a certain percentage of participants passed the test. The causes observed are turned into a statistical claim about likelihood of passing and failing. Secondly, the experiment is done many times, typically involving different individuals each time, and the results of these different experiments are pooled together to produce a further statistical generalisation. Thirdly, experiments will often take slightly different forms. For example, there are other false belief tests which do not involve dolls. Additionally, the experiment with the dolls can be made more complicated. Rather than the participant needing to demarcate between X's knowledge of Y's knowledge, the participant might need to demarcate between X's knowledge of Y's knowledge of Z's knowledge. This means quite different causes are observed in these experiments but they are all taken as adding to the statistical generalisations about causal reactions. All these steps then produce multiple overall statistical results. We demarcate the statistical probability of autistic individuals failing false belief tests of varying levels of complexity and at particular ages. These statistical generalisations are then combined with wider evidence and theorising about the mind, such as evidence from genetics and neurology. All of this is combined to posit that autistic people have theory of mind-deficits whereby a part of the mind which most people have is either absent or working at a significantly reduced degree in autistic people. This shows how the observation we observe in the conceptually constituted world is quite different to the cause we end up positing. We cannot see theory of mind-deficits either in those experiments or in experiments in genetics or neuroscience. Rather, we take the observations of causation and bring it under other concepts and principles to posit theory of mind-deficits.

Theory of mind-deficits as a cause is quite nebulous (see Fellowes Forthcoming for discussion). Firstly, evidence suggests that whether and the degree to which autistic individuals express symptoms taken as being caused by theory of mind-deficits can be heavily dependent upon

environmental factors. Secondly, the degree to which an autistic person can see another perspective might depend upon the level of effort they make. They might need to make conscious reflection in a manner that non-autistic people will often not need since perspective taking often, though not always, works unconsciously for non-autistic people. Thirdly, the degree to which an autistic person can potentially make that effort to understand other perspectives might depend upon the level of spare energy which the autistic person has. Unfortunately, given the stress and strains of navigating a world that is primarily set up for non-autistic people, autistic people often have less spare energy than non-autistic people. Fourthly, there is good evidence that non-autistic people also often fail to understand the perspective of others, including the perspective of autistic people. There is something quite loose about claiming that autistic people have the cause of theory of mind-deficits.

My account of theory of mind-deficits fits my discussion of neo-Kantian notions of causation. The causes we observe around us are very different to the cause of theory of mind-deficits. There are many steps of abstraction and idealisation between what we observe and positing theory of mind-deficits. Additionally, given all the steps between theory of mind-deficits and what we observe, it seems that we never encounter pure theory of mind-deficits, just like how Kant thought we cannot observe pure earth or pure water. Theory of mind-deficits is a thought entity in the sense outlined in Chap. 5. It is an intellectual construct which is distant from what we observe in the conceptually constituted world, a hypothesised causal entity which groups together various genetic, neurological and psychological claims and is one cause among many interacting causes, both internal and external to the autistic person, which can play a role in the behaviour exhibited by autistic people.

6.5 Psychiatric Diagnoses and Causes

Having outlined an example of constructing causes in psychiatry, I now relate this to psychiatric diagnoses. Ideally, we would want psychiatric causes to nicely match up with psychiatric symptoms and psychiatric diagnoses. As outlined in Chap. 2, a leading goal in psychiatry since the

1980s was finding individual causes with individual effects. It has been a long running assumption within psychiatry that the same cause will result in the same condition (see for analysis Blashfield and Keeley 2010, p. 325; Vanheule 2017, p. 45). Consequently, it is assumed that the reason why psychiatric diagnoses are so heterogeneous is because they group together individuals with heterogeneous causes (Cuthbert and Insel 2009, p. 3; Volkmar et al. 2009, p. 109). As discussed in Chap. 2, this goal broadly seems to have failed. Recent evidence suggests there is a more complicated relationship between causes and symptoms within psychiatry. The same symptom can arise from multiple causes whilst the same cause can result in multiple symptoms (Cuthbert and Insel 2009, p. 3; Sanislow et al. 2010, p. 632). Here I will give more detail about the manner in which one cause, one affect can be considered to have failed in psychiatry. The key issue is that we might see one cause, one effect when we make observations in the conceptually constituted world, such as observing a particular person making a particular response to a particular stimuli in a psychological experiment, but this is different to one cause, one effect when it comes to thought entities. That being, if theory of mind-deficits is one thought entity and autism is another thought entity then we do not get one cause, one effect. I now discuss this by considering examples from psychology and biology, and then consider what role psychiatric diagnoses can play in creating causal connections.

I return to my above discussion of theory of mind-deficits in autism. Historically, theory of mind-deficits has been considered a key candidate for giving autism a causal basis. The idea is that autism is caused by theory of mind-deficits. If we defined autism on the presence of theory of mind-deficits then we would have one cause (theory of mind-deficits) and one effect (autism). This approach seems to have failed for multiple reasons. Firstly, experimental evidence suggests that some autistic people do not lack theory of mind-deficits given that they can pass theory of mind tests (Happé and Ronald 2008, p. 276). Secondly, theory of mind-deficits can be found in individuals who do not meet the diagnostic criteria of autism, being associated with diagnoses like schizophrenic (Sprong et al. 2007, p. 10) and deaf individuals (Paterson 2016, p. 142). Thirdly, theory of mind-deficits interacts with many other causes. Fourthly, theory of mind-deficits itself is broadly conceived whereby it

has a level of fit with neurological and genetic evidence but is not itself simply caused by one gene or brain structure. Notions of one cause to one affect do not resemble the relationship between theory of mind-deficits and autism.

I now consider an example from neurology. Hoffman and Zachar outline how a particular part of the brain can be responsible for a wide variety of behaviours. They outline how the amygdala appears to be causally related to behaviours such as fear situations, disgust, emotionally neutral novel information and decision making (2017, p. 71). Biological abnormalities in the amygdala can result in abnormal behaviours in all these domains. However, Hoffman and Zachar outline how there seems no clear correlation between abnormalities of the amygdala and which particular abnormal behaviour or behaviours will arise in a particular individual. Consequently, there seems to be no regular relationship between cause and symptom. It could be responded that this shows the amygdala is insufficiently causally fine-grained. Rather, a more specific part of the amygdala might be causally related to more specific behaviours and so would count as the cause. However, Hoffman and Zachar anticipate this response and doubt its viability. They cite an article which reviews 1469 studies. It states that "a typical cortical region is activated by tasks in nine different domains… even when dividing the cortex into nearly 1,000 small regions" (Anderson in Hoffman and Zachar 2017, p. 72). Those 1000 regions are not each correlated with one behaviour. Additionally, Hoffman and Zachar point out that focusing upon a very small area, such as ten neurons, is unlikely to link to specific behaviours because of brain development over time (2017, p. 72).

Due to the complexity of interacting causal factors, they are often linked to psychiatric diagnoses through notions of a developmental pathway (Casey et al. 2014, 351). This is where a chain of causes starts a causal process which changes dynamically over time, producing a variety of results throughout this process. For example, a developmental pathway might start at one year of age, produce some symptoms of autism, and then the developmental pathway undergoes changes and new symptoms of autism start appearing at age three, then at age four, age five and so on. Such developmental pathways do not work in an easily predictable manner. The route which developmental pathways take might depend upon

the initial causes, and other internal or external causal factors. In one individual, the developmental pathway might have been running for six months before it interacts with an unrelated cause, whereas in another individual the developmental pathway was running for one year before interacting with that cause whilst a third individual never has that cause so no interaction takes place.

All this means that causes typically only work in a probabilistic manner in psychiatry. The same causes can lead to multiple outcomes and the same outcome can be produced by different causes (or combinations of causes) operating independently of one another. Consequently, knowing that a particular cause is present does not mean knowing that a particular outcome will follow. Similarly, knowing that a particular outcome is present does not mean knowing that it was preceded by a particular cause (Cloninger 2014, p. 205; Kendler 2005, p. 1247). Also, a particular cause can be present but whether it activates can depend upon the presence or absence of other causes, such as how the environment can switch on or off genes (Morris et al. 2014, p. 7). Additionally, if it does activate then what results follow can depend upon the presence or absence of other causes. Again, the presence of a particular cause or a particular result does not entail, respectively, the presence of a particular result or cause. Rather, we can find probabilistic associations between them.

I now argue that psychiatric diagnoses can play a role in constituting causes. I argue that psychiatric diagnoses can form probabilistic bridges between causes and symptoms. In doing so they create new causal connections, making possible new causes.

Psychiatric diagnoses can provide a tractable link between probabilistic causes and probabilistic symptoms. Rather than directly linking probabilistic causes with probabilistic symptoms they can instead be linked together through a psychiatric diagnosis. Understood in this way the probabilistic causes which are present in an individual entail a probability that the individual will meet the diagnostic criteria for a psychiatric diagnosis and then the psychiatric diagnosis entails certain probabilities that certain symptoms will be exhibited. If, say, five causes were probabilistically associated with a psychiatric diagnosis then knowing any combination of causes is present (be it just one, some or all) will give a probabilistic indication that the individual affected by those causes will meet the

diagnostic criteria of that psychiatric diagnosis. The psychiatric diagnosis will then give probabilistic indication of particular symptoms occurring given that DSM diagnoses are polythetic whereby someone can meet the diagnostic criteria and only exhibit some of the symptoms.

For example, Schaffner gives an example of how many factors combine together to produce the probabilities of major depression occurring (Schaffner 2008, p. 54). Major depression is only probabilistically linked to symptoms because it has nine symptoms on the diagnostic checklist but only five symptoms are required for a diagnosis (one of which must be low mood) so the diagnosis is only probabilistically linked to any of those nine symptoms. The diagnosis of major depression provides a link between the probabilistic causes and probabilistic symptoms. As a second example, Fitzgerald has suggested, in his study of the neuroscience of autism, how autism consists of the link between causes and symptoms. He writes that autism is "a kind of statistical aggregate suspended across these multiple [causal] levels. This is an argument that autism is not simply *heterogeneous* but that it consists *in* measures of statistical commonality between the various sets of difference, until something looks enough like autism begins, even if only in terms of some statistically significant relation, to hang together" (Fitzgerald 2017, p. 149 emphasis original). Autism is not an entity or thing in any sort of conventionally understood manner, rather, it is simply a statistically significant joining point between probabilistic causes and probabilistic symptoms. Various causes, including theory of mind-deficits, have a probabilistic relationship to autism whereby there is a statistical relationship between any cause being present and autism occurring (additionally, simultaneous presence of two or more causes can increase or decrease the probabilities above and beyond the baseline probability for each cause). Autism then has probabilistic relationships to various symptoms.

We should understand psychiatric diagnoses as thought entities which act as a bridge between causes and symptoms. This seems to heavily fit a neo-Kantian picture of science. The diagnoses of major depressive disorder and autism are constructions that provide links in a system of science, unifying together different parts of the system through being in a statistical relationship to (conceptually constituted) causes and symptoms.

By bridging probabilistic causes and probabilistic symptoms the psychiatric diagnosis is one way of making causal knowledge possible. The probabilistic link between causal factors and the psychiatric diagnosis does not exist outside of employing the psychiatric diagnosis. The probabilistic links between the psychiatric diagnosis and the symptoms also do not exist outside of employing the psychiatric diagnosis. Consequently, the combined set of probabilistic links from causes to psychiatric diagnosis and then from the psychiatric diagnosis to the symptoms does not exist outside of employing a psychiatric diagnosis. Different probabilistic links are made if moving straight from cause to behaviour without the intermediate psychiatric diagnosis acting as a bridge. In this sense the psychiatric diagnosis makes causal knowledge possible which is different to causal knowledge produced by not employing the psychiatric diagnosis. Rather than basing psychiatric diagnoses on charting the immense number of possible relationships between any particular cause and any particular symptom psychiatrists can instead chart the relationship between causes and diagnoses and then chart the relationship between diagnoses and symptoms. This tames much of the complexity and adds much tractability. By acting as an intermediary between probabilistic causes and probabilistic symptoms, psychiatric diagnoses can make useful, tractable causal knowledge possible.

6.6 Categories and Dimensions

The causal knowledge made possible will depend upon whether categorical or dimensional diagnoses are employed. Whereas categorical notions of autism only cover people who meet a cut off, dimensional notions of autism cover all manifestations of autism, meaning it covers everyone who is an instance of categorical autism and many other people beyond that. The causal connections created when psychiatric diagnoses act as a bridge vary depending upon whether categorical or dimensional diagnoses are employed. In this sense they act as an *a priori* principle rooted in reason (as described in Chap. 5) which constitutes knowledge in different ways. If we take all people who meet the diagnostic criteria for categorical notions of autism, such as DSM criteria, we will find a statistical

probability of any cause being present and if we take all people who exhibit the criteria for dimensional notions of autism (that being, they exhibit any characteristic of autism) then we typically find a different statistical probability of the cause being present. As previously discussed, the same cause can be found within categorical notions of autism and dimensional notions of autism. That being, some people who meet the categorical criteria can exhibit various causes, and people who do not meet the categorical criteria can also exhibit those same causes.

It is certainly advantageous to use dimensional notions to track these causes whereby we aim to study who actually exhibits the cause rather than simply ignore the presence of these causes in anyone who does not meet the categorical diagnosis. However, it is also advantageous to use categorical psychiatric diagnoses because they tracks different combinations of causes compared to dimensions. This is important because the causal effect of one cause can change when other causes are present. One cause can affect another cause, making it operate in a different way. Also, one cause can produce another cause. In both these cases, the altered cause or the new cause can then change how a symptom expresses, cause a new symptom to occur or prevent a symptom occurring.

6.6.1 Causes Being Causally Affected by Internal Causes

How the brain develops can be influenced by many different causes. Rather than there being a set of causes present at birth that always have the exact same effect, how those causes operate can be influenced by other causes, both internal and external, which can then change the effects produced. There can be "changes in the structure of existing neurons and the function of neural circuitry in response to intrinsic and extrinsic stimuli" (Morris et al. 2014, p. 6), and "sustained periods of neuronal activation or depression… will eventually lead to tropic effects [growth] or atrophy of the brain area involved" (Spedding et al. 2003, p. 35).

I firstly consider internal causes, such as genetics, neurology and psychology. These are internal in so far as they are present within the individual's body or mind (even if they can be affected by environmental

causes, as I outline in the next section). These internal causes are typically understood as occupying different levels. Genetics, neurology and psychology are all different levels. Causes within a level can interact with one another. The causal effect of one gene can be different depending upon what other genes are present. Additionally, whether a gene switches on or not (i.e. start or stop building proteins) can be influenced by the presence of other genes. Similarly, different neurons interact together whereby the effects of one brain structure might be influenced by other structures to produce further effects. Also, how one area of the brain develops can depend upon the presence or absence of other areas of the brain. One area might depend upon another so it underdevelops if another area is underdeveloped or absent. Alternatively, where one area of the brain is underdeveloped another area may overdevelop to compensate. Finally, different psychological causes can interact together and one psychological state can produce other psychological states, such as fear leading to anxiety. Different levels of the brain can then interact. Genes can switch on or off based upon interactions with neurology and psychology. Brain structures can develop differently depending upon the presence or absence of genes and psychological states. Finally, both genes and neurology can affect psychology. Understanding the relationships between these (and other causes listed below) is immensely complicated.

Categorical and dimensional notions of autism will cover different groupings of interacting causes. Within a categorical notion of autism most causes associated with autism will have a higher probability of being present compared to dimensional notions of autism. Given that a greater number of relevant potentially causally interacting factors are present in categorical notions of autism rather than dimensional notions of autism, these produce different statistical probabilities of which symptoms manifest and how they manifest. The causal probabilities will be different if we consider the causal relationships within everyone who is categorically autistic and compare this with everyone who is dimensionally autistic. If we specify a narrower grouping, such as everyone with a particular cause or symptom, or particular combination of causes and symptoms, then different causal relationships will still follow. If we pick everyone who simultaneously has executive dysfunction and low eye contact in categorical notions of autism then we will cover individuals who generally have

a relatively high number of causes of autism whereas looking at individuals with executive dysfunction and low eye contact within dimensional notions of autism will cover some individuals who exhibit many causes of autism and cover some individuals who exhibit relatively few causes of autism. As such, whether looking at the entire group or a particular combination of causes and symptoms, categorical and dimensional approaches constitute different causal probabilities.

6.6.2 Causes Being Causally Affected by External Causes

Causes can also be affected by external causes. In Chap. 5 I outlined how the environment can causally affect which symptoms are present. Here I am considering how the environment can affect causes, although given that causes affect symptoms this indirectly means environmental causes also affect symptoms. A cause which is internal to the autistic person can be affected by a cause that is external to the autistic person, which then changes the internal cause. "Biological, chemical and physical factors (such as nutrition, minerals and weather) participate in gene-environment interactions. Cultural and educations backgrounds and responses to stress, which subsequently influence lifestyles, also participate in gene-environment interactions" (Abdolmaleky et al. 2005, p. 156). Other factors include "deprivation of normal parental care during infancy, childhood physical maltreatment, childhood neglect, premature parental loss, exposure to family conflict and violence, stressful life events involving loss or threat, substance abuse, toxic exposure and head injury" (Caspi and Moffitt 2006, p. 584). Drugs, psychotherapies and cognitive training (Morris et al. 2014, pp. 12–14) have all been shown to influence the brain. All this shows that neurons and genes can be influenced by the environment whereby different brain structures result and different genes switch on or off. Also, learning, stress and trauma can impact psychology.

How these interactions take place depends upon two factors. Firstly, which internal causes there are at any particular moment for the environment to interact with. Categorical notions of autism have a stronger association with the causes of autism than dimensional notions of autism.

That being, if someone meets the categorical notion of autism they are likely to exhibit a higher number of causes associated with autism whereas someone who is dimensionally autistic may exhibit a relatively high number of causes of autism or may exhibit a relatively low number. This is still true even if we look at individuals with particular causes. If we specify a sample of individuals with three causes of autism those who meet the categorical criteria are still more likely to exhibit other causes of autism compared to those who are dimensionally autistic. Secondly, those who are categorically autistic are likely to be in different environments compared to people who are dimensionally autistic and those environments can then interact with internal causes. As outlined in Chap. 5, someone who is categorically autistic is more likely to receive early intervention, therapy and support. They are also more likely to be home-schooled or be the outsider in a school, and are more likely to be isolated or unemployed later in life. They are also more likely to be in environments which mismatch with their symptoms, whereby societies are set up for a relatively narrow set of characteristics and those who exhibit greater levels of differences to that narrower set of characteristics will likely suffer greater levels of stress and trauma. All this means that categorical notions of autism will constitute different causal probabilities compared to dimensional notions of autism because of differences in internal causes for the environment to interact with and because of greater likelihood of being in different environments.

6.6.3 RDoC and HiTOP

I now consider how RDoC and HiTOP might approach psychiatric diagnoses, bearing in mind my discussion in Chap. 5 that it is not yet clear what a RDoC and HiTOP clinical diagnosis would look like. My discussion here largely resembles that of Chap. 5.

Firstly, are RDoC and HiTOP diagnoses going to be polythetic whereby different combinations of symptoms or causes are being measured by the dimension and not all need to be present for someone to be rated upon that dimension? If they are polythetic then a category might demarcate more specific causes for the environment to interact with or

might demarcate more specific symptoms which influence which environments someone is more likely to be in.

Secondly, will RDoC and HiTOP diagnoses extend into the general population or will there be a sharp break whereby everyone scores 0 or scores between, say, 100 to 70? If they extend into the population then a further categorical divide might again pick out more specific causes that might then interact with the environment or pick out more specific groupings who have more specific probabilities of being in different environments. This might also be true if everyone scores either 0 or between 100 to 70 since people who score above 85 might differ significantly in these ways compared to those who score between 70 to 85.

Thirdly, will people vary significantly in ways not covered by the dimensions? If so then adding categories might specify more specific causes which could interact with the environment and might specify more specific individuals who are likely to end up in unusual environments.

Fourthly, I earlier made a demarcation between causal mechanisms and the specific causal happenings that they cover. Causal mechanisms are abstract and idealised. Causal mechanisms are applicable to multiple particular instances despite those particular instances having differences in the specific details of the specific causal happenings which occur. Adding categories covers more specific manifestations which the environment could interact with and covers individuals being in more specific environments.

6.7 Systematicity

Having shown how categorical and dimensional approaches can play a role in creating causal knowledge I now consider how causes can help establish which psychiatric diagnoses we should formulate. As mentioned earlier, neo-Kantians consider identifying the causes of an entity to convey epistemic strength to that entity. Neo-Kantians value placing scientific entities within a wider system of science. By connecting together different parts of science it becomes more unified whereby seemingly unrelated situations are connected together. Causes are a means of

making such connections. Rather than two phenomena standing independently of one another they are instead placed within a system of science that specifies that one phenomenon causes the other.

Connecting together causes and scientific entities can also be used to formulate scientific entities. One entity can be demarcated from another based upon how they relate to causes. Science involves bringing appearances under concepts. Consequently, some parts of appearances may look similar but once brought under concepts by relating them to another part of science they are formulated as different things. Two people may exhibit different characteristics but once brought under concepts through considering their genetic make-up they might be considered as being instances of the same entity. As described in earlier chapters, science involves a transformation of what we observe by abstracting and idealising to find common patterns in what seems to be otherwise different situations. Causes are a means of doing this.

This means the neo-Kantian philosophy I draw upon fits the desire of psychiatrists to validate psychiatric diagnoses on causes. We can employ causes to potentially give justification to psychiatric diagnoses. By finding the causes of a psychiatric diagnosis we can connect the psychiatric diagnosis greater to a system of science. Ideally, we would connect causes to diagnoses, and diagnoses to symptoms, and symptoms to behaviour, and potentially connect some of these back to causes, creating multiple interconnections. As described in Chap. 2, psychiatrists have heavily emphasised the importance of finding the causal basis of psychiatric diagnoses. This means that neo-Kantianism is in line with this general goal of psychiatry. However, psychiatrists historically have sought a particular type of causal basis for categorical psychiatric diagnoses, that being sensitive and specific causes, and the failure to find these has led to calls to employ dimensional rather than categorical psychiatric diagnoses. I now argue that we can employ non-sensitive and non-specific causes to provide a causal basis for categorical psychiatric diagnoses but in doing so we need to fit psychiatric diagnoses into a wider system of science through applying principles.

I first consider non-sensitive and specific causes in relation to DSM diagnoses. Autism can be formulated on cognitive psychological causes in many different ways. The most promising means of causally formulating

autism has historically been considered to be the three cognitive psychological causes. These are theory of mind-deficits, weak central coherence and executive dysfunction. I shall consider these because they are well known and have been studied at great length, rather than because I endorse them as being good theories. These appear to have a stronger correlation with autism than neurological and genetic factors because autistic individuals are quite likely to have those cognitive psychological causes. Though often present in autistic individuals these causes are not always present (in terms of validity they are not sensitive). Some diagnosed individuals pass tests for theory of mind-deficits (Happé and Ronald 2008, p. 276), executive dysfunction (Happé and Ronald 2008, p. 296) and weak central coherence (Happé and Frith 2006, p. 15). This suggests that some autistic individuals do not have those causes. Additionally, those causes seem to be present in individuals who are not autistic (in terms of validity they are not specific). Theory of mind-deficits are sometimes detected within schizophrenic (Sprong et al. 2007, p. 10) and deaf individuals (Paterson 2016, p. 142). Executive dysfunction can be sometimes detected in schizophrenia (Orellana and Slachevsky 2013, p. 5) and ADHD (Happé and Ronald 2008, p. 296). Weak central coherence can be detected in schizophrenia, Williams syndrome and depression (Happé and Frith 2006, p. 15). Though those cognitive psychological causes have much greater correlation with autism than do neurological or genetic factors they are not always present in autistic individuals and can be present in individuals who are not autistic.

We can consider how the diagnosis of autism links to the three cognitive psychological causes it is associated with. One approach would be to stringently link together causes and diagnoses by creating new psychiatric diagnoses based upon the presence of various types of causes. This would mean accommodating different combinations of cognitive psychological causes to different psychiatric diagnoses, whereby one combination (say, theory of mind-deficits and weak central coherence) is accounted for by one psychiatric diagnosis whilst another has a different combination (say, theory of mind-deficits and executive dysfunction) which was accounted for by a different psychiatric diagnosis. Given three cognitive psychological causes, this results in seven combinations and so would require seven new diagnoses to cover all three.

Formulating an additional seven diagnoses has both advantages and disadvantages on the principles that neo-Kantians employ. It makes part of the system more stringent through increasing the statistical probabilities between different parts of the system. Each particular combination of causes is linked to particular diagnoses to a higher degree. This reduces the gap between the diagnosis and what it covers. For example, the simultaneous presence of weak central coherence and executive dysfunction has a higher probabilistic link with the diagnoses based around both those causes than they do to the current diagnosis of autism. Also, the simultaneous presence of those causes, without also theory of mind-deficits being present, will have a lower probabilistic link with the other six new diagnoses compared to the probabilistic links with current notions of autism. However, this greater stringency depends upon formulating a significantly greater number of psychiatric diagnoses. The diagnosis of autism would need to be reformulated into seven different diagnoses. This compromises the goal of employing as few entities as possible. There is a trade-off between these having diagnoses which are closer to what they cover and the goal of employing as few entities as possible. These two alternatives of employing current notions of autism vs replacing autism with seven diagnoses both have different advantages and disadvantages when measured by the principles neo-Kantians employ. I have only considered three cognitive psychological causes to make this point, however, this becomes more problematic given that there are multiple other potential psychological explanations of autism (see Fellowes Forthcoming for summary) and becomes especially problematic if we instead wished to base autism around the vastly greater number of neuroscientific and genetic causes.

In relation to projects like RDoC and HiTOP, my discussion largely parallels the points I raised in Chap. 5. We can use mathematical and statistical techniques like cluster analysis to try and find the smallest number of factors which best account for the relationships between the data. In Chap. 5 I considered this in relation to symptoms. We can also do this in relation to causes whereby we aim to track mathematically how best to cluster symptoms together to then create the best fit with causes. This would result in a set of dimensional psychiatric diagnoses which cover a set of probabilistically related symptoms and that grouping of symptoms has probabilistic relationships with various causes. Analysing

data using mathematical and statistical techniques heavily fits the neo-Kantian picture of science whereby we apply principles of reason to data to produce a structure of science composed of hypothetical entities. It especially fits Kitcher's notion of producing a system of science which accounts for the relationship between phenomena, including between symptoms and causes, as best as possible whilst containing as few entities as possible. As outlined in Chap. 5, this type of statistical analysis by RDoC and HiTOP might result in biotypes and might result in multiple constructs (i.e. 'Social Communication Difficulties', 'Altered Sensation and Attentional Control' and 'Ritualized Behaviour') which people can be rated upon. These could then be employed clinically as dimensional diagnoses. My question here is whether this system should put categorical demarcations upon the dimensional diagnoses. I outlined earlier four different reasons why categorical diagnoses could be worth putting on dimensional diagnoses. As outlined in Chap. 5, it could be responded that this simply shows that we are not employing enough dimensions and that by employing sufficiently fine-grained dimensions would mean that adding categorical divides would not provide significantly different statistical and causal probabilities. However, doing this goes against the neo-Kantian standard of limiting the number of entities in the system. This means that producing a very large number of psychiatric diagnoses which better track very precise combinations of causes being present would not be desirable on a neo-Kantian approach.

How then should we weigh up this trade-off when deciding what psychiatric diagnoses to employ? As discussed in Chap. 5, I do not believe that a neo-Kantian position entails one answer to this question. The principles that neo-Kantians employ constrain what is an acceptable answer but they do not determine the answer. A psychiatric diagnosis is legitimate when it contributes to a system of science which takes a form in line with the principles that neo-Kantians value. Multiple systems, positing different psychiatric diagnoses, can be considered scientifically legitimate but not all psychiatric diagnoses or systems of psychiatric diagnoses should be considered to follow from attempts to balance principles that neo-Kantians endorse. In relation to autism, the psychiatric diagnosis I have primarily focused upon, I do consider it to be scientifically legitimate but I suggested in Chap. 5 that a system of psychiatric diagnoses would be more scientific if we added subtypes to autism.

6.8 Conclusion

Causes are typically seen as making psychiatric diagnoses worthy of belief and scientifically legitimate. This is because causes are seen as part of mind-independent reality. Basing psychiatric diagnoses upon real causes means that psychiatric diagnoses are themselves real, and being real makes them scientifically legitimate. In this chapter I have employed a neo-Kantian approach to argue that scientific causes are mind-dependent. Reality does not determine how scientific causes are formulated, rather, we need to bring what comes to us from the world under principles to formulate scientific causes. Additionally, scientific causes do not determine which are the correct psychiatric diagnoses, rather, we need to make choices when formulating psychiatric diagnoses upon causes. However, a neo-Kantian position does not leave formulating causes or formulating psychiatric diagnoses on causes as subjective or arbitrary. Neo-Kantians can appeal to principles of reason, such as specificity and homogeneity, when formulating causes or linking causes to psychiatric diagnoses. Such principles are a standard of truth and provide scientific legitimacy. This means that causes formulated upon those principles, and psychiatric diagnoses which are linked to causes through such principles, should be considered worthy of belief and scientifically legitimate.

References

Abdolmaleky, Hamid, Sam Thiagalingam, and Marsha Wilcox. 2005. Genetics and epigenetics in major psychiatric disorders: Dilemmas, achievements, applications, and future scope. *American Journal of Pharmacogenomics* 5: 149–160. https://doi.org/10.2165/00129785-200505030-00002.

Andreasen, Nancy. 1995. The validation of psychiatric diagnosis: New models and approaches. *American Journal of Psychiatry* 155 (2): 159–161. https://doi.org/10.1176/ajp.152.2.161.

Baetu, Tudor. 2015. The completeness of mechanistic explanations. *Philosophy of Science* 82: 775–786. https://doi.org/10.1086/683279.

Beebee, Helen, and Nigel Sabbarton-Leary. 2010. Are psychiatric kinds 'real'? *The European Journal of Analytical Philosophy* 6 (1): 11–27.

Blashfield, Roger, and Jared Keeley. 2010. A short history of a psychiatric diagnostic category that turned out to be a disease. In *Contemporary directions in psychopathology*, ed. Theodore Millon, Robert F. Krueger, and Erik Simonsen, 324–336. Gilford: Gilford University Press.

Boone, Worth, and Gualtiero Puccinini. 2016. Mechanistic abstraction. *Philosophy of Science* 83: 686–697. https://doi.org/10.1086/687855.

Boyd, Richard. 1991. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* 61: 127–148. https://doi.org/10.1007/BF00385837.

Butts, Robert. 1990. Teleology and scientific method in Kant's critique of judgment. *Noûs* 24 (1): 1–16. https://doi.org/10.1007/978-94-015-8188-2_6.

Casey, B., Mary Ellen Oliveri, and Thomas Inse. 2014. A neurodevelopmental perspective on the research domain criteria (RDoC) framework. *Biological Psychiatry* 76 (5): 350–353. https://doi.org/10.1016/j.biopsych.2014.01.006.

Caspi, Avshalom, and Terrie Moffitt. 2006. Gene–environment interactions in psychiatry: Joining forces with neuroscience. *Nature Reviews Neuroscience* 7 (7): 583–590. https://doi.org/10.1038/nrn1925.

Cassirer, Ernst. 1957. *The philosophy of symbolic forms volume three: The phenomenology of knowledge*. Yale University Press.

Cloninger, Robert. 2014. Validation of psychiatric classifications: The psychobiological model of personality as an exemplar. In Alternative perspectives on psychiatric classification, DSM, ICD, RDoC, and beyond, eds. Peter Zachar, Drozdstoj St Stoyanov, Massimiliano Aragona, Assen Jablensky. 201–223. Oxford: Oxford University Press.

Cooper, Rachel. 2005. *Classifying madness: A philosophical examination of the diagnostic and Statistical Manual of Mental Disorders*. Dordrecht: Springer.

———. 2007. *Psychiatry and the philosophy of science*. Stocksfield: Acumen.

Craver, Carl. 2009. Mechanisms and natural kinds. *Philosophical Psychology* 22 (5): 575–594. https://doi.org/10.1080/09515080903238930.

———. 2013. Functions and mechanisms: A perspectivalist account. In *Functions: Selection and mechanisms*, ed. Philippe Huneman, 133–158. Dordrecht: Springer.

Cuthbert, Bruce N., and Thomas R. Insel. 2009. Endophenotypes: Bridging genomic complexity and disorder heterogeneity. *Biological Psychiatry* 66: 988–989. https://doi.org/10.1016/j.biopsych.2009.10.008.

Dupré, John. 1993. *The disorder of things: Metaphysical foundations of the disunity of science*. Massachusetts: Harvard University Press.

Fellowes, Sam. Forthcoming. Philosophical issues raised by autism: Sociality, explanations, heterogeneity and neurodiversity. *Handbook for the Philosophy of Medicine*.

Fitzgerald, Des. 2017. *Tracing autism: Uncertainty, ambiguity, and the affective labor of neuroscience*. University of Washington Press.

Garson, Justin. 2013. The functional sense of mechanisms. *Philosophy of Science* 80: 317–333. https://doi.org/10.1086/671173.

Glennan, Stuart. 2002. Rethinking mechanistic explanation. *Philosophy of Science* 69: 343–353. https://doi.org/10.1086/341857.

Goodyer, I.M. 2012. Mathematical models as an aid for improving the validity of descriptive psychiatry. *British Journal of Psychiatry* 201 (5): 335–336. https://doi.org/10.1192/bjp.bp.112.109538.

Guyer, Paul. 1990. Reason and reflective judgment: Kant on the significance of systematicity. *Noûs*: 17–43. https://doi.org/10.2307/2215611.

Hacking, Ian. 1991. A tradition of natural kinds. *Philosophical Studies* 61: 109–126. https://doi.org/10.1007/BF00385836.

Happé, Francesca, and Uta Frith. 2006. The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders* 36 (1): 5–26. https://doi.org/10.1007/s10803-005-0039-0.

Happé, Francesca, and Angelica Ronald. 2008. The 'Fractionable Autism Triad': A review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychological Review* 18: 287–304. https://doi.org/10.1007/s11065-008-9076-8.

Hoffman, Ginger, and Peter Zachar. 2017. RDoC's metaphysical assumptions: Problems and promises. In *Extraordinary Science and Psychiatry*, ed. Jeffery Poland and Şerife Tekin, 59–86. Massachusetts: MIT Press.

Hyman, Steven. 2010. The diagnosis of mental disorder: The problem of reification. *Annual Review of Clinical Psychology* 6: 155–179. https://doi.org/10.1146/annurev.clinpsy.3.022806.091532.

Kant, Immanuel. 1998. *Critique of pure reason*, The Cambridge edition of the works of Immanuel Kant. Ed. and Trans. Paul Guyer and Allen Wood. Cambridge: Cambridge University Press.

Kant, Immanuel. 2000. *Critique of the power of judgment*, The Cambridge edition of the works of Immanuel Kant. Ed. and Trans. Paul Guyer. New York: Cambridge University Press.

Kendell, Robert, and Assen Jablensky. 2003. Distinguishing between the validity and utility of psychiatric diagnosis. *American Journal of Psychiatry* 35: 139–144. https://doi.org/10.1176/appi.ajp.160.1.4.

Kendlar, Kenneth, Peter Zachar, and Carl Craver. 2011. What kinds of things are psychiatric disorders? *Psychological Medicine* 41: 1143–1150. https://doi.org/10.1017/S0033291710001844.

Kendler, Kenneth. 2005. "A Gene for…": The nature of gene action in psychiatric disorders. *American Journal of Psychiatry* 162: 1243–1252. https://doi.org/10.1176/appi.ajp.162.7.1243.

Kitcher, Philip. 1989. Explanatory unification and the causal structure of the world. In *Scientific explanations*, ed. Philiphp Kitcher and Wesley C. Salmon, 410–506. University of Minnesota Press.

Kotov, Roman, Robert F. Krueger, David Watson, Thomas M. Achenbach, Robert R. Althoff, R. Michael Bagby, Timothy A. Brown, William T. Carpenter, Caspi Avshalom, Lee Anna Clark, Nicholas R. Eaton, Miriam K. Forbes, Kelsie T. Forbush, David Goldberg, Deborah Hasin, Steven E. Hyman, Masha Y. Ivanova, Donald R. Lynam, Kristian Markon, Joshua D. Miller, Terrie E. Moffitt, Leslie C. Morey, Stephanie N. Mullins-Sweatt, Johan Ormel, Christopher J. Patrick, Darrel A. Regier, Leslie Rescorla, Camilo J. Ruggero, Douglas B. Samuel, Martin Sellbom, Leonard J. Simms, Andrew E. Skodol, Tim Slade, Susan C. South, Jennifer L. Tackett, Irwin D. Waldman, Monika A. Waszczuk, Thomas A. Widiger, Aidan G.C. Wright, and Mark Zimmerman. 2017. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology* 126 (4): 454–477. https://doi.org/10.1037/abn0000258.

Kozak, Michael J., and Bruce N. Cuthbert. 2016. The NIMH research domain criteria initiative: Background, issues, and pragmatics. *Psychophysiology* 53: 286–297. https://doi.org/10.1111/psyp.12518.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. Thinking about mechanisms. *Philosophy of Science* 67: 1–35. https://doi.org/10.1086/392759.

Massimi, Michela. 2008. Why there are no ready-made phenomena: What philosophers of science should learn from Kant. In *Kant and philosophy of science today*, ed. Michela Massimi, 1–35. Cambridge: Cambridge University Press.

McNulty, Michael Bennett. 2015. Rehabilitating the regulative use of reason: Kant on empirical and chemical laws. *Studies in History and Philosophy of Science Part A* 54: 1–10.

Meehl, Paul. 1995. Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist* 50 (4): 266–275.

Morris, Sarah, Judith M. Rumsey, and Bruce N. Cuthbert. 2014. Rethinking mental disorders: The role of learning and brain plasticity. *Restorative Neurology and Neuroscience* 32 (1): 5–23. https://doi.org/10.3233/RNN-139015.

Murphy, Dominic. 2006. *Psychiatry in the scientific image*. Massachusetts: Massachusetts Institute of Technology.

———. 2014. Natural kinds in fold psychology and in psychiatry. In *Classifying psychopathology*, ed. Harold Kincaid and Jacqueline A. Sullivan, 105–122. Massachusetts: MIT Press.

Orellana, Gricel, and Andrea Slachevsky. 2013. Executive functioning in schizophrenia. *Frontiers of Psychiatry* 4 (35): 1–15. https://doi.org/10.3389/fpsyt.2013.00035.

Overton, James. 2011. Mechanisms, types, and abstractions. *Philosophy of Science* 78: 941–954. https://doi.org/10.1086/662274.

Paterson, Candida. 2016. Empathy and theory of mind in deaf and hearing children. *Journal of Deaf Studies and Deaf Education* 21 (2): 141–147. https://doi.org/10.1093/deafed/env058.

Pies, Ronald. 2008. Beyond reliability: Biomarkers and validity in psychiatry. *Psychiatry* 5 (1): 48–52.

Psillos, Stathis. 1999. *Scientific realism: How science tracks truth*. Oxon: Routledge.

Rosenberg, Alex. 2018. Making mechanisms interesting. *Synthese* 195: 11–33. https://doi.org/10.1007/s11229-015-0713-5.

Sanislow, Charles, Daniel S. Pine, Keven J. Quinn, Michael J. Kozak, Marjorie A. Garvery, Robert K. Heinssen, Philip Sung-En Want, and Bruce N. Cuthbert. 2010. Developing constructs for psychopathology research: Research domain criteria. *Journal of Abnormal Psychology* 199 (4): 631–639. https://doi.org/10.1037/a0020909.

Schaffner, Kenneth. 2008. Etiological models in psychiatry: Reductive and non-reductive approaches. In *Philosophical issues in psychiatry, explanation, phenomenology, and nosology*, ed. Kenneth S. Kendler and Joseph Parnas, 48–89. Baltimore: Johns Hopkins University Press.

Spedding, Michael, Isabelle Neau, and Laszlo Harsing. 2003. Brain plasticity and pathology in psychiatric disease: Sites of action for potential therapy. *Current Opinion in Pharmacology* 3 (1): 33–40. https://doi.org/10.1016/s1471-4892(02)00008-5.

Sprong, Mirjam, Patricia Schothorst, Ellen Vos, Joop Hox, and Herman Van Engeland. 2007. Theory of mind in schizophrenia. *British Journal of Psychiatry* 191: 5–13. https://doi.org/10.1192/bjp.bp.107.035899.

Tabb, Kathryn. 2019. Philosophy of psychiatry after diagnostic kinds. *Synthese* 196: 2177–2195. https://doi.org/10.1007/s11229-017-1659-6.

Theurer, Kari. 2013. Seventeenth-century mechanism: An alternative framework for reductionism. *Philosophy of Science* 80: 907–918. https://doi.org/10.1086/673894.

Tsou, Jonathan. 2016. Natural kinds, psychiatric classification and the history of the DSM. *History of Psychiatry* 27 (4): 406–424. https://doi.org/10.1177/0957154X1665658.

Vanheule, Stijn. 2017. *Psychiatric diagnosis revisited: From DSM to clinical case formulation*. Cham: Palgrave Macmillan.

Volkmar, Fred, Matthew State, and Ami Kiln. 2009. Autism and autism spectrum disorders: Diagnostic issues for the coming decade. *Child Psychology and Psychiatry* 50 (1–2): 108–115. https://doi.org/10.1111/j.1469-7610.2008.02010.x.

7

Individual People

7.1 Introduction

I have outlined how psychiatric diagnoses are the product of multiple abstractions and idealisations. We observe others in the conceptually constituted world around us. We take the behaviour they exhibit and idealise and abstract it into symptoms whereby behaviour is removed or distorted. We then take the symptoms exhibited by particular individuals and rearrange them into psychiatric diagnoses which are different to the specific symptoms exhibited by people who meet the diagnostic criteria. This means there is a significant gap between psychiatric diagnoses and actual people. This raises the question why to employ psychiatric diagnoses rather than just directly understand actual people. Even if psychiatric diagnoses can generate knowledge in some sort of idealised and abstract systematised sense, why not just ignore this given it is distant from actual people and instead just focus upon actual people? In this chapter I consider the claim that we do not need to employ psychiatric diagnoses. I then show how psychiatric diagnoses can play a role in understanding actual people. I also consider how there are significant aspects of people which cannot be understood through psychiatric diagnoses, showing

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_7

significant limits to psychiatric diagnoses. In doing so I consider how a neo-Kantian would understand those elements which cannot be captured by psychiatric diagnoses.

7.2 Individuals vs the DSM and ICD

Psychiatric diagnoses are the product of abstracting and idealising observations of people. Also, the types of observations used to construct psychiatric diagnoses are very selective ones. The observations we make primarily relate to conceptually constituted observable behaviour. This then can provide information about things like level of social skills and repetitive behaviour from which we can construct (by bringing under further concepts) DSM style symptoms. This, however, leaves out much which is relevant. For example, it does not cover issues like their past life experience, their present life situation and their future life goals. All these can be important for understanding the mental health that someone has. It can be crucial for understanding things like trauma and life satisfaction. Additionally, the observations used to construct psychiatric diagnoses do not provide information about what meaning an individual places upon their behaviour, their symptoms, their diagnosis or their wider life. They might embrace their diagnosis and symptoms or they might try and resist them. They may see them as an integral part of who they are or as something to be overcome. It also does not cover what meaning they have of their life more generally, not covering issues like their culture, any religious beliefs, political attitudes, social attitudes and moral attitudes, and also what stance they take to all of these. The meaning or lack of meaning, and the nature of any meaning, that an individual holds to any aspect of their life is almost completely not covered by the types of observations we use to construct psychiatric diagnoses. There is very much a "deemphasis on the patient's subjective experiences in DSM and ICD" (Aftab et al. 2024, p. 449). As Tekin notes, "the self has rarely been the *object* of scientific research [in psychiatry]" (Tekin 2019, p. 2198 emphasis original; see also Johnstone 2018, p. 33; Kinderman et al. 2013, p. 2).

It is very much true that on my neo-Kantian account that psychiatric diagnoses are idealisations and abstractions which miss out important

aspects of particular individuals. Establishing that more detailed picture is extremely important and should be done. As such, categorical psychiatric diagnoses are not enough and need supplemented with something more. Later I will consider what a neo-Kantian approach would make of that additional element. Before this, I consider the claim that we could simply dispense with the psychiatric diagnosis altogether to focus upon the specific individual.

A potential alternative to employing categorical psychiatric diagnoses is a psychological formulation. This is the

> process of co-constructing a hypothesis or 'best guess' about the origins of a person's difficulties in the context of their relationships, social circumstances, life events, and the sense that they have made of them… Formulation draws on two equally important sources of evidence: the clinician brings knowledge derived from theory, research, and clinical experience, while the service user brings expertise about their own life and the meaning and impact of their relationships and circumstances. (Johnstone 2018, p. 32; see also Vanheule 2017, p. 172)

The formulation aims to describe specific behaviours and feelings of an individual and an explanation why they act and feel that way (Johnstone 2018, p. 32; Vanheule 2017, p. 104). In line with my discussion in earlier chapters, I see person centred approaches as another way of constituting people. We can constitute people as individuals rather than constitute them as part of a wider grouping. By studying the specific behaviour exhibited by the individual and studying the specific causes present in that individual, and by learning about behaviour exhibited and causes present in the past, we can build a set of inductive and causal probabilities. These will be specific to the individual rather than based upon the probabilities associated with a wider class.

I strongly believe that formulations should be employed alongside categorical diagnoses. However, some critics of the DSM believe that formulations should replace categorical diagnoses. Some quotes highlight this:

> if a psychosocial formulation can provide a reasonably complete explanation for the experiences that have led to a psychiatric diagnosis, what ever

> they may be, then there is no place or need for a competing hypothesis that says 'and by the way, it is also because she has schizophrenia.' The diagnosis becomes redundant. (Johnstone 2018, p. 39)

and

> two people with a diagnosis of 'schizophrenia' or 'personality disorder' may possess no two symptoms in common, [so] it is difficult to see what communicative benefit is served by using these diagnoses. Surely a description of a person's real problems would suffice? A description of an individual's actual problems would provide more information and be of greater communicative value than a diagnostic label. (Kinderman et al. 2013, p. 3)

All this raises a challenge to categorical psychiatric diagnoses. If we can provide a specific description of particular individuals through a formulation then why employ the non-specific idealised and abstract psychiatric diagnosis which is distant from actual people? I now outline two further scientific advantages to employing categorical and dimensional diagnoses.

Firstly, general trends that are not always applicable to specific instances can be better established by abstracting and idealising away from particular individuals. This helps us build a more detailed picture of how particular phenomena interact with one another above and beyond how they manifest in particular individuals. I previously described how symptoms can interact with one another which then alters the manifestation of the symptoms or produces new symptoms. Understanding how this occurs is enhanced by studying both when this occurs and when this interaction does not occur. By studying that symptom X does interact with symptom Y but not symptom Z across many cases we gain a better understanding of how symptom X interacts in general. Similarly, we get a better understanding of a gene not just by considering cases where it interacts with another particular gene but also all the genes it does not interact with. The same point is also true in relation to symptoms interacting with the environment, causes interacting with the environment and symptoms and causes interacting with one another. All of this is especially true given that these interactions are typically probabilistic rather than being the case of two things always interacting or never interacting. Understanding

these sorts of interactions is enhanced by abstracting and idealising away from particular cases to study general trends that occur across many cases.

Secondly, building probabilistic relations between symptoms, between causes and between symptoms and causes can help with detecting symptoms and causes. This is particularly important because symptoms and causes can be difficult to notice. They may take a very subtle form which is difficult to distinguish from something else, they may only occur in specific situations and they might only be detectable using specific methods. Knowing that someone meets the diagnostic criteria for a psychiatric diagnosis entails various probabilities of exhibiting particular symptoms and causes. If we detect some symptoms and causes of psychiatric diagnoses in an individual which are associated with a particular psychiatric diagnosis then we have good reason to look for other symptoms and causes that are associated with the psychiatric diagnosis. This means that we might first establish the individual exhibits relatively easy to detect symptoms and causes and then use this information to help us detect harder to spot symptoms and causes. Psychiatric diagnoses can helpfully guide our awareness to harder to detect symptoms and causes due to the probabilistic associations between the psychiatric diagnosis and both symptoms and causes (for more details see Fellowes 2022).

There are also numerous practical advantages to associating an individual with a generalised class of inductions even though some of those inductions will not be applicable to that particular individual. Firstly, it can help alleviate constraints imposed by time. There can be instances where there is literally no time to outline which symptoms someone has. For example, a police officer or medical personal at the scene of a crime or an accident may benefit from knowing an individual involved has various symptoms. However, the relevant individual may not have time or may not be sufficiently composed to list their symptoms. In contrast, mentioning their diagnoses can very quickly transmit a significant level of important, if non-specific, inductive generalisations. Similarly, in casual social situations it is much easier to inject into the flow of a conversation the name of their diagnosis rather than mention multiple symptoms. Mentioning their psychiatric diagnosis will allow others to make non-trivial inductions about the individual without then dominating the conversation with high levels of detail.

Secondly, psychiatric diagnoses can help with advance knowledge. We might need an indication of what characteristics someone might have years in advance of that individual actually being present and before any particular individual is identified. For example, designing a building or a support service will require knowledge of possible characteristics of individuals before those individuals use the eventual building or service. In both cases a psychiatric diagnosis again can provide inductive probabilistic knowledge of possible behaviour which could be valuable for making relevant decisions.

Thirdly, there can be situations where it is desirable to only indicate that a range of symptoms may be present rather than specify which symptoms are actually present. For example, when applying for a job an individual may wish to indicate that they exhibit symptoms. However, they might not be aware of which symptoms are relevant to the job. As a second example, they might be aware that they are perceived to behave in abnormal or undesirable ways in social situations but not be aware of which symptoms, or how those symptoms, lead to that behaviour. In both these cases it may be more helpful for them to mention that they have a particular diagnosis, entailing that a range of possible symptoms could affect their work and socialising, rather than attempt to specify that a particular symptom has a particular effect in a particular situation.

Fourthly, there can be strong personal reasons to not specify which symptoms someone exhibits. Most people conceal significant aspects of themselves. They rarely provide personal details about themselves to, for example, work colleagues or classmates, and may also conceal personal information from friends and family members. Additionally, when transmitting personal information it is often preferable to give a general indication about personal matters rather than provide a more specific description. Someone might prefer to simply signal to their colleagues that they are feeling low, or signal that they are feeling low because of a dispute with a romantic partner, rather than provide significant details upon what the dispute is about. Similarly, a psychiatric diagnosis is a means of giving a general indication about personal details which lack the specificity of outlining which symptoms of the psychiatric diagnosis they exhibit. It may feel uncomfortable providing significant detail by

specifying which symptoms are present to strangers or even close friends and family members.

Having argued that psychiatric diagnoses can assist understanding actual people, I now consider how people are much more than their psychiatric diagnosis and then consider how a neo-Kantian approach would understand those aspects of people that are not covered by psychiatric diagnoses.

7.3 The Status of Individuals

I now consider a neo-Kantian approach to understanding people as individuals. It might seem like much of the neo-Kantian architecture I have drawn upon is inapplicable to understanding individuals. Throughout this book I have emphasised neo-Kantianism as involving taking conceptually constituted observations and bringing them under concepts and principles to produce idealisations and observations which are distant from what we observe. However, when we aim to understand an individual we do not want an idealised and abstract description of them. We want to understand them as concrete individuals. This raises the question of what use a neo-Kantian position is here. Whilst much of the methods I have previously described will be inapplicable, a neo-Kantian approach can still fit. I now show that observations of individuals are still conceptually constituted, and I consider how empathy or symbolic understanding might be a non-scientific means of understanding individuals.

Some neo-Kantian philosophers have aimed to understand how individual phenomena can be understood. Rickert and Windleband developed neo-Kantian accounts of history. In doing so they demarcated between the natural sciences which aimed for universality whereby a description of an electron is taken as applicable to all electrons, and the historical sciences whereby a description of an historical event or a historical individual, such as the French revolution or Rosa Luxemburg, is only applicable to that event or individual. They emphasised that the natural sciences and the historical sciences need to use different methods. My account of psychiatric diagnoses has very much placed them in the domain of natural sciences, portraying them as idealisations and

abstractions which are distant from actual people. It is useful to note when considering the philosophical status of individuals that there is a difference between understanding something as being an instance of an entity and understanding someone as being an individual. To know someone is an instance of autism is different to understanding them as an individual. To understand the individual I draw upon their neo-Kantian account of the historical sciences. Rickert writes that "[i]t belongs to the character of the natural scientific concept that it is valid for objects that exist in different places and different times. The unique and the individual, on the other hand, always exist in a specific place and at a specific time" (1986 [1902], p. 47). Although both employ different methods, Rickert emphasises that

> [t]he fundamental difference between natural science and history is that natural science forms concepts with a general content, whereas history forms concepts with an individual content. Or, the former generalizes and latter individualizes. But this does not mean that the particular has *no* significance for natural science and the general *no* significance for historical science. Not only are the concepts of the general and the particular relative, but no science at all is possible without general concepts. (1986 [1902], p. 114; see also Broome 2008; Windleband 2015 [1894])

Psychiatric diagnoses are general whereas individuals are individual, but despite this difference we still need general concepts to understand individuals.

The psychiatrist and philosopher Karl Jaspers heavily emphasised the distinction between the individual and psychiatric diagnoses. Jaspers was influenced by many different sources (Walker 2014). Some of his main concepts were directly inspired by Kant's philosophy, he was a student of Windleband and he drew upon Rickert. He was also heavily inspired by sociologist Max Weber who was in turn influenced by Rickert. However, it should be emphasised that Jaspers was also influenced by a much wider range of philosophers, especially in relation to his philosophy of existentialism.

Jaspers made a distinction between understanding and explanation in psychiatry. Understanding relates to meaningful psychic life whereas

explanation relates to the underlying causes of that psychic life. Our psychic life contains many elements which we have varying degrees of awareness of, such as perceptions, thoughts and feelings, and also the meaning we ascribe to them. In contrast, we cannot experience within our consciousness the causes which are producing those perceptions, thoughts and feelings. Those causes can be understood only through scientific investigation. All this means that there is a sharp distinction between descriptions of our psychic life and the causes of our psychic life (Jaspers 1997a [1913], p. 530). As such, full knowledge of the causes of our psychic life does not then reveal the experience and meaning of our psychic life.

Jaspers emphasised that individual people are more than we can ever perceive. Drawing upon Kant's discussion of wholes vs parts, Jaspers argues that people exist as wholes but we can only ever perceive them as parts (1997a [1913], p. 756). We can only ever perceive a limited aspect of their psychic life whereas their psychic life "is not just an agglomeration of separable and isolated phenomena, but presents a total relational context which is in constant flux and from it we isolate our particular data in the very act of describing them" (1997b [1913], p. 58). This means that the "[i]solation of phenomena [of psychic life] makes them sharper and clearer than they really are" (1997b [1913], p. 58; see also 1997b [1913], p. 225). This point is true in relation to understanding raw phenomenological experience at any particular moment or in aiming to understand the biographical details of someone at any particular moment. We can only grasp the parts and not the whole, but we cannot truly understand those parts without also understanding the whole in which people exist.

Additionally, Jaspers was influenced by Max Weber's notion of an ideal type. Max Weber was a sociologist who drew upon numerous philosophical positions, including Kant (Weber 1949 [1904], p. 106). Weber believes that there are an infinite number of causes which influence events in the world and that nothing in the causes themselves can determine which of the infinite causes are the relevant ones. As such, we need to apply our values to decide what is and is not relevant (Weber 1949 [1904], p. 78). To understand the world we apply ideal types which are a "mental construct… [that] cannot be found empirically anywhere in

reality" (Weber 1949 [1904], p. 90). An ideal type is an idealised picture of the world which we can use to make judgements about particular instances in the world. We have an idealised notion of capitalism which we use to understand specific manifestations of capitalism. In doing this we can make judgements about the degree to which particular instances are similar or dissimilar to the ideal type. Jaspers drew upon Weber's notion of ideal types when thinking about psychiatric diagnoses (Walker 2014, p. 323). When we consider someone as an instance of a psychiatric diagnosis we employ an ideal type. We can then ask how far the individual is from the ideal type and find that people have varying degrees of fit with ideal types. This shows that for Jaspers there is a clear separation between people themselves and psychiatric diagnoses since people do not fully resemble their psychiatric diagnoses.

Psychiatric diagnoses are a restricted means of understanding actual people. They can play some role but they also have significant limits. We can instead aim to understand people as individuals rather than as an instance of a diagnosis or alongside understanding them as an instance of a diagnosis. I have shown how understanding individuals still requires concepts and still does not deliver us knowledge of how people actually are. I strongly support understanding people as individuals, considering this equally important as understanding them as instances of psychiatric diagnoses. However, my neo-Kantian account shows that we should consider neither psychiatric diagnoses nor seeing people as individuals as describing people as they actually are. In this sense, notions that focusing upon individuals describes something real whereas psychiatric diagnoses do not describe anything real should be rejected. Focusing on individuals may be less abstract and idealised than psychiatric diagnoses but that does not make it any more real.

7.4 Empathy and Symbols

I now consider two alternative ways of understanding individuals which may challenge a neo-Kantian approach. These are understanding individuals through empathy and understanding them through symbols.

When aiming to understand individuals we can attain a state of empathy with them. One aspect of empathy is where you subconsciously attune with the emotions of another, feeling what they feel. If someone is in pain then you too also feel this pain (Coplan 2011, p. 4). Empathy potentially conflicts with a neo-Kantian approach in two ways. Firstly, there is a tradition of seeing empathy as pre-conceptual. You do not need any concepts to feel what another feels, you just feel it (Coplan and Goldie 2011, p. xxiii). This challenges the neo-Kantian notion that any knowledge of the external world involves concepts. Secondly, some phenomenologists claim that you can literally see the emotions that someone else exhibits (Gallagher 2017, p. 161). If empathy literally involves feeling what another feels then it seems like we have direct knowledge of something in the world. Kant denies that we can know what the world is like in itself, outside of what we contribute to knowledge, but if empathy literally involves feeling what another feels then this might allow a direct experience of something in the external world, namely the feeling of other people.

Also, we might also be able to attain a symbolic understanding of individuals. Rather than attaining a scientific understanding or even a conscious understanding of another, we might instead understand them through symbols. Having argued that individuals are "ineffable" (1997a [1913], p. 754) and "cannot be systematised" (1997a [1913], p. 767), Jaspers argues that some type of deeper understanding of individuals is possible, but only through a symbolic understanding rather than a scientific understanding. Jaspers says that we cannot understand individuals in "the categories of event, causality, substance, force" (1997a [1913], p. 771) but instead a symbolic understanding allows us to understand them (1997a [1913], p. 579). Jaspers writes that

> Only through a lifelong study of poets such as Shakespeare, Goethe, the ancient dramatists and such moderns as Dostoevsky, Balzac, etc., do we arrive at the required intuition, and gain a sufficient store of images and symbols and the ability to exercise an understanding imagination necessary to guide the concrete understanding of the moment… As an investigator into meaning I am conditioned by the sources of my understanding, by such confirmation as I find and by my own problems. These all decide

> whether I remain tied to banal simplifications and rational schemata or whether I endeavour to comprehend men in their most complex manifestations. (1997b [1913], p. 314)

Jaspers thinks that some type of knowledge of people is possible without using the concepts which Kant thinks we must structure the world into.

I now consider how to interpret people using empathy and symbols on a neo-Kantian approach. Empathy does involve concepts and understanding through symbols is still ultimately conceptual, even if both involve imprecise concepts. These concepts, however, would be very different to the sort of concepts typically used in science whereby the concepts of object and causality are inapplicable and some other type of concepts which are vaguer than these are present. On this understanding, empathy and symbols are not grasping anything transcendentally real like an individual in itself but are only grasping something conceptually constituted. Kant says that we bring objects in the outer world under space and time whereas we bring our mental states under time since they do not exist in space. This means our mental states are conceptually constituted and so understanding the mental states of others does not relate to the transcendental. Similarly, in relation to history, Rickert says that particular historical individuals are often understood intuitively rather than through a precise manner present in the natural sciences (1962 [1926], p. 73) but he thinks this is still ultimately conceptual.

This would leave psychiatric diagnoses as having a quite different status compared to knowledge of individual people. Psychiatric diagnoses relate to the external world of conceptually constituted scientific entities which are produced by applying the same sort of concepts which we apply to all science. We can still apply those same concepts to individual people but we soon find there are some aspects of individual people which defy such concepts. At this point we need to use something more. This means employing very different concepts. It means that we need to use different means to understand individuals at a deep level compared to those used to formulate psychiatric diagnoses.

I strongly think that empathy and symbolic understanding should be used to supplement psychiatric diagnoses. As an autistic individual I unfortunately have very limited ability to form empathetic relationships

in the sense of feeling what another feels. However, as I understand it, there are depths to the human psyche which can be attained by empathy and are inaccessible to thought, cognition and reflection. There may be some people who have limited level of emotion for empathy to pick up upon (I suspect this is the case with me) but generally empathy has the potential to reveal something important which cannot be captured by psychiatric diagnoses. Similarly, the symbolic understanding Jaspers talks of convey something which cannot be captured by a psychiatric diagnosis. Additionally, such symbolic understanding can be used both for understanding others and understanding your own self. I have defended the legitimacy of psychiatric diagnoses and argued they can help us understand individual people. They do, however, have very significant limitations. To truly understand people we need far richer sources than the DSM. Tarkovsky and Dostoevsky may deliver truths in a way the DSM cannot.

7.5 Psychiatric Diagnoses and Identity

It could be responded that psychiatric diagnoses are integral to personhood. A psychiatric diagnosis is an inseparable part of who the person is, an inseparable part of their identity. As such, notions of people existing above and beyond their psychiatric diagnosis do not work. I now consider in what manner psychiatric diagnoses could be identities.

Traditionally, psychiatric diagnoses have been seen as in principle separate from the person who received the psychiatric diagnosis. There is the underlying person and the psychiatric diagnosis is something additional to them. To say there is an autistic person is to mention two separate things, the person and autism. On this understanding, autism is an entity or thing which is different to the person. This way of thinking suits some ways of doing psychiatry. By being separable from the person it raises the possibility of curing the diagnosis. It suggests we could remove the entity from someone, meaning someone could be cured of autism and there is still the same underlying person intact.

The neurodiversity movement has heavily challenged this picture. Neurodiversity advocates argue that psychiatric diagnoses are not

somehow a separable entity from the person (see Simkulet 2013, for discussion). There is not the person as one entity and the psychiatric diagnosis as a separate entity, rather, there is just one entity which combines both. Psychiatric diagnoses are an integral part of who someone is. A common notion is that 'you cannot cure me of autism because if you removed autism you would remove me'. By removing autism from someone you would not simply have the same person minus autism, rather, you would have a different person. Put simply, psychiatric diagnoses are part of someone's identity. The diagnosis is an integral part of who someone is. This then leads to various ethical claims. Just as it would be wrong to try and cure, suppress or disregard someone's sexuality, gender or ethnicity, so too it would be wrong to do this to someone's psychiatric diagnosis. In line with these and other characteristics, someone should gain various rights due to their psychiatric diagnosis being part of their identity.

Assessing the notion that psychiatric diagnoses are part of identity is difficult to do without being clear upon what notion of identity is being employed. Notions of identity in philosophy are typically extremely complicated and it is unclear to me how we might translate the notion of identity employed by activists into those employed by philosophers. For example, is identity rooted in scientific claims, rooted in metaphysics, rooted in our language usage, rooted in our ethical practices or something else? What exactly makes something an identity rather than just some separable characteristic? There are lots of different ways of understanding identities and most of them will not be vulnerable to a neo-Kantian critique. However, one answer is that identities are rooted in the fundamental nature of reality. What makes something an identity is based upon facts in the mind independent world. On this understanding, we might be wrong about what is an identity, and what identity someone has, but when we correctly understand that something is an identity then we are tracking something about the nature of reality. If psychiatric diagnoses were real parts of the world and if psychiatric diagnoses were identities then psychiatric diagnoses might be inseparable from people. If in reality someone is an instance of a psychiatric diagnosis, and if in reality that psychiatric diagnosis is an identity, then to remove the psychiatric diagnoses would be to remove the identity, meaning that you end up with

a different person. I argue that we should not see psychiatric diagnoses as identities in this sense.

I have argued that psychiatric diagnoses are distant from what we observe. They are not accurate descriptions of actual people or descriptions of things in themselves. This means we cannot justify psychiatric diagnoses being part of identity on realist grounds because psychiatric diagnoses are an abstraction or idealisation away from our observations of people. Also, they do not reflect mind-independent entities in the sense of a Kantian things in themselves. As such, it seems that psychiatric diagnoses cannot be identities in the sense of being metaphysically part of the fundamental nature of a person based upon metaphysics as relating to the nature of reality.

If, on a neo-Kantian approach, we stop thinking of science as describing the metaphysical nature of reality then we can instead see psychiatric diagnoses and identities as part of the conceptually constituted world? This might mean they are inseparable conceptually rather than metaphysically. If autism is something that can be investigated scientifically and if people are something that can be investigated scientifically then it seems that both are part of the same domain. This means that potentially the two of them could be linked together. Whether the two actually are linked together would depend upon how both of them are conceptually constituted, but its seems possible that we could conceptually constitute them in a manner that puts a necessary link between psychiatric diagnoses and personhood. However, I have appealed to Rickert and Jaspers when discussing how we should understand individuals. Both emphasise that individuals are more than we can know of them scientifically. The concepts used in science can only take us so far and any knowledge of individuals beyond that means using something more than scientific concepts. If psychiatric diagnoses and personhood require very different concepts, or personhood might only be reached through non-conceptual means, then it is difficult to see how a necessary link between them could arise.

There may be very good reasons to see psychiatric diagnoses as identities based on other grounds, such as how they play a practical role in our social and legal systems. Additionally, even if psychiatric diagnoses were not identities, this does not then mean curing psychiatric diagnoses or

denying rights to people with psychiatric diagnoses would be morally acceptable. I simply argue that seeing psychiatric diagnoses as identities on metaphysical or scientific grounds does not work.

7.6 Conclusion

Critics of psychiatric diagnoses argue that there is no need to employ them if we instead can capture relevant information about people through a formulation. Psychiatric diagnoses are abstractions and idealisations which are distant from people so they need supplementing by formulations but there is good reason to still employ psychiatric diagnoses for various scientific and practical purposes. Additionally, I outlined that we need to use concepts when understanding individual people. They are less abstract and idealised than psychiatric diagnoses but they still fit my neo-Kantian account. However, the type of concepts used when understanding individual people are very different to those used to formulate psychiatric diagnoses. This means neo-Kantianism gives reason not to see psychiatric diagnoses as part of identity on metaphysical or scientific grounds.

References

Aftab, Awais, Konrad Banicki, Mark Ruffalo, and Allen Frances. 2024. Psychiatric Diagnoses: A Clinical Guide to Navigating Diagnostic Pluralism. *Journal of Nervous and Mental Disease* 212 (8): 445–454. https://doi.org/10.1097/NMD.0000000000001791.

Broome, Matthew R. 2008. Philosophy as the Science of Value: Neo-Kantianism as a Guide to Psychiatric Interviewing. *Philosophy, Psychiatry, & Psychology* 15 (2): 107–116. https://doi.org/10.1353/ppp.0.0172.

Coplan, Amy. 2011. Understanding Empathy: Its Features and Effects. In *Empathy: Philosophical and Psychological Perspectives*, ed. Amy Coplan and Peter Goldie, 3–18. Oxford: Oxford University Press.

Coplan, Amy, and Peter Goldie. 2011. Introduction. In *Empathy: Philosophical and Psychological Perspectives*, ed. Amy Coplan and Peter Goldie, IX–XLVII. Oxford: Oxford University Press.

Fellowes, Sam. 2022. The Value of Categorical Polythetic Diagnoses in Psychiatry. *The British Journal for the Philosophy of Science 73* (4): 941–963. https://doi.org/10.1086/714801.

Gallagher, Shaun. 2017. Empathy and Theories of Direct Perception. In *The Routledge Handbook of Philosophy of Empathy*, ed. Heidi L. Maibom, 158–168. New York: Routledge.

Jaspers, Karl. 1997a. *General Psychopathology*. Vol. 1. Baltimore, MD: JHU Press.

———. 1997b. *General Psychopathology*. Vol. 2. Baltimore, MD: JHU Press.

Johnstone, Lucy. 2018. Psychological Formulation as an Alternative to Psychiatric Diagnosis. *Journal of Humanistic Psychology* 58 (1): 30–46. https://doi.org/10.1177/00221678177222.

Kinderman, Peter, John Read, Joanna Moncrieff, and Richard P. Bentall. 2013. Drop the Language of Disorder. *Evidence Based Mental Health* 16 (1): 1–3. https://doi.org/10.1136/eb-2012-100987.

Rickert, Heinrich. 1962. *Science and History: A Critique of Positivist Epistemology*. Canada: Van Nostrand Company.

———. 1986. *The Limits of Concept Formation in Natural Science: A Logical Introduction to the Historical Sciences*. Cambridge: Cambridge University Press.

Simkulet, William. 2013. Neurodiversity and Personhood. In *Ethics and Neurodiversity*, ed. C.D. Herrera and Alexandra Perry, 206–216. Newcastle upon Tyne: Cambridge Scholars Publishing.

Tekin, Şerife. 2019. The Missing Self in Scientific Psychiatry. *Synthese 196*: 2197–2215. https://doi.org/10.1007/s11229-017-1324-0.

Vanheule, Stijn. 2017. *Psychiatric Diagnosis Revisited: From DSM to Clinical Case Formulation*. Cham: Palgrave Macmillan.

Walker, Chris. 2014. Karl Jaspers on the Disease entity: Kantian Ideas and Weberian Ideal Types. *History of Psychiatry* 25 (3): 317–334.

Weber, Max. 1949. "Objectivity" in Social Science and Social Policy. In *Methodology of the Social Sciences*, ed. Edward A. Shils and Henry A. Finch, 50–112. The Free Press: New York.

Windleband, William. 2015. History and Natural Science. In *The Neo-Kantian Reader*, ed. Sebastian Luft, 287–298. London: Routledge.

8

Conclusion

Throughout this book I have challenged a common attack on categorical, polythetic psychiatric diagnoses which rejects their scientific legitimacy. The view that I challenge associates scientific legitimacy with mind-independent reality. It is often assumed that a scientific concept needs to describe something in mind-independent reality for that concept to be scientifically legitimate. These mind-independent entities are taken to be the basis of science. Scientists should seek out those mind-independent entities and base their beliefs upon what they discover about those objects. Doing this turns science from mere opinion or speculation into something that we should consider knowledge. It means that science does not simply describe the opinions of scientists but actually describes how the world is. Additionally, this is said to provide justification to important parts of science such as inductions and causal claims. The reason why we should trust the inductive and causal claims that scientists postulate is that they describe actual states of affairs of mind-independent reality. Since there is good reason to believe that psychiatric diagnoses do not describe anything in mind-independent reality this then leads many to claim they should be rejected.

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5_8

This approach has been a very influential way to understand psychiatric diagnoses. This is most evident within the notions propagated by the 1970s neo-Kraepelians and adopted by the 1980 DSM-III. The real, validated psychiatric diagnoses were portrayed as being disease entities whereby each psychiatric diagnosis had its own cause or causes that were present in everyone who meets the diagnostic criteria and not present in anyone who did not meet the diagnostic criteria. Only those psychiatric diagnoses which had been validated upon such sensitive and specific causes would be worthy of belief or scientifically legitimate. Those causes would determine which entities exist and the opinions of psychiatrists would not influence which entities exist. Additionally, we would have reason to believe in the inductive and causal claims of scientists because the natural similarities of the entity make those inductive and causal claims true. Measured by these standards, we should not see currently employed DSM psychiatric diagnoses as being real. They are far too heterogeneous in symptoms covered and underlying causes. Following this, interest has been growing in alternatives to the DSM like RDoC, HiTOP and person centred approaches. These approaches appear to take an alternative view of reality, whereby the domain of psychiatry actually exists in a dimensional form or exists as individual people. Appeals to reality suggest we need to develop an alternative to the DSM.

In this book, I have defended the use of categorical, polythetic psychiatric diagnoses by rejecting the central assumption of their critics. Critics assume that science should describe mind-independent entities but by drawing upon a neo-Kantian approach I have rejected that assumption. Such mind-independent entities would be things in themselves which Kant describes as "nothing for us" (Kant 2002, 4:336). To a neo-Kantian, science cannot describe such mind-independent entities. For a neo-Kantian, scientific entities are constructed through taking what comes to us from the world and applying concepts and principles to it. Constructing scientific entities requires multiple transformations. This means that scientific entities are neither what we experience nor mind-independent entities. Rather, they have been constructed by us through concepts and principles. As such, we should not judge scientific entities on whether they are or are not mind-independent entities. Instead, they should be

judged upon how well they conform to the concepts and principles that come from us.

I have argued that we should see categories, dimensions and person centred approaches not as how things are in reality but as forms we apply to the world. We can constitute the world as categorical (present or absent), dimensional (present to a degree) or as individuals (not part of a group). Additionally, we can always place further categorical and dimensional divides upon any grouping we have. Having constituted the world as categorical or dimensional, we can then add further categorical or dimensional divides. Additionally, at any point we can stop considering wider groupings and instead focus upon individuals. All these three approaches make possible different knowledge, whereby the inductive and causal probabilities change depending upon what combination of categories or dimensions are employed, or depending upon whether we employ a person centred approach.

When deciding upon which specific categories and dimensions to employ we need to employ principles. I have outlined the neo-Kantian notion that we should aim to build a system of hypothetical entities that aim to cover as much phenomena as possible whilst being constrained by employing as few entities as possible. In doing so, we might do a statistical analysis of a domain and find that categories or dimensions result. Having done this, we might then still desire to add further categorical or dimensional divides to constitute more specific inductive and causal probabilities. Further divisions are made but these then better accommodate more specific descriptions of probabilities between phenomena into the system.

By rejecting the assumption that scientific entities should describe mind-independent entities I can then undermine many critics of psychiatric diagnoses. The numerous, strong arguments which critics have produced which show that psychiatric diagnoses are not mind-independent entities do not entail that they are scientifically illegitimate. With this in mind we can now return to the concerns of critics which I outlined in the introduction.

Firstly, there is the claim that real causes produce real entities and that two instances of an entity are instances of the same entity if they have the same causes. The causal heterogeneity of currently employed psychiatric

diagnoses means that they cannot be real entities and so are not scientifically legitimate. In response I have argued that stance is untenable if it relies upon scientific entities reflecting mind-independent entities. Those would be the things in themselves which no scientific theory can describe. An alternative standard is required when judging which causes and entities are scientifically legitimate. I have outlined how neo-Kantians employ principles of reason to demarcate entities and causes. What constitutes a scientifically legitimate cause or entity are those which contribute towards a system of science that employs as few notions as possible whilst also describing things in as much detail as possible. When doing this we need to consider how causes relate to psychiatric diagnoses, attempting to balance the desire that each psychiatric diagnosis is causally homogeneous, with the desire that the smallest number of causes and diagnoses are employed as possible. Arguably, many currently employed psychiatric diagnoses would not fit such a system. However, my argument shows that it is in principle possible to employ scientifically legitimate categorical, polythetic psychiatric diagnoses even if they are not causally homogeneous. Additionally, I have provided autism as an example of a diagnosis which helpfully contributes to a system of science.

Secondly, there is the concern that psychiatric diagnoses are the product of a decision making process. Critics believe that scientists should discover what entities there are rather than decide what entities there are. Since it is well documented that currently employed psychiatric diagnoses are the product of a decision making process it seems they cannot be real or scientifically legitimate. However, the neo-Kantian approach I have outlined shows that decision making is needed in all sciences. Decisions are needed over what concepts to employ, and how to employ them. Decisions also need to be made over how to balance principles used when building a system of science.

How then should we weigh up this trade-off when deciding what psychiatric diagnoses to employ? As outlined in Chap. 4, I do not believe that a neo-Kantian position entails one answer to this question. It is potentially possible that factor analysis or cluster analysis could be seen to entail one correct set of entities providing we had all possible data and employed those methodologies in the best way possible. However, I do not believe this would be the case even in principle since various

judgements need to be made about what to idealise and abstract, and how to apply these methodologies. As such, rather than seeing the neo-Kantian approach I have developed as being able to produce the correct psychiatric diagnoses, it instead outlines constraints upon what is a legitimate psychiatric diagnosis. A psychiatric diagnosis is legitimate when it contributes to a system of science which takes a form in line with the principles that neo-Kantians value. Multiple systems, positing different psychiatric diagnoses, can be considered scientifically legitimate but not all psychiatric diagnoses or systems of psychiatric diagnoses should be considered to follow from attempts to balance principles that neo-Kantians endorse.

Thirdly, critics believe that psychiatric diagnoses should describe the world and so cannot miss out significant details and still be real or scientifically legitimate. Since psychiatric diagnoses fail to describe important aspects of people like past life history, present life circumstances and future life goals they are not describing the world so are not real and are not scientifically legitimate. However, I have shown that scientific entities are the product of a series of transformations. This means what we experience is transformed, and in doing so much of what we experience is abstracted away. The particular individual that we experience (after we have conceptually constituted it) contains specific details which are not present in the more abstract thought entity. As such, there is no requirement that the abstract thought entity covers all scientifically relevant aspects of particular instances of the thought entity. Psychiatric diagnoses can be scientifically legitimate despite being distant from many aspects of individuals. Additionally, given the abstract nature of the thought entity, there is good reason to investigate particular aspects of instances of thought entities which are not covered by the thought entity itself. There is thus good reason to investigate the aspects of diagnosed individuals which are not covered by the psychiatric diagnosis.

In this book I have countered important criticisms of categorical psychiatric diagnoses. I have argued that in principle categorical, psychiatric diagnoses can be scientifically legitimate in-spite of those criticisms. Additionally, I have outlined criteria upon which categorical, psychiatric diagnoses should be formulated, indicating what a good DSM style psychiatric diagnosis would look like. I have also argued that dimensional

and person centred approaches should be employed. Whilst psychiatry is moving in the direction of these, I have argued that categorical psychiatric diagnoses can still play an important role. I have argued that making the DSM dimensional would be an improvement compared to current approaches but I also argued that adding categorical divides to those DSM dimensions would likely constitute important inductive and causal probabilities. Additionally, I have considered the more radical approach of RDoC and HiTOP which aims to dispense with DSM diagnoses. I think these projects broadly fit the neo-Kantian philosophy I have developed. I am hopeful that they will produce dimensional diagnoses that help better account for the diversity present in the domain of psychiatry. I think that dimensional diagnoses produced by RDoC and HiTOP would have a good chance of being superior to the dimensional diagnoses in a version of the DSM 5 which was dimensional. I have, however, argued that there would likely still be good reason to put categories onto the dimensional diagnoses which RDoC and HiTOP eventually produce. Finally, by portraying both categorical and dimensional diagnoses as idealisations and abstractions which are very distant from actual people I have given good reason to employ person centred approaches. Whilst I believe person centred approaches are still abstract and idealised, they do contain more specific details of particular people that should be sought. As such, whilst I broadly think RDoC and HiTOP are a good means of implementing the neo-Kantian philosophy I advocate for, and I think the neo-Kantian philosophy also gives good reason to employ person centred approaches, I believe categorical psychiatric diagnoses still have a role to play in a future, superior version of psychiatry.

Reference

Kant, Immanuel. 2002. *Theoretical Philosophy After 1781* (Kant, Immanuel, 1724–1804. Works. English 1992). Ed. and Trans. Henry Allison and Peter Heath, Trans. Gary Hatfield and Michael Friedman. Cambridge: Cambridge University Press.

Index[1]

[1] Note: Page numbers followed by 'n' refer to notes.

S. Fellowes, *In Defence of Psychiatric Diagnoses*,
https://doi.org/10.1007/978-3-031-74478-5