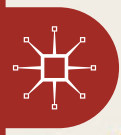

PALGRAVE STUDIES IN THE THEORY AND HISTORY OF PSYCHOLOGY

Embodied, Embedded, and Enactive Psychopathology

Reimagining Mental Disorder

Kristopher Nielsen

palgrave
macmillan

Palgrave Studies in the Theory and History of Psychology

Series Editor
Thomas Teo
Department of Psychology
York University
Toronto, ON, Canada

Palgrave Studies in the Theory and History of Psychology publishes scholarly books that draw on critical histories and theoretical concepts and methods, from a variety of approaches in the psychological humanities, to examine the discipline, profession, and practice of psychology.

This series publishes scholarly books that use historical and theoretical methods to critically examine the historical development and contemporary status of psychological concepts, methods, research, theories, and interventions. Books in this series are characterised by one, or a combination of, the following: (a) an emphasis on the concrete particulars of psychologists' scientific and professional practices, together with a critical examination of the assumptions that attend their use; (b) expanding the horizon of the discipline to include more interdisciplinary and transdisciplinary work performed by researchers and practitioners inside and outside of the discipline, increasing the knowledge created by the psychological humanities; (c) "doing justice" to the persons, communities, marginalized and oppressed people, or to academic ideas such as science or objectivity, or to critical concepts such social justice, resistance, agency, power, and democratic research. These examinations are anchored in clear, accessible descriptions of what psychologists do and believe about their activities. All the books in the series share the aim of advancing the scientific and professional practices of psychology and psychologists, even as they offer probing and detailed questioning and critical reconstructions of these practices. The series welcomes proposals for edited and authored works, in the form of full-length or short monographs; contact beth.farrow@palgrave.com for further information.

Kristopher Nielsen

Embodied, Embedded, and Enactive Psychopathology

Reimagining Mental Disorder

palgrave
macmillan

Kristopher Nielsen
Te Kura Mātai Hinengaro - School of Psychology
Te Herenga Waka - Victoria University of Wellington
Wellington, New Zealand

ISSN 2946-2452 ISSN 2946-2460 (electronic)
Palgrave Studies in the Theory and History of Psychology
ISBN 978-3-031-29163-0 ISBN 978-3-031-29164-7 (eBook)
https://doi.org/10.1007/978-3-031-29164-7

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

Acknowledgments

Thank you to the community who supported me during this project. Particular thanks to: Tony Ward, Daniel Wegerhoff, Emma Ashcroft, Róisín Whelan, and Alice Leader. Thank you also to Georgia O'Grady for keeping me humble.

Praise for *Embodied, Embedded, and Enactive Psychopathology*

"This is a highly topical book that convincingly applies the new embodied and enactive approaches to mental disorders. Instead of an internalist, reductionist view, the author understands such disorders as complex interactions of subject and environment; thus, from the subject's perspective, they appear not only as suffering but also as doing—an important prerequisite for changing unfavorable interaction patterns in therapy. I strongly recommend this book to all mental health professionals who are seeking a forward-looking way to understand mental illness."

—Thomas Fuchs, *Karl Jaspers Professor for Philosophy and Psychiatry, University of Heidelberg, Germany*

"Nielsen uses the enactive toolkit to develop an original framework of psychopathology that emphasizes organizational causality, naturalized normativity, contextuality, complex dynamic networks, and explanatory pluralism. His ability to weave together the natural and the social into a coherent whole, helps dissolve deeply entrenched binaries that have held the psy-sciences back. 3E psychopathology represents the avant-garde in our philosophical understanding of mental disorders and will be a source of explanatory insights and new research questions for years to come."

—Awais Aftab, *Clinical Assistant Professor of Psychiatry, Case Western Reserve University, USA*

"In this book Kristopher Nielsen, a clinical psychologist dissatisfied with current accounts of what mental disorders are, undertakes an ambitious and systematic probe of a vast literature in philosophy of mind, science, and psychiatry to locate a set of conceptual resources sufficient to ground an adequate understanding of mental disorders. From the vantage point of the innovative enactive framework that Nielsen constructs, human organisms are dynamic multi-scaled systems embedded in physical-socio-cultural environments that function well when they can fulfill their needs by adapting to changing circumstances within themselves and their environments. Mental disorders, then, arise in instances in which attempts to self-maintain in the face of a changing environment falter, thereby thwarting need fulfillment and culminating in dysfunction. The implications of Nielsen's conceptual framework are thus considerable; understanding mental

disorders, identifying their cause(s), and developing effective interventions for treating them, importantly and fundamentally requires a holistic multi-scale examination of an individual's brain-body-environment system. Moreover, achieving such conceptual, explanatory, and therapeutic goals is not possible within the confines of a single scientific or humanistic discipline; rather, it requires practitioners from many different disciplines to coordinate their efforts. Nielsen's novel and engaging contribution is thus poised not only to revolutionize how mental disorders are understood, explained, and treated, but also how practices surrounding mental health research and clinical treatment are structured. The book is a must-read for philosophers, mental health researchers, and clinicians, and will be of interest to all whose lives have been touched by mental illness."

—Jacqueline Sullivan, *Associate Professor, University of Western Ontario, Canada*

"Trying to comprehend the complexity of psychopathology is daunting task. A strength of the 3E worldview is that it presents an easy to navigate scaffolding that makes the complexity of psychopathology more manageable and as a writer Kristopher Nielsen is 3e psychopathology's clearest expositor. His concern with how 3E psychopathology can be used in actual clinical situations is an extra bonus."

—Peter Zachar, *Ida Belle Young Professor of Psychology, Auburn University Montgomery, USA*

Contents

List of Figures

List of Tables

1

Conceptualization as a Core Task of Psychopathology Research

Even if not affected ourselves, the vast majority of us will know someone who carries the weight of a mental disorder with them. Clearly, what we currently call mental disorders demand the development of effective treatments—as well as preventative and management strategies—as soon as we are able. But what exactly is a mental disorder, and what does it mean to have one? This is the central question addressed by this book. In particular, this book seeks to express a new way of thinking about mental disorder grounded in something called 'an embodied, embedded, and enactive view of human functioning' that I call '3e Psychopathology'—but I will return to the particulars later. For now, it is important to note that, while scientific in intent, such work represents a conceptual or even philosophical endeavor. In this opening chapter I will therefore first consider why such conceptual and philosophical work is a necessary and important part of the sciences of psychopathology in the first place and highlight some wider commitments. I will then example some general conceptual questions to spark the reader's conceptual intuitions, and briefly orient the reader to the wider intentions of this work. Finally, I will outline the structure of the book chapter by chapter.

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7_1

1.1 The Task of Conceptualization

In order to develop effective *treatments* for mental disorders, we should ideally be working from a good understanding of how exactly they emerge and persist; i.e., we must have good *explanations* of mental disorders. Further, due to their complexity, developing such explanations necessitates reasonably coordinated action by researchers around the globe. Before it will be possible to explain mental disorders effectively then, there will likely need to be common sets of labels and concepts that ensure that researchers are seeking to explain the same things, i.e., we must have ways of *classifying* mental disorders[1]. These three tasks—classification, explanation, and treatment—are often seen as the three core tasks of psychopathology research.

The task of *classification* is concerned with finding some degree of order in the tangled and complex range of behaviors and experiences that appear to be 'disordered', so that we may diagnose and study them effectively (Berenbaum, 2013; L. A. Clark et al., 2017; Zachar & Kendler, 2017). The current dominant classification systems are the Diagnostic and Statistical Manual of Mental Disorders (DSM), currently in its fifth edition[2] and the International Classification of Diseases (ICD), currently in its eleventh edition. These two systems are very similar and I will generally refer to them together as DSM-ICD. The classification of psychopathology is currently at a conceptual crossroads. It is increasingly becoming accepted that fundamental flaws in the DSM-ICD are resulting in it struggling to pick out 'real' mental disorders as opposed to artificially selected clusters of symptoms (Lilienfeld & Treadway, 2016; Zachar & Kendler, 2017). Alternative classification systems are being developed, and the DSM-ICD's continued positioning as the bedrock documents of psychopathology is in serious doubt (Casey et al., 2013;

[1] Classification may not necessarily involve developing a typology of diagnoses ('diagnostic kinds'). There are current arguments for shifting away from diagnostic kinds all together and focusing on a wider set of 'psychiatric' kinds and their complex relations. The Research Domain Criteria [RDoC] represents one such shift that will be discussed in this book, but there are other flavors to this shift away from diagnostic kinds. See Tabb (2016) for a review.

[2] The International Classification of Diseases [ICD] largely parallels the DSM's content, but using a prototype model of description rather than a list of criteria.

Cuthbert, 2014; Insel et al., 2010). Theoretical work within the field of classification is currently asking important questions such as: should our diagnostic systems simply give labels to patterns of signs and symptoms, or try to map onto the causal structures underlying disorder?; should our diagnostic systems attempt to be theoretically neutral, or be open and honest with their theoretical commitments?; are mental disorders simply brain disorders? (Banner, 2013; Insel & Cuthbert, 2015): and, how should our diagnostic systems be responsive to their political and social purposes outside of diagnosis and research (Zachar, 2018)?

The task of *explanation* meanwhile, is concerned with the postulation and validation of theories that make the behaviors and experiences observed in mental disorders less surprising and more comprehendible (Haig, 2014). Whether grounded in neuroscience, psychology, or some other discipline, good explanations of mental disorder point to opportunities for treatment by tracking factors that either *cause* or *maintain* mental disorder. Current and historic attempts at explanation in psychopathology have resulted in limited success. To illustrate this point very briefly, compare current understanding of the causal processes involved in bio-medical illnesses such as the flu or cancers, to prototypical mental disorders such as depression and schizophrenia. We may not have 'cures' for any of these problems, but at least within the bio-medical examples we have some clear ideas about what is going on—i.e., infection by an influenza virus and the uncontrolled division of cells due to mechanisms that are coming to be understood. Comparatively, almost all mental disorders lack agreed-upon causal structures. Aside from the development of actual explanatory theory, (meta-)theoretical work in the area of explanation and philosophy of science more broadly is currently asking questions such as: what are the role of 'mechanisms' in explanations of mental disorder (Glennan & Illari, 2017; Hartner & Theurer, 2018; Thomas & Sharp, 2019)?; what exactly should we be trying to explain—i.e., disorder syndromes, symptoms, brain malfunctions, clinical phenomena, functional processes, or something else entirely (Elbau et al., 2019; Hawkins-Elder & Ward, 2019; Insel et al., 2010; Nielsen & Ward, 2020b; T. Ward & Clack, 2019)?; are detailed explanations always better than general ones (Craver & Kaplan, 2018; Potochnik, 2016,

2017)?; and most generally, how might we go about explaining things as complex and unknown as mental disorders (Insel et al., 2010; Kendler, 2008, 2012; Murphy, 2017)?

Finally, the task of *treatment* involves the development and validation of efficacious interventions for mental disorder; either pharmacological, psychotherapeutic, or through some other means. The current project focuses less on this task, although I will touch on it in the latter half of Chap. 6 where I explore how clinicians develop individualized explanations or 'formulations' for their patients' difficulties. Aside from questions regarding the efficacy of various treatments, some of the big questions regarding the task of treatment currently include: the respective roles of diagnosis versus formulation in targeting treatment (Eells, 2015; Johnstone, 2018); how to utilize the rapidly growing research base to guide treatment choices effectively when our diagnostic concepts are so limited?; how does psychotherapy actually work? (Baier et al., 2020; Kendall et al., 2016; Kraemer et al., 2002); what is the therapeutic mechanism of selective serotonin reuptake inhibitors [SSRIs]?; is the aim of treatment to reduce simply symptoms or should it be something more?. A core idea underlying this project is that developing targeted and efficacious treatments, as well as improving on current treatment approaches, will be much easier when the earlier tasks have been performed well. Well-considered classification systems and valid explanations will provide a strong foundation for the task of treatment.

One of the founding observations of the current project is that this three-task model of psychopathology is incomplete. The elementary yet missing question seems to be: What is mental disorder in the first place? Before we can classify mental disorders—or explain and treat them—we must have some concept of what counts as a mental disorder, why this is the case, and what sort of things they are. What we take mental disorder to be, either explicitly or implicitly, directly informs how we go about the tasks of classification, explanation, and treatment. Our understanding of the nature of mental disorders is a metaphysical commitment (read: 'educated starting guess') that will bias our epistemological strategies, such as how we go about designing studies and reasoning about their findings (Hochstein, 2019). Conceptual links between the nature of mental

disorder and the tasks of classification and explanation mean that elucidating the nature of mental disorder will likely help address some of the mentioned questions currently plaguing these areas. There is, therefore, a need to bring our understandings to the surface and study them directly, so that we may be aware of their biases. The task of *conceptualization* then, while sometimes taken as merely part of the task of classification, is better thought of as its own endeavor (see Fig. 1.1). It is primarily within this task of conceptualization that the current project is situated.

To be clear, the idea of focusing on the task of conceptualization is not itself novel. Much previous work has been done in this area, particularly with the development of conceptual models which I will review in Chap. 2. The claim I am making is that the value of conceptual work, as well as the pressing need for its continued development, is not sufficiently recognized in mainstream psychopathology. Encouragingly, recognition of the need for good conceptual work does seem to be slowly on the rise, as

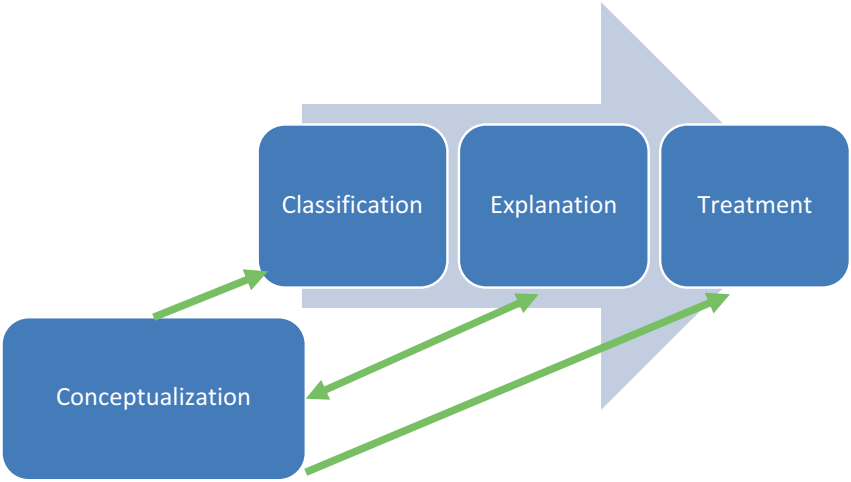

Fig. 1.1 A Four-Task Model of Psychopathology Research. Each task is represented by the boxes. This project is concerned with the task of conceptualization (left), and its implications for later tasks. Conceptualization of mental disorder provides foundational assumptions for our efforts in classification, explanation, and treatment

seen in the emergence of 'philosophy of psychiatry' as an inter-disciplinary field over the last few decades (Fulford et al., 2013; Radden, 2006; Tekin & Bluhm, 2019).

1.2 Some General Questions to Get Started

When people hear the term 'mental disorder', most seem to confidently assume that they know what this term means. Or perhaps, if we can't quite find the words, we can surely have confidence that the relevant scientists and clinicians have a clear conceptualization. A concern underlying the current project is that this confidence may be somewhat misplaced. In Chap. 2, I will review common understandings of what is meant by the term 'mental disorder' and evidence that, while all of the varying views have their strengths, so too do they feature inconsistencies and weaknesses. However, to briefly motivate recognition of the importance of conceptual work upfront (and before things get too technical), it is useful to consider some fundamental questions framed in everyday language that those doing such conceptual work may try to answer. It is not my intention to provide answers here. Instead, I will return to these questions in the closing chapter.

Firstly, are mental disorders something you *get* or something you *do*? In other words, does somebody 'have' depression or are they themselves depressed? This question is important because it has direct implications for how society and individuals, respond to someone experiencing/having/enacting a mental disorder. If a mental disorder is a disease or a lesion in someone's brain, the afflicted person is considered to have little control over it. It also then seems like the sort of thing that might be treated with medication. If, however, a mental disorder is something people do, the afflicted person is considered to have more control over their actions (Kvaale et al., 2013). They may therefore be able to *learn to do things differently*, i.e., it is the sort of thing that might be treated with therapy. Thus, on such a view there is a greater perception of agency, but also responsibility for actions taken.

Next, does a mental disorder exist inside someone's brain or is it dispersed across their brain, body and environment? For example, imagine

someone is working in stressful conditions, and that this stress is maintaining 'symptoms' of depression and anxiety. If they are taken out of this workplace, they may no longer be depressed and anxious. This raises the question, Were they disordered or was their environment dysfunctional?

One final question to example the need for conceptual work in psychopathology: are mental disorders defined by brute facts or by social norms and values? In the 1960s and '70s, psychiatrist and philosopher Thomas Szasz famously made the claim that mental disorder was a myth (Szasz, 1960, 1963, 1974). By this he meant that genuine 'disorder' is, by definition, medico-physical, thus leaving no space for disorders that are purely 'mental'. Rather, disorders with no physical basis are, according to Szasz, simply 'problems in living' and their medicalization a fantasy. For Szasz, this begged questions as to the function of this myth in society. Optimistically he considered whether this medicalization helped society believe in a naturally ordered state of life; one where significant problems-in-living are aberrations rather than the norm. Others in the so-called anti-psychiatry movement however took a more pessimistic view, arguing that mental disorders are simply constructed labels for people that don't follow the unspoken rules of society, and viewing psychiatry as society's tool for dismissing those that refuse to conform (Foucault, 2003/1961). If, however, mental disorders are not based on social norms and values, instead picking out 'real' states or entities in the world, what exactly are they? Are genuine mental disorders required to be diseases or brain abnormalities, or may they be a different kind of thing entirely?

These are just some of the more general questions one might ask when considering the nature of mental disorder. I have placed them here to demonstrate to those that are not familiar with debates in this area that answers to the question 'what is mental disorder?' are still hotly debated. The current project is positioned in response to the need for better answers to this question. It will also explore links to the later tasks of classification, explanation, and treatment. In accordance with this purpose, the aim is to develop and argue for a novel concept of mental disorder and to explore ramifications for the later tasks.

1.3 A Wider Commitment to Pluralism

Before going any further, I want to briefly voice a commitment to pluralism within each of these four highlighted tasks of psychopathological science. Whatever we think mental disorders are, it is clear that they are complex sorts of things. As such there are likely many different ways to think about them that will provide useful insight. Similarly, there are very likely different ways to classify, explain, and treat mental disorders. I am therefore committed to: *conceptual pluralism* (the view that there are multiple valid ways to think about mental disorder), *classificatory pluralism* (the view that there are likely multiple valid ways to divide up mental disorders/targets of enquiry), *explanatory pluralism* (the view that, for any given mental disorder, having multiple explanations or models will add to our understanding), as well as more obviously *methodological* and *treatment pluralism* (the views that we can study and treat mental disorders in various valid ways). This is of course not to say that 'anything goes'. Within each of these tasks and perhaps for different particular mental disorders, different epistemological tools are likely to be more or less helpful. Looking specifically at the conceptual area for example, while there are likely many helpful ways to think about mental disorders, some frameworks will still be more fruitful, justifiable, or useful, than others. As such, we are still justified in generating new ways of thinking about mental disorder, and in arguing about the validity and utility of different frameworks. The spirit of this argument however, should be one of pragmatism and diversity rather than one of 'winner takes all'. Given that the current project seeks to develop such a new way of thinking about mental disorder and argue for its utility, I mention these commitments early to stave off the potential perception of fanaticism.

1.4 Is This a Scientific or Philosophical Project?

Conceptual work, and the current project specifically, is non-explanatory and—while it is responsive to empirical matters—not itself empirical. It is therefore reasonable to consider whether it can truly be considered

scientific in nature, or whether it is a more philosophical enterprise. However, this evokes a much larger question as to the relationship between science and philosophy, and whether a hard distinction between the two is warranted in the first place. While it is not my intention to tackle such a large question here, it will be worth orienting the reader to my perspective on such questions. This will help to clarify why I see such conceptual work as part and parcel of good science.

In short, I see philosophy and science as largely continuous. While philosophy may also be interested in analytical truths and 'big questions', a large part of what philosophy is trying to do is to improve our understandings of the world (Thagard, 2017). This is a purpose aligned with that of science. From my perspective, this more natural side of philosophy is hard to distinguish from science. While each clearly involve different methods and evoke different images in our minds (e.g., a microscope versus armchairs and rigorous argument), the purposes of both science and natural philosophy are to explain the world around us. It therefore makes just as much sense to me to consider science a branch of natural philosophy that has particularly well-developed empirical muscles, as it does to consider natural philosophy a kind of science with particularly well developed analytical/theoretical muscles.

This then leads to wider consideration of the role of a 'theoretical scientist'. I see myself as a 'theoretical psychopathologist', and such theoretical roles are not well instantiated in the discipline of psychology. It is worth noting however that we do not bat an eye when we hear someone referred to as a 'theoretical physicist'. What is the role of a theoretical scientist if they are not manning the microscopes and gathering empirical facts about the world? In brief, I see this role as one that sits in the artificial middle space between science and philosophy. For example, a theoretical scientist within psychopathology may variously focus on important tasks as broad as but not limited to:

- Developing explanatory theories.
- Critiquing and comparing explanatory theories.
- Weaving together theories from diverse areas within psychology and neuroscience, or subsuming theories under broader theories, thus performing an integrative role across psychology.

- Engaging with work outside of their domain thus cross-pollinating the field with potentially useful ideas.
- Critiquing and evaluating the concepts used within theories; i.e., are the concepts we are utilizing metaphysically plausible? Are they coherent with what we know about the rest of the world? Are their simpler or more elegant ways to think of things?
- Generating new concepts.
- Linking theory and empirical findings to normative and practical concerns, such as developing practice frameworks for clinicians.

1.5 Conceptualization and Conceptions of Human Functioning

A primary observation that motivated the current project is that the concept of mental disorder that an individual subscribes to tends to track the individual's *conception of human functioning* in general. Put more simply, someone's implicit or explicit understanding of how the human mind works seems to inform their understanding of the human mind as not working properly. This points to an important conceptual co-determinacy between frameworks of human functioning and frameworks of mental disorder. As a slightly contrived example, if I got in a time machine, visited René Descartes, and asked him what mental disorders are, I assume that his answer would be grounded in his dualistic understanding of the mind-body. Perhaps he would suggest that mental disorders represent some sort of mechanistic breakdown in the soul's connection with the body. Indeed in his mediations on first philosophy he alludes to madness as an altered perception of reality caused by 'bilious vapors' affecting proper functioning of the brain.

I will return to further demonstrate this conceptual co-determinacy between what someone understands mental disorder to be and how they understand human functioning in Chap. 2. For now, this observation opens up the question of what happens if we consciously position ourselves within an understanding of human functioning that seems fit for purpose. Rather than considering humans as simply units in an evolutionary process, as brains driving our bodies around like cars, or as leaves

on the wind of social processes, perhaps we should seek to consider human functioning in a richer and more integrative way? If we do so we may come to a more comprehensive understanding of what mental disorder is. This is one of the key underlying ideas that inspired the current project.

1.6 Structure and Argument of This Book

The underlying justification for the current project can be broken down into three key points. Firstly, there is significant room for improvement in the conceptual understanding of mental disorders. This is an important task that underpins our approach to other key tasks in psychopathology, as well as how society understands mental disorder and responds to those that suffer them. Secondly, what we take mental disorder to be is conceptually related to our underlying assumptions about human functioning. In other words, if one understands humans to work in a particular way, then ones understanding of how humans can 'breakdown' is likely related to this. This raises the possibility that some understandings of human functioning might be more useful than others for generating understandings of mental disorder. Thirdly, the philosophical orientation known as *3e cognition* or *enactivism* seems to be a good candidate for this role as a guiding framework of human functioning within psychopathology. I will more fully unpack what exactly I mean by '3e cognition' in Chap. 3. For now it is enough to say that 3e cognition is a perspective on human functioning and 'how the mind works' that allows for the integration of biological, psychological, socio-cultural, and environmental causal factors in a naturalistically plausible way.

In this book, I develop 3e Psychopathology, an understanding of what mental disorder is from the perspective of 3e cognition. I argue that this perspective produces a rich and flexible understanding of mental disorder that can compete well against current popular approaches. Throughout this book, it is *not* my intention or purpose to argue for 3e cognition/enactivism as a philosophy of mind, only its fruitfulness for considering the nature of mental disorder. Breaking this wider argument down by chapter, the book is structured as follows.

In the current chapter I have introduced the general topic area, briefly considered whether this is a philosophical or scientific project, and expressed the need for greater focus on conceptual work in the sciences of psychopathology. In Chap. 2, I will explore current conceptual models of mental disorder and show that, while they all have their strengths, all have room for improvement. These various models are separated into those that address structural questions (i.e., *what* kinds of things are mental disorders and *where* do they exist) and those that answer normative questions (i.e., *why* should something count as disordered or dysfunctional). In Chap. 2, for reasons of space, I have chosen not to review the understanding of mental disorders implicit in major classification systems such as the DSM-ICD, as this has been well discussed in wider literature. If you are interested in my perspective on this then please see Nielsen (2020). At the end of Chap. 2, I argue that there is a need for a broader framework of human functioning in which to situate a concept of mental disorder. To put it simply, if we want to conceptualize *dysfunction* we must first formulate a concept of what it is to be *functional,* or otherwise not disordered.

In Chap. 3, I introduce 3e cognition/enactivism and argue that this position has potential to serve as the broader framework of human functioning needed. I then discuss other recent attempts to consider mental disorder as a wider concept from this 3e view. Specifically I consider the work of Thomas Fuchs (2017), Sanneke de Haan (2020), and Michelle Maiese (2021), considering the strengths and weaknesses of their frameworks.

Chapter 4 is referred to as 'The Bones of 3e Psychopathology'. In this chapter I present the core ideas of my own conceptual understanding of mental disorder. I first expand on several theoretical tools that I believe make 3e cognition a valuable framework of human functioning from which to approach mental disorder. I then present an understanding of the general causal structure of mental disorders under a 3e view, similar to that originally presented in Nielsen and Ward (2018). This is followed by an overview of how I understand that normative basis of mental disorder under the 3e view, similar to that originally presented in Nielsen and Ward (2020a).

Chapter 5 aims to glue the bones together and flesh out the fuller conceptual model within 3e Psychopathology. Drawing on ideas of pluralism and perspectivism (Massimi, 2021; Massimi & McCoy, 2020), I aim to describe the concept using multiple styles of language and some metaphors in order to try to offer a thorough understanding of the model being presented. I analyze the concept using a conceptual taxonomy presented by Zachar and Kendler (2007), offer an analysis of anxiety as an illustrative example, and compare the conceptual model to a selection of models reviewed in Chap. 2. I then address several questions relating to the task of classification.

Chapter 6 addresses implications of 3e Psychopathology for the task of explanation. Chapter 6 is split into two parts. The first part addresses implications for explanation at a research level. I argue that 3e Psychopathology demands a pluralistic approach to research and the development of explanatory theory, and summarize the core ideas from the Relational Analysis of Phenomena (the RAP)—a meta-methodological framework for developing explanatory theories congruent with the principles of 3e Psychopathology. The second part of Chap. 6 then explores how a 3e Psychopathology orientation may influence the way clinicians approach the development of formulations, i.e., individually tailored explanations that guide treatment decisions. This second part of Chap. 6 takes a clinical and practical focus to show how conceptual work can have real and useful impact on day-to-day matters. In this section I occasionally draw on my own experience as a clinical psychologist and present two 3e Psychopathology inspired tools to assist with formulation practices.

Finally, in Chap. 8 I summarize and draw the book to an end. I summarize the book, present a summary example of disordered eating and some of the various ways we can conceptualize this, and highlight some limitations of 3e Psychopathology. I return to briefly summarize how 3e Psychopathology would answer the initial questions posed earlier in this current chapter. I then close the book with an emphasis on the need for continued conceptual refinement if the sciences of psychopathology and its associated clinical approaches are to progress.

References

Baier, A. L., Kline, A. C., & Feeny, N. C. (2020). Therapeutic alliance as a mediator of change: A systematic review and evaluation of research. *Clinical Psychology Review, 82*, 101921.

Banner, N. F. (2013). Mental disorders are not brain disorders. *Journal of Evaluation in Clinical Practice, 19*(3), 509–513.

Berenbaum, H. (2013). Classification and psychopathology research. *Journal of Abnormal Psychology, 122*(3), 894.

Casey, B., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S., & Ressler, K. J. (2013). DSM-5 and RDoC: Progress in psychiatry research? *Nature Reviews Neuroscience, 14*(11), 810.

Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest, 18*(2), 72–145.

Craver, C., & Kaplan, D. M. (2018). Are more details better? On the norms of completeness for mechanistic explanations. *The British Journal for the Philosophy of Science, 1*, 287–319.

Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry, 13*(1), 28–35.

de Haan, S. (2020). *Enactive psychiatry*. Cambridge University Press.

Eells, T. D. (2015). *Psychotherapy case formulation*. American Psychological Association.

Elbau, I. G., Binder, E. B., & Spoormaker, V. I. (2019). Symptoms are not the solution but the problem: Why psychiatric research should focus on processes rather than symptoms. *Behavioral and Brain Sciences, 42*, E7.

Foucault, M. (2003). *Madness and civilization*. Routledge.

Fuchs, T. (2017). *Ecology of the Brain: The phenomenology and biology of the embodied mind*. Oxford University Press.

Fulford, K., Davies, M., Gipps, R., Graham, G., Sadler, J., Stanghellini, G., & Thornton, T. (2013). *The Oxford handbook of philosophy and psychiatry*. OUP Oxford.

Glennan, S., & Illari, P. (2017). *The Routledge handbook of mechanisms and mechanical philosophy*. Taylor & Francis.

Haig, B. D. (2014). *Investigating the psychological world; scientific method in the behavioural sciences*. Massachusetts Institute of Technology.

Hartner, D. F., & Theurer, K. L. (2018). Psychiatry should not seek mechanisms of disorder. *Journal of Theoretical and Philosophical Psychology, 38*, 189–204.

Hawkins-Elder, H., & Ward, T. (2019). Theory construction in the psychopathology domain: A multi-phase approach. *Theory & Psychology*.

Hochstein, E. (2019). How metaphysical commitments shape the study of psychological mechanisms. *Theory & Psychology, 29*(5), 579–600. https://doi.org/10.1177/0959354319860591

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry, 167*, 748–751.

Insel, T., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science, 348*(6234), 499–500.

Johnstone, L. (2018). Psychological formulation as an alternative to psychiatric diagnosis. *Journal of Humanistic Psychology, 58*(1), 30–46.

Kendall, P. C., Cummings, C. M., Villabø, M. A., Narayanan, M. K., Treadwell, K., Birmaher, B., Compton, S., Piacentini, J., Sherrill, J., & Walkup, J. (2016). Mediators of change in the Child/adolescent anxiety multimodal treatment study. *Journal of Consulting and Clinical Psychology, 84*(1), 1.

Kendler, K. (2008). Explanatory models for psychiatric illness. *American Journal of Psychiatry, 165*(6), 695–702.

Kendler, K. (2012). Levels of explanation in psychiatric and substance use disorders: Implications for the development of an etiologically based nosology. *Molecular Psychiatry, 17*(1), 11.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877–883.

Kvaale, E. P., Haslam, N., & Gottdiener, W. H. (2013). The 'side effects' of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review, 33*(6), 782–794.

Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology, 12*, 435–463.

Maiese, M. (2021). An enactivist reconceptualization of the medical model. *Philosophical Psychology, 34*, 1–27.

Massimi, M. (2021). *Perspectival realism*. Oxford University Press.

Massimi, M., & McCoy, C. D. (2020). *Understanding perspectivism: Scientific challenges and methodological prospects*. Taylor & Francis.

Murphy, D. (2017). Can psychiatry refurbish the mind? *Philosophical Explorations, 20*(2), 160–174.

Nielsen, K. (2020). *What is mental disorder? Developing an embodied, embedded, and enactive psychopathology*. PhD thesis, Victoria University of Wellington. http://hdl.handle.net/10063/8957

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 8*(6), 800–822. https://doi.org/10.1177/0959354318808394

Nielsen, K., & Ward, T. (2020a). Mental disorder as both natural and normative: Developing the normative dimension of the 3e conceptual framework for psychopathology. *Journal of Theoretical and Philosophical Psychology, 40*(2), 107–123. https://doi.org/10.1037/teo0000118

Nielsen, K., & Ward, T. (2020b). Phenomena complexes as targets of explanation in psychopathology: The Relational Analysis of Phenomena (RAP) approach. *Theory & Psychology, 30*(2), 164–185.

Potochnik, A. (2016). Scientific explanation: Putting communication first. *Philosophy of Science, 83*(5), 721–732.

Potochnik, A. (2017). *Idealization and the aims of science*. University of Chicago Press.

Radden, J. (2006). *The philosophy of psychiatry: A companion*. Oxford University Press.

Szasz, T. S. (1960). The myth of mental illness. *American Psychologist, 15*(2), 113.

Szasz, T. S. (1963). *Law, liberty, and psychiatry: An inquiry into the social uses of mental health practices*. Syracuse University Press.

Szasz, T. S. (1974). *The myth of mental illness: Foundations of a theory of personal conduct, Rev*. Harper & Row.

Tabb, K. (2016). Philosophy of psychiatry after diagnostic kinds. *Synthese*, 1–19.

Tekin, S., & Bluhm, R. (2019). *the bloomsbury companion to philosophy of psychiatry*. Bloomsbury Publishing.

Thagard, P. (2017). *Natural philosophy: From social brains to knowledge, reality, morality, and beauty* (draft 3).

Thomas, J. G., & Sharp, P. B. (2019). Mechanistic science: A new approach to comprehensive psychopathology research that relates psychological and biological phenomena. *Clinical Psychological Science, 7*(2), 196–215. https://doi.org/10.1177/2167702618810223

Ward, T., & Clack, S. (2019). From symptom to clinical phenomena. *New Ideas in Psychology, 54*, 40–49.

Zachar, P. (2018). Diagnostic nomenclatures in the mental health professions as public policy. *Journal of Humanistic Psychology*. https://doi.org/10.1177/0022167818793002

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164*(4), 557–565.

Zachar, P., & Kendler, K. S. (2017). The philosophy of nosology. *Annual Review of Clinical Psychology, 13*, 49–71.

2

Current Conceptual Models of Mental Disorder

In this chapter I review prominent conceptual models of mental disorder, commenting on their strengths and weaknesses. These are models that provide answers to the question 'what are mental disorders?'. My focus here is on *formal* conceptual models—i.e., those presented as such. I have structured the presentation of these formal views in a way that highlights two different ways that we can understand the question 'what are mental disorders?'. I first present what I refer to as the *structurally oriented concepts*. These concepts focus on the nature of mental disorders in the ontic sense; on what mental disorders are in terms of their physical or causal structure. This is opposed to what I refer to as the *normatively oriented concepts,* which I present next. These normatively oriented concepts focus on why something should be (or should not be) considered a disorder. Please note that I have chosen not to review less formal models in this text. By this I refer to those models presented explicitly or implicitly within frameworks that have boarder non-conceptual purposes such as the DSM-ICD, or the Research Domain Criteria [RDoC]. I have previously presented such reviews in Nielsen (2020) and Nielsen and Ward (2018), but have not included them here as they are less central to the wider and more immediate argument. In closing this chapter, I make

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7_2

some observations that support the use of an embodied, embedded, and enactive view as a framework of human functioning through which to consider mental disorder. A key role of this chapter is to demonstrate that while having a multitude of conceptual models at our disposal is useful (i.e., conceptual pluralism), this does not negate the need for conceptual refinement and the development of better models.

2.1 Structurally Oriented Concepts

Haslam (2002) presents a conceptual taxonomy that usefully organizes differing perspectives on the structural nature of psychopathology. Haslam ultimately argues for a conceptual pluralism, whereby different mental disorders are seen to likely have different structural natures; for example, that borderline personality disorder and bipolar disorder are not just different types of mental disorder, but different *kinds* of types, with the latter being much more homogenous and disease-like, and the former being much more heterogeneous and socially weighted in its etiology. In accordance with this, Haslam sees pragmatic value in the plurality of structural views available, and his taxonomy is intended as a first pass attempt to collate the different kinds in a meta-structural way. He clusters the views under the labels: 'non-kinds/continua' (phenomena that don't form a kind but differ on a single spectrum, e.g., color/wavelength, neuroticism); 'practical kinds' (phenomena that can be clustered together because it is useful to do so, e.g., flying creatures, mood disorders); 'fuzzy kinds' (phenomena that can roughly be clustered together based on similarity even though all the instances aren't the same, e.g., board games, sandwiches); 'discrete kinds' (phenomena with no essences that can still be clearly identified as members or non-members most of the time, e.g., biological males[1]); and 'natural kinds' (phenomena with defined essences,

[1] Biological sex is an arguable case of a discrete kind but is a good illustrative example in that it has no single essence, instead being composed of multiple related components (e.g., xx/xy chromosomes, hormone levels, internal and external physiology) that *tend* to bifurcate into male and female camps in *most* cases. This is not to deny the existence or validity of intersex persons in anyway. One could also argue that biological males or females are examples of fuzzy kinds. I am less convinced that there is truly a clear demarcation between fuzzy and discrete kinds, but I include reference here to stay true to Haslam's taxonomy.

e.g., atomic elements). I will unpack these labels further when discussing them below.

In this section I use an adaption of Haslam's (2002) taxonomy to organize my overview of the structurally oriented conceptual models. The key change I have made is that I have excluded 'practical kinds' from this section, instead discussing them in the following section on normatively oriented concepts. I also give more room to the discussion of fuzzy kinds, as this is a complicated concept which will be important in later chapters. I will further explain the differences between the kinds at the start of each sub-section. Note that all structural models discussed necessarily assume realism about mental disorders[2] (Kendler, 2016). Finally, note that the use of Haslam's taxonomy brings with it a focus on the degree of kindship/homogeneity of the underlying causal structures of mental disorder. This is as opposed to demarcating different conceptual positions by the etiological domains they emphasize (e.g., mental disorders are genetic diseases, neurological conditions, social problems)[3]. Where relevant I therefore point out recognized conceptual positions that are not only committed to a particular degree of homogeneity, but also to the primacy of particular etiological domains (e.g., biological essentialism, biopsychosocial holism).

[2] 'Realism' refers to the view that there are ontic things in the world to which the label 'mental disorder' could refer, that these things, whatever form they take, are 'discovered' and exist independently of our attempts to classify them (i.e., they are not *entirely* socially constructed or pragmatic). I briefly discuss social constructionism and pragmatism in the following section on normatively oriented concepts. Socially constructed kinds could possibly be discussed in this section as, while they are constructed, they still have an ontic reality in the form of a pattern of behavior (Mallon, 2016); for example see the controversial socio-cognitive model of dissociative identity disorder (Gleaves, 1996). I cover social constructionist models in the normative section due to their association with anti-psychiatry.

[3] By discussing two separate ideas/dimensions in proximity I risk conflating them here. The idea of a continuum of homogeneity (simple/essentialist—complex/emergent) and the idea of a 'continuum' of etiological domain (biological-social) are in fact separate ideas that are often conflated (although it is interesting to consider if there is actually a possible relationship between these dimensions). Also note that the idea of particular mental disorders existing at *one place* on a organic-to-social continuum is a strongly criticized idea, mental disorders from schizophrenia to borderline personality are better seen as 'dappled' across this spectrum, each with mechanisms at a variety of scales (Kendler, 2012).

Non-kinds/Continua

Haslam (2002) begins his taxonomy with a category that captures those concepts in psychiatry that *do not* count as kinds, i.e., things that are completely continuous and are therefore *non-kinds* or *continua*. Such concepts are often referred to as dimensional. A good example of a non-kind is neuroticism. There is no non-arbitrary level of neuroticism at which someone counts as 'neurotic' or not, rather people can be more or less neurotic, with no clear 'tipping point' at which one can be labeled. Neuroticism therefore is a case of a pure continuum rather than a kind.

Most concepts utilized across psychology are continuous in a certain sense, and better modelled as dimensional rather than categorical (Haslam et al., 2012; Kotov et al., 2017). This also includes many diagnostic concepts, for example someone can be more or less depressed; depression comes in degrees. However, this level of continuity is subtly different to a non-kind where *no* meaningful point of demarcation or tipping point between members and non-members of the class is assumed to be present. There are few conceptual models of disorder that subscribe to this radical continuity, with most models assuming at least a fuzzy degree of categorical kindship across members of a class. The exceptions to this are some of the *practical kind* models which I will discuss in the section on normatively oriented concepts.

Natural/Essentialist Kinds

Haslam (2002) draws a distinction between *natural kinds* proper and *discrete kinds* (which I will discuss next). Within his taxonomy, natural kinds have a clear common causal structure; a single 'latent variable', or 'essence' underlying them. From philosophy, the classic example of natural kinds in this strict sense are atomic elements which are clearly defined by the number of protons present, for example, gold always has seventy-nine protons while helium always has two. When referring to this kind notion, I prefer to use the term *essentialist kinds*. The reasons for this choice of terminology are multiple. Firstly, my general use of the term 'natural kind' is a lot broader than Haslam's (2002) use. My use of 'natural kind'

refers to a kind concept that picks out something real as opposed to conventional, selecting out a class of things which share properties to the degree that labeling them can be useful for our scientific purposes (i.e., correct application of the label to a thing allows for inductive inference as to other properties that the labeled thing may hold). This conception therefore encompasses both strictly natural and discrete kinds in Haslam's terms[4] (and even many 'fuzzy' kinds). Secondly, there is a lot of controversy over what authors actually mean when the term 'natural kind' is used, with some uses signaling a restrictive essentialist concept as in Haslam's taxonomy, and others a more open concept like my general use of the term (Bird & Tobin, 2018). Finally, sometimes there can be difficulty with the use of the term *natural* kind regarding whether such a concept can encompass social or mental phenomena. Rightly or wrongly, one criterion often discussed concerning natural kindship is that of 'mind independence'[5] (Khalidi, 2013). This is seemingly due to a false dichotomy intuitively drawn between what is 'natural' versus 'human' and can produce some difficulties when studying mental and social phenomena such as mental disorders.

Current conceptual models that propose mental disorders to be essentialist kinds tend to be those that model mental disorders on physical disorders, so called *biological essentialism*. These approaches assume that there are yet to be discovered biological disease processes or abnormalities underlying mental disorders. When uncovered, such biological lesions will reveal that mental disorders are essentially physical disorders (presumably of the brain) that manifest mental and behavioral symptoms. The idea is that revealing these latent biological variables will allow for clear and etiopathologically valid categorization. A structural conceptualization such as this can be implicitly seen in explanatory theories such as the—now highly contested—serotonin hypothesis concerning

[4] My orientation here is parallel to a natural kind position argued for by Boyd (1991) and by Magnus (2014a, 2014b), whereby some, but not all, natural kinds are Mechanistic Property Clusters or MPCs (which will be discussed when covering fuzzy kinds).

[5] Khalidi (2013) offers a discussion of this issue, arguing for a shift away from mind independence as a criterion for natural kindship and toward consideration of whether a kind is categorized together based on causal relation/similarity versus categorized together as a matter of convention. Many social kinds (war, money, racism) can indeed be natural despite their mind dependence.

depression. This theory holds that depression is essentially a dysfunction in the serotonergic systems of the brain (Albert et al., 2012; Gardner & Boles, 2011). More explicitly, such essentialist conceptions can be seen in the work of authors like Insel and Cuthbert (2015), who—on the basis of the success of 'precision medicine' in areas such as oncology, where genotyping and targeting of specific cancer sub-types is becoming more common—argue for the need to make our diagnostic categories more precise. Up until this point Insel and Cuthbert's arguments represent a reasonably consensus view. The essentialist (and theory-reductionist[6]) step these authors take is their next one, where they argue that the only way to achieve such precision is through adopting a biologically focused model of psychiatry; a model in which mental disorders are simply brain disorders with behavioral, cognitive, and emotional symptoms. Implicit in this step is the idea that, when it comes to mental disorders, the brain is where the money is; that there are undiscovered neurological essences to what we label (wrongfully in their mind) *mental* disorders[7]. Notably, biomedical notions of mental disorder seem to be gaining in popularity, both within psychopathology and with lay people (Lebowitz & Appelbaum, 2019).

Biological essentialism is not the only kind of essentialist position one could take in regard to mental disorder. For example, psychoanalytic approaches to the explanation of mental disorder represent an essentialist approach, but with the dominant latent variable being some underlying psychological factor (a 'neurosis'), rooted in past experience. The neurosis here, is in effect acting as a psychological essence and could therefore be termed a form of *psychological essentialism*. To use a more mainstream example, *cognitive models* of psychopathology—those that hold mental disorder to boil down to errors or biases in thinking—can also be understood as examples of psychological essentialism. For example, think of therapists that utilize Cognitive Behavioral Therapy [CBT] with clear

[6] 'Theory-reductionism' is the view that the different domains of science can be reduced to the more 'fundamental' sciences, i.e., that psychology is applied biology, is applied chemistry, is applied physics, is applied math.

[7] Another component of their argument is the need to unclip research efforts from current diagnostic categories. This is a point I agree with and will be covered more in Chaps. 6 and 7 which are more focused on explanation.

emphasis on the cognitive over the behavioral. Such therapists see behavioral interventions only as a tool to shift problematic patterns in cognition (to use a common turn of phrase, they do CBT with a capital 'C' and a small 'b'). Such therapists are implicitly taking a psychological essentialist position. Beck and Bredemeier's (2016) unified cognitive model of depression is a good example of a theory that also falls under this conceptual position. For the most part however, the idea that mental disorders are essentialist kinds tends to co-occur with the idea that the essences in question lie within the brain.

Discrete Kinds

Haslam (2002) uses the term *discrete kinds* to distinguish things that feature clear membership conditions, but that—in contrast to essentialist kinds—are not defined by a single causal factor or essence. Instead, discrete kinds have complex underlying causal structures, but due to the dynamics of the causal structure in context they bifurcate into members and non-members of the kind. Thus, discrete kinds still produce a clear boundary with very few ambiguous cases. Haslam (2002) gives the example of melancholic depression. This is a diagnostic concept, present in the DSM-5 as a sub-type of depression, featuring dominant anhedonia and vegetative symptoms. Haslam cites taxometric evidence that melancholic depression is clearly categorical in nature but notes that this does not necessarily imply the existence of an underlying essence, instead arguing that this may be an example of a discrete kind. This is unfortunately the only diagnostic example Haslam mentions, and the concept of a discrete kind has not, to my knowledge, been picked up by other authors. It is also not clear what categorically separates a discrete kind from an essentialist kind with a particularly complex essence (or alternatively a reasonably homogenous Mechanistic property Cluster [MPC] kind, discussed later). I mention it here as it remains an interesting idea, and to be true to Haslam's taxonomy.

Fuzzy Kinds

Fuzzy kinds are real and objective categories that exist in nature and are thereby very different to non-kinds/continua. However, the point of demarcation between what is and isn't counted as a token of the kind is blurry, or rather 'fuzzy'. Rather than a single tipping point, or 'joint' in nature, that separates members of a fuzzy kind and non-members, there is a *zone of ambiguity*; a gentle curve of demarcation rather than a defined point. Fuzzy kinds then, represent "real, discoverable discontinuities" in the world (Haslam, 2002, p. 208), and are therefore not non-kinds. Fuzzy kinds however, admit to intermediate or borderline cases. As an example, the concept of a 'teddy bear' is meaningful. There are clear cases of objects that are teddy bears such as Mr. Bean's 'Teddy', and there are clear cases of objects that are not teddy bears such as my foot. However, there are also in-between cases such as a soft-toy Koala. Koalas are not proper bears yet are sometimes referred to as such. If I showed a soft-toy Koala to a selection of people, some would categorize it as a teddy bear and some would not. But this does not mean that there is no meaningful difference between teddy bears and other objects. Teddy bears can therefore be said to be fuzzy, not just because of their texture, but because they admit ambiguous membership. It is important to note here that it is not the fact that people have difficulty identifying the members of a kind in itself that makes the kind fuzzy, but rather its *actual* in-between status. I am talking here about ontological fuzziness rather than epistemological fuzziness.

A concept being fuzzy suggests that the causal structures underlying the phenomena referenced by the concept are reasonably complex (Haslam, 2002). If some phenomenon is supported by a single causal factor or 'essence' then its identity tends to be clear-cut (i.e., discrete or essential kinds). For example, a given atom either is an example of gold or is not, depending on a single factor (i.e., the number of protons present). For fuzzy kinds, the existence of borderline cases suggests that more than one 'defining' factor is at play. For example, what counts as a teddy bear is dependent on not just one factor but many: does it have a snout, is it cute, is it squishy, does it have round ears? While 'teddy bear' is still

a meaningful category, soft-toy Koalas also exist with enough of these properties to be meaningfully akin to teddy bears, but to not quite be 'proper' teddy bears. If a mental disorder (e.g., depression) differs meaningfully from both normality and other mental disorders (e.g., anxiety), yet there are messy in-between cases (e.g., anxious-depression, or people who are just a little bit depressed) then the fuzzy kind label may be appropriate[8]. When considering mental disorders this idea seems appealing given that such a messy reality is exactly what we find; i.e., high rates of apparent artefactual co-morbidity and diagnostic ambiguity (Andrews et al., 2002; Lilienfeld & Treadway, 2016).

Given this association with complexity, a position intuitively associated with the idea of a fuzzy kind is the biopsychosocial movement (Bolton & Gillett, 2019; Borrell-Carrió et al., 2004; Engel, 1977). This movement is a broad approach to health and wellbeing, born in reaction to the growing biological reductionism of medicine in the middle of the twentieth century. Originally proposed by Engel (1977), the biopsychosocial movement emphasizes the need for holism, and the need to recognize that mental disorders (and physical disorders) generally arise from, or are influenced by, complex non-linear interactions between multiple factors, and that these factors range across different scales of analysis (from molecular to socio-cultural). The movement also emphasizes a congruent focus on the person above and beyond their disease and genuine care and concern during patient-professional interaction. The biopsychosocial movement then, is anti-reductionistic and encourages broad and agentic considerations. Considering the structure of mental disorder through the biopsychosocial lens may therefore bring certain ethical advantages, perhaps producing a more compassionate psychiatry that is more mindful of the person-as-a-whole, rather than simply the mechanics of their disease processes. Despite the value and importance of this approach however, considering the biopsychosocial movement as a structural model of mental disorder is currently problematic. The only structural commitment this approach really makes is to the general facts that 1) factors across the different scales of analysis are likely relevant, and that

[8] The difficulty here is ruling out other possibilities such as anxious-depression being something different all together, or depression simply being radically continuous (i.e., a non-kind).

2) these factors may interact in complex ways. This is in no doubt true, certainly there is a need to recognize the complexity at hand. The problem here is that, in making no firm commitment to the nature of these interactions above and beyond their complexity, the biopsychosocial movement offers very little guidance for attempts at classification, explanation or treatment, other than to 'look at *all* the things' (Ghaemi, 2009). An exception to this is Bolton and Gillett's biopsychosocial model of health and disease (2019), which seeks to further specify how biological, psychological, and social causes can exist and interact to shape human functioning or ill-health. However, as it stands this model does not present a fleshed-out conception of what it takes mental disorder to be. For further discussion of this model and comparison to the model expressed in this book, see Aftab and Nielsen (2021). In summary, despite how often we may hear it spoken of, it is not clear if there is really such a thing as 'the biopsychosocial model of mental disorder'. Such references are better thought of simply as a call to widen our perspective and consider the complex reality of the phenomena we call mental disorders. How this is to be done and what it means for our concept of mental disorder continues to be under specified. The conceptual product of this book represents one possible step forward.

One structural model of mental disorder that puts the fuzzy kind idea to work with greater specificity is the view that mental disorders are *Mechanistic Property Clusters* or 'MPC kinds'[9]. This model was applied to mental disorder by Kendler, Zachar and Craver (2011), building upon the philosophical work of Boyd (1991). MPC kinds are constituted by clusters of properties held together or caused by a mutually reinforcing *network of mechanisms*. For example, the kind 'sheep', in being a biological species, is often assumed to be a meaningful and categorical kind. But what makes a sheep a sheep? Well, for one, sheep are wooly, and have four legs. One problem with this answer is that if I have a three-legged sheep and shave it bare, it still seems like this poor creature, no matter its condition, is still a sheep in a meaningful sense. The properties of being wooly and having four legs then, don't seem to be the 'essence' of what it means

[9] Following Boyd (1991), the philosophical terminology is homeostatic property cluster (HPC), but here I use Kendler et al.'s label (MPCs) as this is conventional in the psychopathology literature.

to be a sheep. Boyd's answer to this problem was to change tack; not to look for the 'essence' of the sheep—the 'necessary and sufficient conditions' that define a sheep—but rather to propose that what makes a sheep a sheep is the fact that all sheep share an evolutionary lineage, representing overlap in the causal structures that led to any one sheep's existence. A slightly different example, given by Magnus (2012, 2014a, 2014b), would be pools of water. Pools of water do not necessarily share a causal lineage, e.g., a pool of water may form here on earth, as well as on a completely different planet. However, a very similar causal process underlies their formation (e.g., an affinity between H_2O molecules due to their dipole structure, processes of condensation, some process of containment). The mechanism (or set of mechanisms) that leads to the formation of such pools is the same or features significant similarity. Cases such as these are referred to as *type-causal* MPCs because the underlying causal pattern occurs multiple times; it is a 'type' of causal pattern that leads to members of the kind sharing properties. The previous example of a biological taxon (a sheep) is referred to as a *token-causal* MPC because there is a single causal cascade (in this case an evolutionary history) shared by all members and leading to their overlapping properties (Magnus, 2012, 2014a, 2014b).

On this MPC view then, mental disorders are fuzzy sets of properties (i.e., properties of people, presumably signs and symptoms) and a network of causal mechanisms that holds these properties together in a wider possibility space (Kendler et al., 2011). This causal network may consist of the symptoms themselves, as well as underlying states and processes. Importantly, the factors playing a role in this causal network may cross boundaries of scale—evolutionary, physiological, psychological, social, etc.—with no *a priori* privilege given (Kendler et al., 2011). Kendler et al. also highlight the flexibility of this position, leaving room for more or less homogenous MPC kinds:

> "In the limit of simplicity and determinacy, MPCs tend toward essences, with properties and mechanisms common to all and only members of the kind. At the other extreme, cluster kinds tend toward constructed or practical kinds, where the boundaries of categories are often defined with

respect to the classificatory practices of some interested party." (Kendler et al., 2011, p. 1146)

Note that more homogenous MPC kinds would likely be captured by Haslam's concept of a discrete kind (Haslam, 2002). The MPC concept is therefore very flexible in its reference.

The MPC view is currently popular when considering the structural nature of mental disorder. It offers a possible reason why no dominating causal factors or clearly defined causal networks underlying any modern mental disorders have been found. Mirroring the study of physical disorder and disease, it has been historically assumed that the discovery of such 'essences' is the ultimate goal of psychopathology research. The MPC view and other such 'fuzzy' models suggest that maybe the reason we are failing to find such essences is that they simply may not exist. Fuzzy models allow us to consider this without giving up on kindship altogether, instead suggesting that mental disorders may be different to many physical disorders, not just because they concern behavior and 'the mind', but because of their complexity. In other words, that they may be heterogeneous categories with no definable essence but that meaningful and useful patterns can still be found. The major issue facing the MPC and other fuzzy views is parallel to that faced by the biopsychosocial approach. If we recognize this degree of complexity, where do we start? Will some scales of analysis be more useful than others? Which mechanisms should be focused on? Despite being more specified than the biopsychosocial approach, the MPC view still does not offer much *guidance* in this respect. As will be seen in later chapters, the concept of mental disorder developed in this book is structurally very similar to an MPC view, while placing issues of complexity and normativity much more in the foreground. The perspective developed attempts to address this issue with guidance, not by prioritizing any scale of analysis *a priori*, but through consideration of the normative dimension of mental disorder and its intersection with the structural.

Before moving on, one currently popular idea that attempts to put the notion of an MPC to work is that of the Symptom Network Model of mental disorders (SNWM). The SNWM approach assumes that many mental disorders are best understood as *networks of symptoms,* which can

be statistically modeled. Symptoms within these networks are hypothesized to cause each other, with recursive feedback resulting in the relative stability of the network over time (Borsboom et al., 2018; Cramer et al., 2010; McNally, 2016). Recent years have seen a significant increase in SNWM research, with many examples being used successfully in empirical studies (Fried et al., 2017). This approach is presented by its proponents as a radically new way of conceptualizing psychopathology; as a model of mental disorder that rejects the search for underlying cause/s of psychopathology, i.e., the essentialist or latent variable model (Borsboom et al., 2018). However, there is considerable debate over whether this is the case, or whether SNWM is simply a new and promising measurement tool that tracks statistical relationships between symptoms (Bringmann & Eronen, 2018; Epskamp et al., 2017; Fried & Cramer, 2017; Haig & Vertue, 2010; Humphry & McGrane, 2010; Molenaar, 2010; T. Ward & Fischer, 2019). These concerns seem warranted, especially given that, conceptually, the SNWM seems very much like an MPC model that restricts itself to the level/scale of signs and symptoms. I will now shift to overviewing a selection of normative conceptual models.

2.2 Normatively Oriented Concepts

The conceptual models covered in this section focus on *why* something should be considered a mental disorder and are mostly not covered by Haslam's (2002) taxonomy as this was oriented predominantly towards structural concepts. Another way to think of these normatively oriented models is that they try to provide understandings of mental disorder with 'conceptual validity' (Wakefield, 2014b). Conceptual validity refers to the ability of a concept or framework to correctly distinguish between 'normal' functioning on one side and *dis*order, *dys*function, or pathology, on the other[10]. The use of 'correctly' here comes from Wakefield's definition and I take it to be synonymous with 'well-reasoned/justified'. To label someone's thoughts and behavior's as 'broken' or 'bad' in anyway invites stigma and has a huge impact on people's lives and

[10] This is not to pre-suppose a categorical difference. In fact, the divide seems likely to be continuous.

self-understandings. As the arbiters of such labels, psychiatry and clinical psychology need explicit ethical guidance, a necessary part of which is a clear understanding of what counts as mental disorder and what doesn't. For this and many other reasons[11], the conceptual pluralism prescribed when discussing the structural nature of mental disorder can seem less applicable when discussing the normative nature of mental disorder. By this I mean that if we are going to label someone as 'dis'-anything, we ought to be able to provide good reasons for doing so, and we ought to seek to be correct in making this distinction (whatever that may turn out to mean).

Even if there is 'one correct' way to understand the normative nature of mental disorders, conceptual pluralism may still be the best way forward given the complexity at hand. Fulford and Colombo (2004) give the analogy of a complex mural on the wall in a dark room, with the mural representing the 'correct' concept of mental disorder. There are six people in the room and each one is given a flashlight. The beam of each flashlight, through taking a different perspective, reveals a different facet of the mural. With enough flashlights we may hope to perceive the entire mural, but each individual flashlight likely has value in this task. I would add to this however, that given the ethical weight of our task alluded to above, critical care is required; we need to make sure that someone isn't pointing their flashlight at the wrong wall.

In what follows I overview some of the conceptual models offered as justification for use of a mental disorder label, or those that attempt to offer guidance as to what should count as mental disorder. It is not my intention to cover all normatively oriented models available as this is not a comprehensive review. For example, I do not cover models that see mental disorder as an entirely moral or religious concept, nor do I cover those reason-based models that see mental disorder as defined in some way by irrationality[12] (Graham, 2013; Megone, 1998). I also do not cover

[11] See Telles-Correia, Saraiva, and Gonçalves (2018) and Wakefield (1992a, 2007) for discussions surrounding the need for a precise definition. Contrariwise see Bingham and Banner (2014).

[12] Briefly, my key issue with these reason-based-models is that they commit to an understanding of the 'rational man' as an ideal from which to contrast disorder. This seems very culturally specific, and it seems there is a risk that this may illegitimately pathologize cultural variance. Megone's (1998) model in particular is also reliant on unfavorable ideas such as Aristotelian teleology (final causes as a function of essence), and human exceptionalism (the idea of a unique and vital difference between humans and animals).

Roschian models that hold mental disorder to be a multi-dimensional cluster concept, centered around a prototype rather than necessary and sufficient conditions[13] (Lilienfeld & Marino, 1995; Walker & Rogers, 2018). I focus instead on families of conceptual models that are currently or recently popular, and that together offer the reader a general overview of the conceptual landscape. I first briefly cover anti-psychiatric or *deflationary* positions as these historically provided the impetus for the development of the other models in this section. I then cover *statistical functionalism*, followed by *evolutionary functionalism*. I then discuss *evaluative* concepts, and finally *practical kinds*. Note that some of these normatively oriented models draw from the philosophy of medicine, and are often concerned with disorder, dysfunction, or disease in general rather than just mental disorder. Because of this I occasionally draw on examples across both physical and psychiatric medicine.

Anti-psychiatric/Deflationary Positions

In overviewing understandings of what makes mental disorder 'disordered', it would be remiss to not highlight those views that hold the label of disorder to be unjustified and/or unethical. Because of their use by persons and groups opposed to the institution of psychiatry through the latter half of the twentieth century, these positions are often referred to as *anti-psychiatric*. However, 'anti-psychiatry' is quite a loaded term, and it is important to distinguish between opposition to psychiatry as a whole, and principled disagreement with the concept of mental disorder. For these reasons it may be better to refer to these positions as *deflationary*. These deflationary positions are responsible for much of the debate concerning the normative justification for the mental disorder label as they represent the null hypothesis: that in important ways the label 'mental disorder' fails to refer to anything in nature. To be clear, what is 'deflated' within such perspectives is the notion of mental disorder as a real/

[13] Briefly the issue with these Roschian/Wittgensteinian models is that they are overly flexible, thereby providing very little specificity or guidance. This is a similar weakness to the pragmatic concepts that I will discuss. I will briefly return to Roschian models when discussing the work of de Haan in later chapters.

natural/worthwhile concept, not our notion of truth itself. This is important to clarify because I am borrowing this term from philosophy where it is often used in this slightly grander way.

The psychiatrist and philosopher Thomas Szasz is responsible for the most famous of these deflationary positions (Szasz, 1960). The core of Szasz's position is that real illness or disorder is necessarily a bodily phenomenon. If this is assumed, then the category 'mental disorder' seems problematic. What we refer to as mental disorders will either turn out to have a physiological cause—and thus be disorders of the brain or body—or they will turn out to have no basis in the body, and therefore not qualify as genuine instances of illness/disorder. For Szasz then, 'mental' disorder is an impossibility and our use of the term must be a 'myth'. While, in public discussion, Szasz is often implied to be some sort of radical social constructionist, his issue with the concept of mental disorder actually stems from a position of biological disease realism. Szasz's use of the word 'myth' is very intentional and has a double meaning. On one side he is referring to the apparent impossibility of *mental* disorder (as explained), and on the other he is speculating that we use the notion of mental illness/disorder to distance ourselves from the harsh realities of our society. The idea here is that the labeling of genuine but normal 'problems in living' as medical issues, and thereby as uncontrollable deviances from the norm, allows us to believe that the society we have constructed is kinder than it really is.

Another famous deflationary position is that of the philosopher Michel Foucault (2003/1961). Foucault's study of the development of the concept of madness in Europe lead him to the conclusion that the modern label of mental disorder is primarily a label for social deviance, and a tool for controlling those whom society disvalues. While we have come to see a categorical difference between those that suffer mental disorder and those that do not, Foucault's analysis suggests that such objectification of these differences has in part arisen because of the way we have historically separated those viewed as 'mad'—alongside political dissidents and criminals—from the rest of society through the practice of institutionalization.

While neither of these views is currently popular in the mainstream psychopathology literature[14], it is somewhat unfair to say they have failed simply because the institution of psychiatry still stands. Many of the normatively oriented concepts I will explore in this section were conceived of as responses to the concerns of these deflationary positions. These deflationary views helped to highlight why the sciences of psychopathology need a strong conceptual base, including a principled reason to demarcate the disordered from the benign. Without such a reason, those of us currently working with mental health diagnoses are practicing on the basis of a non-natural and/or unjustified conceptual framework. In other words, these deflationary positions demonstrate that without a convincing positive understanding of what mental disorders are, psychologists and psychiatrists potentially lack sufficient ethical justification for their practice.

Statistical Functionalism

One common understanding of what counts as mental disorder is that it has something to do with deviation from the statistical norm. This view is apparent when we use the term 'abnormal psychology' as synonymous with 'dysfunctional' or 'disordered' psychology. Unfortunately, by itself such a view does not get us very far. This is because it cannot distinguish between 'good' and 'bad' forms of abnormality, e.g., being abnormally good at mathematics or abnormally good at giving speeches does not seem to count as a mental disorder. For this reason, conceptual models of what counts as mental disorder based around typicality have to further specify what kind of abnormalities or typicalities are relevant to disorder and why. *Functionalism* of some stripe or another often fills this position and will be discussed in the current section. In the following sections I will also discuss models that use *values* or *pragmatics* to fill this space.

The most well-known position of the *statistical functionalist* variety is the Bio-Statistical Theory of Health (BST) developed by Christopher

[14] Such views are expressed elsewhere in academia. One notable example from within psychopathology is the Power Threat Meaning Framework (Johnstone et al., 2018) which takes a similar deflationary perspective on mental disorder.

Boorse (1975, 1977, 2014). This is a conceptual model of health and 'disease' in general but can be used to inform a view of mental disorder. Under the BST, a disease is an *internal* state that impairs health by bringing about reduced efficiency of so-called *normal functions* relative to a *reference class*. Reference classes are members of the same species, sex, and age group[15], thus making normal functions effectively things that others like you can do that contribute to survival or reproduction (Boorse, 1977; Nordenfelt, 2007). If you go bald at the age of 13 while other teenaged humans of your sex do not, then this would count as disease under the BST (so long as hair can be assumed to serve a biological function such as keeping the sun off your head and/or helping to attract mates). The general gist of the BST is that "diseases are internal states that interfere with functions in the species design" (Boorse, 1977, p. 558). Boorse developed this concept to be explicitly value-free; as a concept that sees diseases as empirical facts rather than value-based distinctions[16]. For Boorse then, 'disease' is a theoretical/technical concept and should be distinguished from a more general sense of 'illness' which he does see as value-laden[17]. In other writings he has used the alternative term 'pathology' to refer to disease/disorder (Wakefield, 2014a).

While he does not make it a focus of the theory it is important to note that Boorse (1977) limits the kinds of things that can count as diseases under the BST to inefficiencies/difficulties with *physiological* functions. Thus, I refer to the BST as an example of *physiological statistical functionalism.* For example, someone with abnormally high blood pressure

[15] Boorse indicates that ethnicity should sometime be considered insofar as the differences in functional design across ethnic groups are relevant (Boorse, 1977).

[16] Both Kingma (2007) and Varga (2011) counter Boorse's claim that the BST is in fact value-free by pointing that the use of sex, age, and ethnicity to define the reference class is not itself based on empirical fact but on intuition, and thereby is likely importing value into the process. For example, one common criticism of the BST is that is seems to define homosexuality as a disease on the basis of its statistical deviance and the resulting lower rates of reproduction. Kingma points out that the addition of sexual orientation to the defining attribute of the reference classes would change this entirely. Those that include sexual orientation in the reference class selection would view homosexuality as entirely normal, and those that do not would view it as a disease. Really the BST is only potentially value-free post the selection of a reference class.

[17] Fulford (2001) criticizes the BST, for one arguing that, even if it does produce an internally consistent value-free concept of disease it fails to recognize that the term 'disease' is *used* evaluatively, even by Boorse himself.

relative to a standard developed by measuring the blood pressure of others of the same sex and age could be said to have a disease (hypertension) under the BST, whereas someone with abnormally low empathy would not *necessarily* be seen to have a disease under the BST. In order to be seen as diseases under the BST an assumption has to be made that abnormal mental conditions are causally supported by an abnormal physiological structure (usually in the brain). On this view then, mental disorders are not 'mental diseases' but rather physiological diseases, not yet understood, that happen to feature mental and behavioral outcomes (hence why they are sometimes referred to as 'disease models'). The BST, and other (*physiological*) *statistical functionalist* views—e.g., Reitschel (2014) and the RDoC movement, see Insel et al. (2010) and Nielsen (2020) for further discussion—are typically associated with a clearly categorical or even essentialist structural view, whereby mental disorders are assumed to have yet to be discovered dominant causal factors or essences. It is this exclusion of the possibility of independent mental dysfunction/disorder (mental difficulties without a physiological abnormality as a basis) that opens such views to charges of reductionism.

Not all views that could be labeled as varieties of statistical functionalism are restricted to physiological deviations. For example, Bergner (1997, 2004)—continuing the original work of Ossorio (1985)—proposes a *disability concept* of mental disorder[18]. A key part of their definition is that mental disorder involves significant restriction in a person's ability to engage in deliberate behaviors that that they *ought* to be able to engage in. Regarding this use of 'ought', Bergner (1997) explains that 1) this is purposefully ambiguous in order to accommodate clinical judgement, but also that 2) the idea is that the behaviors one 'ought' to be able to engage in are specified in a sense that is "highly developmental and highly contextual" (p. 240). The essence of what Bergner is claiming seems to be that mental disorder concerns *deliberate behaviors that others can typically perform but that the sufferer cannot*, while excluding any such restrictions on behavior that can be explained in reference to contextual factors (e.g., age, culture, immigrant status, physical environment).

[18] For further (empirical) support of this disability view see Bergner and Bunford (2017), for a critique see Wakefield (1997b).

Direct parallels are clear here to the BST and the idea of relativizing disease to a reference class (although the 'reference class' in this model is much more specific). It is for this reason that I consider Bergner to be proposing a form of *behavioral statistical functionalism*[19].

The key difficulty with statistical functionalism applied to mental disorder can be summed up by the question 'why should being normal matter?' In both varieties of statistical functionalism espoused here, the typicality of some state or action is used to infer that this state is the way that our bodies *ought* to be, or that this action is the way we *ought* to act. Problematically, the link from the 'is' of the statistical norm, to the 'ought' of claiming that a biological state of affairs is *better or worse* than another—what I will refer to as *the normative gap*[20]—seems reasonably thin and unclear. For Bergner, this normative gap goes virtually unrecognized, while for Boorse, the (tentative) link has to do with the normal state representing species design/baseline health: "…the normal is the natural" (Boorse, 1977, p. 555). This does not seem like a big problem when considering physical disorders because at this level what is 'good' versus 'not good' is generally quite clear. As a simple example, most people agree that a heart attack is just plain bad. When speaking of behavior, thought, and emotion however, there is not always one right way to function. Cultural variation is a good example of this. In explicitly evaluative words unavailable to these authors, there is a diversity of legitimate values in the psychological realm that is not present in the physiological (Fulford, 2001). For example, statistical functionalism is often argued to erroneously capture homosexuality under the banner of mental disorder given

[19] This label is by no means a perfect fit, for example, I am not sure whether Bergner and Ossorio would agree with the use of 'functionalism' here. I could label it *contextualized behavioral statistical-ism* or something similar. However, in so far as behaviors one 'ought' to be able to do can be referred to as functions the label used seems acceptable. The current label also highlights important similarities across divergent views; just as the BST contrasts the individual's physiology against a reference class, this view contrasts the individual's capacities against similar others in similar contexts. Further, my sense is that Bergner would disagree that context can ever really be sufficiently captured by use of a reference class nor any statistical means, and that therefore clinical judgement will always be required in diagnosis. He is probably right, but how do we go beyond the statistics while maintaining clarity, rigor, and a common language? This is another reason why a richer conceptual model/framework is required.

[20] This normative gap is of course nothing new—it is simply the domain-local version of Hume's 'ought-from-an-is' problem (Hume, 1978/1738)

it is statistically deviant and results in less offspring. This all suggests very strongly that the use of statistical normality, even if applicable to the definition of dysfunctional physiology, is not applicable in the definition of dysfunctional psychology.

At this impasse there are two options standardly recognized: 1) move away from statistical normality and attempt to plug the normative gap with a better story of how functions can naturally arise. I will explore this option in the next section on evolutionary functionalism. Alternatively, 2) recognize that values do play a role in defining mental disorder, as explored in the following section of value-laden concepts. At the end of this chapter I will suggest that there is another, less recognized, option available to us.

Evolutionary Functionalism

Under evolutionary functionalism, what is disordered is that which fails to perform its evolved function. Rather than deriving ideas of function from that which is statistically normal as above however, this position holds that functions are capacities that parts of the body or mind have, *due to their being selected for across the evolution of the organism*. Evolutionary functionalism then, attempts to plug the normative gap using evolutionary theory. The most well-known conceptual model of this type is Jerome Wakefield's harmful dysfunction (HD) analysis, or more specifically the 'dysfunction' component of this model (1992b, 2007, 2014a). The HD analysis is a two-part model. It holds that mental disorder is 'dysfunction' plus 'harm'. In this section I will discuss the dysfunction component of Wakefield's HD analysis as it is a good example of the pitfalls that arise for the evolutionary functionalist, despite the positions intuitive appeal (I will explore the harm component in the value-laden concepts section).

On the HD view then, dysfunction is a necessary but not sufficient component of disorder (Wakefield, 1992b, 2007, 2014a). This is contrary to the BST in which dysfunction by itself is sufficient for attributing disorder (or rather disease/pathology in BST terminology). The dysfunction component of the HD analysis is defined evolutionarily, requiring that mental disorders include a part or behavior of the organism that

doesn't do what it has been selected to do by the evolutionary process: "A 'dysfunction' exists when an internal mechanism is unable to perform one of its natural functions" (Wakefield, 2007, p. 152). Comparing to the BST once again, the key difference here is the use of the term '*natural* function' as opposed to '*normal* function'. The former are products of random mutation and natural selection across time, and the latter are statistically derived (Boorse, 1977; Wakefield, 1992b). Specific to mental disorder, Wakefield describes the internal mechanisms concerned as 'mental mechanisms'; as evolved tendencies and capacities in behavior, motivation, cognition, perception, or emotion, that have been selected for due to their serving the survival and reproduction of the species and their ancestors[21]. Mental dysfunction within the HD analysis then, is when evolved mental mechanisms don't function as designed by natural selection (with *disorder* being ascribed when the dysfunction results in socio-culturally defined harm). For example, genuine cases of depression, for Wakefield, represent a malfunction in the psychological mechanisms evolved to regulate emotion, leading to a set of behaviors and experiences society deems harmful (Wakefield, 1997a). Hence, Wakefield's well-known criticism of the removal of the bereavement exclusion in the DSM-5 depression criteria: grief following bereavement is not a dysfunction, but rather an evolved mechanism acting as it should (Wakefield, 2013).

Despite the popularity of the HD analysis, many critiques have been made of this approach to understanding dysfunction. Unfortunately these critiques tend to be quite complex (perhaps supporting the HD analysis' continued popularity). Due to this complexity I do not have room to fully unpack these critiques here. For those interested a fuller summary of these critiques can be found in Chap. 2 of my PhD thesis (Nielsen, 2020). To offer the briefest of summaries, there are three different modes of critique launched at the notion of dysfunction within the HD analysis. The first simply attempts to generate counterexamples to the HD notion of dysfunction. Many such examples have been argued, such as cases of depression or conduct disorder where difficulties appear

[21] This use of 'mechanism' is again bio-functional, a common intent. Broader definitions of mechanism are in use so it is important to specify (Andersen, 2014a, 2014b; Garson, 2017; Illari & Glennan, 2017).

to arise not from any dysfunctional mechanism but by normal processes of mood regulation or learning occurring in a pathogenic environment (Lilienfeld & Marino, 1995; Murphy & Woolfolk, 2000; Nesse, 2001; Varga, 2011). Unfortunately this mode of critique has tended to devolve into complex backwards and forwards arguments between Wakefield and his critics, where one gets the sense that both sides think they have bested the other.

The second mode of critique is targeted at Wakefield's use of evolutionary theory within the HD analysis. These critiques target keystone claims within Wakefield's framework, namely that evolution can be said to provide 'designs' or attributes with 'purposes', and moreover that we can confidently claim to know about such designs/purposes. These critiques argue that deviation from an organism's apparent 'design', as well as large degrees of contingency and randomness, are such vital components of the evolutionary engine that the human notion of 'design' seems to be somewhat of an inappropriate analogy. Further, even if such designs or purposes exist, it is doubtful that we could ever be confident in our knowledge of them given the complex and historical nature of even the most basic evolutionary adaptions (Lilienfeld & Marino, 1995; Murphy & Woolfolk, 2000; Sadler, 1999; Sadler & Agich, 1995).

Finally, the third mode of critique attempts to undercut the claim that HD-style dysfunction is value-free (Murphy & Woolfolk, 2000; Sadler & Agich, 1995). These critiques argue that, because of the inherent difficulties with figuring out something's evolutionary function, values will always permeate in the actual application of the HD analysis. Given our current (and likely future) inability to confidently know the evolutionary functions of a behavior, the HD notion of dysfunction can offer very little guidance in practice. Worse, it may encourage us to generate evolutionary stories that implicitly align with our values and biases. As an example, homosexuality could conceivably be considered a dysfunction in Wakefield's sense, given it presumably leads to lower reproductive success. While there are evolutionary theories as to the possible adaptive function of homosexuality, these are (and likely will continue to be) speculative and contested. The HD notion of dysfunction therefore offers unacceptably little guidance as to whether homosexuality should be considered a dysfunction.

Before moving on, it is useful to consider an evolutionary functionalist position different to that of HD-style dysfunction; that of Troisi and Macguire (2002). I mention this here because, in generating their own position of 'Darwinian Psychiatry', these authors demonstrate awareness of some of the mentioned epistemological issues with evolutionary functionalism that hamper Wakefield's analysis of dysfunction. In particular, Troisi and Macguire point out the vital role of phenotypic variability in the evolutionary process, as well as that the evolutionary fitness of a behavior is highly contingent and nigh on impossible to measure directly. In doing so they acknowledge our epistemological limits concerning the evolved functionality of a behavior. As such they suggest a need to measure *functional consequences in the individual* rather than inferring whether they were adaptive for the species in the ancestral context. The problem with this of course is that 'functional consequences' in a Darwinian frame boil down to the number and quality of the offspring produced. Due to obvious time constraints we can't sit around and wait while counting the number of off-spring someone has and/or how long they live. Troisi and Macguire' solution is to suggest the use of 'the achievement of short-term biological goals' as a proxy measurement for evolutionary success. 'Darwinian Psychiatry' then is a much more successful but much less ambitious variation of evolutionary functionalism in comparison to the HD analysis. More importantly for the current discussion however, the limitations these authors place on themselves stem directly from their understanding of the messy realities of evolution. These limitations highlight nicely where Wakefield's concept of dysfunction arguably oversteps what evolutionary theory can truly provide.

Evaluative Concepts

The normative conceptual models explored so far have all been attempts at *naturalizing* mental disorder; of limiting the normative scope of the concept to exclude values, especially individual and culturally specific

values[22]. Many authors argue however, that attempting to do so is futile and we should instead be open and honest about the role of values in psychiatric diagnosis (Doust et al., 2017; Fulford, 2002; Sadler & Agich, 1995; Stier, 2013). Metaphorically, these positions are bridging the normative gap with values, sourcing their claims about the 'goodness' or 'badness' of human thought and behavior from socio-cultural value structures. Moreover, those who hold this position tend to claim that everyone else is doing this too, only without realizing it. Positions that recognize the role of values in this way are broadly known as *evaluative* in nature. In contrast, the collective term for those who attempt to naturalize mental disorder—to see it as purely factual—are most typically known as *descriptivists* (Fulford, 2002). In line with Zachar and Kendler (2007) however, I will refer to this position as *objectivism* in order to avoid using multiple senses of 'descriptivist' across this project.

Generally speaking, evaluativists are motivated by two observations. The first of these observations is that values are almost certainly playing a role in the conception and application of current diagnostic concepts (Foucault, 2003; Sadler, 2005; Stier, 2013; Szasz, 1960). If this is true, this means that when a clinician or psychiatrist makes a diagnosis, there seems to be a very real sense in which they are evaluating the client rather than simply describing their state. Objectivists find this conclusion unsettling, preferring that diagnosis be a purely factual matter (for example see; Hucklenbroich, 2014). A workable objectivist rebuttal here is that evidencing the value-laden nature of current concepts and diagnostic practice speaks only to an understanding of concepts and practice *as they are*, not necessarily *as they should be* (Muders, 2014). This thereby leaves room for the possibility that, despite the role of values in current diagnostic concepts, there is a way to consider them as wholly objective and that perhaps such a way is preferable.

The second observation that often motivates evaluativism is simply that popular objectivist approaches, such as the two brands of

[22] The popularity of such naturalized value-free models may well be a reaction to the arguments of the anti-psychiatry movement who questioned the concept of mental disorder predominantly on the basis of its evaluative (and therefore on their view non-scientific) conceptual nature (Varga, 2011).

functionalism explored above, seem to fail to distinguish between disorder and non-disorder effectively. For example, Doust et al. (2017) explore three examples of conventionally accepted medical disorders and demonstrate that functionalism offers very little guidance as to where the boundaries of disorder should be placed. Instead, they propose, the answer to this question seems to revolve around the values at play. Therefore, they argue that our conceptual models should openly recognize the role of values in demarcating disorder. If they do not do so, we meet the same problem we saw with the HD notion of dysfunction where values may creep in unannounced and therefore unconsidered. Problematically however, Doust et al. offer no framework for how this recognition of the role of values could be achieved.

There are generally three different evaluative stances, taken in response to the acceptance of these observations, as to what a concept of mental disorder should be. I refer to these stances as: *weak-evaluativism*, *strong-evaluativism*, and *anti-psychiatric evaluativism*.

Weak-evaluativism simply recognizes that terms like dysfunction and disorder are evaluative *in a limited sense*. Specifically, weak-evaluativism does not prescribe the inclusion of socio-culturally and individually specific values in consideration of what counts as disorder. According to the weak-evaluativist then, cases where socio-cultural values are playing a role in diagnosis—e.g., see Stier (2013)—are in error. Under weak evaluativism, the values at play are assumed to be universal and therefore not particularly contentious. This brand of evaluativism seems potentially workable for bio-medical disorders where values are relatively agreed upon—e.g., it doesn't seem contentious to say that brain tumors are bad—but seems much less workable in the domain of mental disorder where values are exponentially more diverse (Fulford, 2002).

Strong-evaluativism, in contrast to the weak form, accepts that socio-cultural and individual values should and do play a role in demarcating disorder. The immediate problem with this position however, is that it introduces a high degree of relativism (Jefferson, 2014). This is where what counts as disorder changes across cultures and time periods, dependent on the local value set. For example, under a strong-evaluativism, the labeling of homosexuality as disordered within the bounds of a conservative culture seems concerningly uncontestable. This relativism also

opens-up boundary issues, i.e., how do we know whose values to use, and where does one culture stop and another start? It is potentially due to these issues of relativism that very few strongly evaluativist concepts have been proposed as formal conceptual models of mental disorder.

Finally, the third evaluativist position that can be taken is anti-psychiatric evaluativism. This position holds that concepts of mental disorder are so value-laden that they do not refer to anything 'real', that they are ethically unacceptable, and that we should therefore discontinue their use. Foucault's (2003/1961) position mentioned in the deflationary section would be an example of this kind of evaluativism.

One unique approach to strong-evaluativism that seems to successfully contain the threat of relativism is the HD analysis (Wakefield, 1992b, 2007, 2014a). By specifying that both harm and dysfunction are necessary for an attribution of disorder, but that neither is individually sufficient, Wakefield incorporates socio-cultural values into his conceptual model while staving off unconstrained relativism. Under the HD analysis, harm is considered in explicitly culturally relative terms:

> "…disorder lies on the boundary between the given natural world and the socially constructed world; a disorder exists when the failure of a person's internal mechanisms to perform their functions as designed by nature impinges harmfully on the person's wellbeing as defined by social values and meanings." (Wakefield, 1992b, p. 373).

The general gist of this idea—how it utilizes both components to constrain the other—is regarded highly. For example, renowned author in this area, Peter Zachar, refers to the HD idea as "parsimonious, elegant, and useful" (2014, p. 121); three descriptive terms of which I would certainly agree with the first two. The issue, as we saw in the previous section, is primarily with the workability of the dysfunction component. It is not clear whether this notion of dysfunction represents an acceptable use of evolutionary theory, nor whether we can ever obtain the deep knowledge of evolutionary processes required to utilize it. Hence, with the dysfunction component virtually defunct, the parsimony of the HD idea, and how it attempts to put strong-evaluativism to work in a suitably constrained manner, ultimately falls flat.

Before moving on I should note that a core assumption of the current project is that, in the demarcation of disorder, the question of whether norms and values have a role to play *at all* is somewhat trivial. At its simplest, a diagnosis is a claim that something is *wrong* with a person. On my view it is therefore *necessarily* normative/evaluative, and I therefore reject total objectivism (although not, as I will show, the allure of naturalization). In Chap. 4, I will attempt to carve new ground between the weakly and strongly evaluative positions. The resulting view will include certain socio-cultural values as relevant to mental disorder on a principled basis, while maintaining a thoroughgoing naturalism. This will be achieved through the use of a framework that subscribes to value-inclusive naturalism, allowing us to move beyond the dichotomy of objectivist versus evaluativist positions (Thornton, 2000).

2.3 Practical Kinds

Faced with the many competing normatively oriented concepts explored above, some authors have suggested turning to pragmatism for solutions. A *pure* or *radical* pragmatic view holds that the underlying structure of mental disorders is either that of 1) non-kinds and therefore continuous with normal human behavior, or 2) totally socially constructed. Nonetheless the pragmatist holds that it is *useful* for our purposes as explainers and clinicians (who work within socio-legal environments that often demand categorical identifiers) to treat them as more 'real' and categorical than they may be. On this view then, it is the *usefulness* of mental disorder concepts that justifies their use, despite the fact that they may not refer to any real kind in nature (Haslam, 2002; Kendler, 2016). To return to our metaphor, the pragmatists are skipping over the normative gap and saying 'let's just do what seems useful'.

In this radical form, pragmatism risks total conventionalism (in the sense that they have no referent in the natural world and are thereby empty labels). This where what counts as mental disorder are simply those things that we, or a particular group, *label* as mental disorders. For example, O'Connor (2017) presents the idea that mental disorders are

practical psychiatric kinds. By this he means that mental disorders are those categories that psychiatry invents because they are useful for psychiatry's purpose of helping people. This position is not intended to be a deflationary one; rather than define psychiatry as the profession that treats mental disorder, O'Connor flips this around and defines mental disorder as that which psychiatry treats. Psychiatry in turn is defined in a broader sense as the profession that aims to "…help those with emotional or psychological impairments who seem to be unable to help themselves." (O'Connor, 2017, p. E-8)[23]. This position rejects naturalism about mental disorder, both in the sense that mental disorder may represent natural dysfunction/s, and in the sense that mental disorders may be understood structurally as natural kinds. Rather for O'Connor, mental disorder concepts are the products—and tools—of psychiatric practice which, in turn, he seems to see as a broadly moral enterprise. While this may represent a valid—if slightly disparaging—perspective on the nature of current diagnostic concepts in mental health, it still leaves mental disorders as totally conventional entities and thus provides next to no guidance as to what kinds of things we should or shouldn't count as mental disorder.

In response to this issue of conventionalism, some pragmatist positions take only a *partially* pragmatic approach by incorporating other normative or structural elements. One such model would be Zachar's Practical/MPC hybrid model (2015). This model combines the concept of a fuzzy MPC kind with pragmatism:

> "Concepts for psychiatric disorders are constituted by discoveries *and* decisions. There is an interaction between what the world produces and what we find useful to notice." (Zachar, 2015, p. 289).

Under this model, paradigm mental disorders are seen to be likely tracking MPC like structures in human behavior. The fuzzy nature of MPCs provides instances of ontological indeterminacy, in response to

[23] There is a charge of circularity that can be made against this position. For example, what exactly defines an 'emotional or psychological *impairment*? This seems to be another term for a mental disorder. I take this to be representative of O'Connor's point—on his view mental disorder is a conceptually thin notion, constructed through the practice of a morally defined institution.

which classificatory decisions are made in accordance with our pragmatic purposes. For example, if, for the moment, we assume that depression and its melancholic subtype are MPC kinds whose properties overlap, there is a genuine sense in which the decision to treat these entities as having a type-subtype relation is somewhat arbitrary. We could alternatively treat them as different entities with similar symptom profiles. This is not a totally conventionalist position as there are structures in nature to which mental disorder labels are thought to refer, but Zachar's model highlights that many such arbitrary or pragmatic decisions have, over time, shaped our diagnostic systems[24].

Again however, a pressing issue with Zachar's (2015) model concerns the lack of guidance it provides. It is undeniable that our current diagnostic concepts are partially historical in nature; that their current form is contingent upon past human affairs and decisions rather than representing naturally separable phenomena. Pragmatism helps us recognize this but doesn't necessarily treat it as a problem, let alone provide a solution. This is because, other than their usefulness, pragmatism doesn't commit to any particular notion of what a diagnosis of mental disorder *should* represent. Pragmatic notions of mental disorder seem too thin in that they fail to provide an ideal; they are 'unambitious' in this way (Kendler, 2016). If tomorrow, we discover a new putative mental disorder, pragmatism offers us very little help in deciding whether to include it in our diagnostic systems or not.

This concludes the review of the dominant positions available when considering the conceptual nature of mental disorder. All of the models presented can tell us something interesting about the nature of mental disorder, but all face significant problems. Again, please note that I have chosen to not review the concepts of disorder present in projects such as the DSM-ICD and the RDoC, as these have been sufficiently reviewed elsewhere and are not immediately central to the current work. Such reviews can be found in Nielsen (2020).

[24] Zachar explicitly recognizes this partial nominalism/historicism in his Imperfect Community Model, where mental disorders are seen to be clustered under a single banner partially due to genuine family resemblance, but partially due to pragmatic and historical factors (Zachar, 2014).

Returning to Human Functioning

Considering the conceptual positions reviewed in this chapter, we see evidence for the observation in Chap. 1 that there is a conceptual co-determinacy between conceptions of mental disorder and wider views of human functioning. Foucault, for example, was interested in the relation between individuals and society, believing that behavior is strongly regulated by socially generated norms and concepts (and therefore that the production of these norms and concepts is where true power lies in society). His understanding of mental disorder as a socially constructed label for certain kinds of deviance makes sense in light of this. As a further example, consider Insel and Cuthbert (2015) who argue for a biologically focused model of mental disorder as a route to precision medicine in psychiatry. Note how their essentialist assumptions make perfect sense given the medically minded and brain-focused approach to human functioning that they ground themselves in.

This same conceptual co-determinacy is most clear when considering the functionalist positions. The very idea of these positions is to contrast disorder against an understanding of the things humans should be able to do if they are functioning normally. For the statistical functionalist these things are derived from an understanding of what most others can do, for the evolutionary functionalist these things are derived from an understanding of what is evolutionarily successful. The connection is also clear in the evaluativist position. The evaluativist's central claim is that all objectivist positions fail because they miss the irreducible role of values in our lives. In essence they are saying something like 'we are more than our statistical normality, more that our ability to pass on our genes; we have values'. The claim then is that the objectivist does not hold a rich enough (i.e., value-inclusive) understanding of human functioning by which to contrast mental disorder. In sum, how we conceptualize mental disorder/dysfunction appears deeply related to our basic assumptions about human functioning. This simple observation was the initial impetus for the current project. It raises the specter of a possible way forward in our conceptual understanding of mental disorder. That is, through grounding ourselves in a rich and integrative understanding of human functioning novel insights may arise.

The Normative Gap May Be Artefactual

A second observation is that the 'normative gap' observed between simply describing human behavior and being able to say that some behaviors are disordered or bad in some way, may in-part be an artifact of how we talk about values. Typically, we talk about values as if they are entities that somehow transcend matters-of-fact, but assuming naturalism this simply cannot be the case. This observation has been made before, and put in much clearer terms by Thornton (2000). Thornton considers the debate between those who see mental disorder as necessarily evaluative (e.g., Fulford, Sadler) and those that are attempting to 'naturalize' mental disorder through the concept of a natural function (e.g., Boorse, and to a lesser extent Wakefield). The functionalists think, very roughly, that incorporating values into the concept of disorder/dysfunction is to admit that it is not a natural/scientific phenomenon. Hence, they are trying to show they can *reduce* this notion of mental disorder to a more basic, purely factual language. The evaluativists meanwhile disagree, believing that there is an irreducibly evaluative element to mental disorder. Thornton however, points out that in doing so, both sides tend to agree that *values are not natural*. Thornton's proposal is that a non-reductionistic understanding of naturalism does not rule out an understanding of values as part of the natural world: "…although mental illness cannot be reduced to the realm of law, it is no less real for that." (2000, p. 75). While he does not go into detail, what Thornton is implying here is that 'values' may be real things in the world, emergent at levels of organization higher than physics or chemistry. Further, he seems to be suggesting that the adoption of a naturalized but non-reductionistic worldview may help to resolve, or in other ways move beyond, the apparent evaluative-objective divide.

What this is calling for is a naturalized but non-reductionistic conception of human functioning; one that can incorporate the obvious fact that humans have values and that our functioning is deeply normative. Such a framework could conceivably plug the normative gap in a naturalistic way without leaving us making do with an impoverished notion of what it means to be human. This second observation then, is pointing in

a similar direction to the first. If we want a fuller understanding of mental disorder, we need to situate ourselves within a value inclusive understanding of human functioning. One framework that may be able to serve this role is *enactivism/3e cognition*.

References

Aftab, A., & Nielsen, K. (2021). From Engel to Enactivism: Contextualizing the biopsychosocial model. *European Journal of Analytic Philosophy, 17*(2), M2–M22.

Albert, P. R., Benkelfat, C., & Descarries, L. (2012). *The neurobiology of depression – Revisiting the serotonin hypothesis. I. Cellular and molecular mechanisms*.

Andersen, H. (2014a). A field guide to mechanisms: Part I. *Philosophy Compass, 9*(4), 274–283.

Andersen, H. (2014b). A field guide to mechanisms: Part II. *Philosophy Compass, 9*(4), 284–293.

Andrews, G., Slade, T., & Issakidis, C. (2002). Deconstructing current comorbidity: Data from the Australian National Survey of Mental Health and Wellbeing. *The British Journal of Psychiatry, 181*(4), 306–314.

Beck, A. T., & Bredemeier, K. (2016). A unified model of depression: Integrating clinical, cognitive, biological, and evolutionary perspectives. *Clinical Psychological Science, 4*(4), 596–619.

Bergner, R. M. (1997). What is psychopathology? And so what? *Clinical Psychology: Science and Practice, 4*(3), 235–248.

Bergner, R. M. (2004). An integrative framework for psychopathology and psychotherapy. *New Ideas in Psychology, 22*(2), 127–141.

Bergner, R. M., & Bunford, N. (2017). Mental disorder is a disability concept, not a behavioral one. *Philosophy, Psychiatry, & Psychology, 24*(1), 25–40.

Bingham, R., & Banner, N. (2014). The definition of mental disorder: Evolving but dysfunctional? *Journal of Medical Ethics, 40*(8), 537–542.

Bird, A., & Tobin, E. (2018). *Natural kinds*. The Stanford Encyclopedia of Philosophy.

Bolton, D., & Gillett, G. (2019). *The biopsychosocial model of health and disease: New philosophical and scientific developments*. Springer Nature.

Boorse, C. (1975). On the distinction between disease and illness. *Philosophy & Public Affairs, 5*, 49–68.

Boorse, C. (1977). Health as a theoretical concept. *Philosophy of Science, 44*(4), 542–573.

Boorse, C. (2014). A second rebuttal on health. *Journal of Medicine and Philosophy, 39*(6), 683–724.

Borrell-Carrió, F., Suchman, A. L., & Epstein, R. M. (2004). The biopsychosocial model 25 years later: Principles, practice, and scientific inquiry. *The Annals of Family Medicine, 2*(6), 576–582.

Borsboom, D., Cramer, A., & Kalis, A. (2018). Brain disorders? Not really… Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences, 42*, 1–54.

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies, 61*(1–2), 127–148.

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review, 125*(4), 606.

Cramer, A. O., Waldorp, L. J., van der Maas, H. L., & Borsboom, D. (2010). Complex realities require complex theories: Refining and extending the network approach to mental disorders. *Behavioral and Brain Sciences, 33*(2–3), 178–193.

Doust, J., Walker, M. J., & Rogers, W. A. (2017). Current dilemmas in defining the boundaries of disease. *Journal of Medicine and Philosophy, 42*(4), 350–366.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science, 196*(4286), 129–136.

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network pschometrics: Combining network and latent variable models. *Psychometrika, 82*(4), 904–927.

Foucault, M. (2003). *Madness and civilization*. Routledge.

Fried, E. I., & Cramer, A. O. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science, 12*(6), 999–1020.

Fried, E. I., van Borkulo, C. D., Cramer, A. O., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology, 52*(1), 1–10.

Fulford, K. (2001). 'What is (mental) disease?': An open letter to Christopher Boorse. *Journal of Medical Ethics, 27*(2), 80–85.

Fulford, K. (2002). Values in psychiatric diagnosis: Executive summary of a report to the chair of the ICD-12/DSM-VI Coordination Task Force (Dateline 2010). *Psychopathology, 35*(2–3), 132–138.

Fulford, K., & Colombo, A. (2004). Six models of mental disorder: A study combining linguistic-analytic and empirical methods. *Philosophy, Psychiatry, & Psychology, 11*(2), 129–144.

Gardner, A., & Boles, R. G. (2011). Beyond the serotonin hypothesis: Mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(3), 730–743.

Garson, J. (2017). Mechanisms, phenomena, and functions. In *The Routledge handbook of mechanisms and mechanical philosophy* (pp. 122–133). Routledge.

Ghaemi, S. N. (2009). The rise and fall of the biopsychosocial model. *The British Journal of Psychiatry, 195*(1), 3–4.

Gleaves, D. H. (1996). The sociocognitive model of dissociative identity disorder: A reexamination of the evidence. *Psychological Bulletin, 120*(1), 42.

Graham, G. (2013). *The disordered mind: An introduction to philosophy of mind and mental illness*. Routledge.

Haig, B. D., & Vertue, F. M. (2010). Extending the network perspective on comorbidity. *Behavioral and Brain Sciences, 33*(2–3), 158–158.

Haslam, N. (2002). Kinds of kinds: A conceptual taxonomy of psychiatric categories. *Philosophy, Psychiatry, & Psychology, 9*(3), 203–217.

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine, 42*(5), 903–920.

Hucklenbroich, P. (2014). Medical criteria of pathologicity and their role in scientific psychiatry – Comments on the articles of Henrik Walter and Marco Stier. *Frontiers in Psychology, 5*, 128.

Hume, D. (1978). A treatise of human nature [1739]. *British Moralists*, 1650–1800.

Humphry, S. M., & McGrane, J. A. (2010). Is there a contradiction between the network and latent variable perspectives? *Behavioral and Brain Sciences, 33*(2–3), 160–161.

Illari, P., & Glennan, S. (2017). Varieties of mechanisms. In *The Routledge handbook of mechanisms and mechanical philosophy* (pp. 109–121). Routledge.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry, 167*, 748–751.

Insel, T., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science, 348*(6234), 499–500.

Jefferson, A. (2014). Mental disorders, brain disorders and values. *Frontiers in Psychology, 5*, 130.
Johnstone, L., Boyle, M., Cromby, J., Dillon, J., Harper, D., & Kinderman, P. (2018). *The power threat meaning framework*. British Psychological Society.
Kendler, K. (2012). The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. *Molecular Psychiatry, 17*(4), 377.
Kendler, K. (2016). The nature of psychiatric disorders. *World Psychiatry, 15*(1), 5–12.
Kendler, K., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*(6), 1143–1150.
Khalidi, M. A. (2013). *Natural categories and human kinds: Classification in the natural and social sciences*. Cambridge University Press.
Kingma, E. (2007). What is it to be healthy? *Analysis, 67*(2), 128–133.
Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., & Clark, L. A. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454.
Lebowitz, M. S., & Appelbaum, P. S. (2019). Biomedical explanations of psychopathology and their implications for attitudes and beliefs about mental disorders. *Annual Review of Clinical Psychology, 15*, 555–577.
Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology, 104*(3), 411–420.
Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology, 12*, 435–463.
Magnus, P. (2012). *Scientific enquiry and natural kinds: From planets to mallards*. Springer.
Magnus, P. (2014a). Epistemic categories and causal kinds. *Philosophy Faculty Scholarship*. https://doi.org/10.1016/j.shpsc.2014.10.001
Magnus, P. (2014b). NK≠ HPC. *The Philosophical Quarterly, 64*(256), 471–477.
Mallon, R. (2016). *The construction of human kinds*. Oxford University Press.
McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy, 86*, 95–104.
Megone, C. (1998). Aristotle's function argument and the concept of mental illness. *Philosophy, Psychiatry, & Psychology, 5*(3), 187–201.
Molenaar, P. C. (2010). Latent variable models are network models. *Behavioral and Brain Sciences, 33*(2–3), 166–166.

Muders, S. (2014). On the concept of the normative in the assessment of mental disorder. *Frontiers in Psychology, 5*, 129.

Murphy, D., & Woolfolk, R. L. (2000). The harmful dysfunction analysis of mental disorder. *Philosophy, Psychiatry, & Psychology, 7*(4), 241–252.

Nesse, R. M. (2001). On the difficulty of defining disease: A Darwinian perspective. *Medicine, Health Care and Philosophy, 4*(1), 37–46.

Nielsen, K. (2020). *What is mental disorder? Developing an embodied, embedded, and enactive psychopathology*. PhD thesis, Victoria University of Wellington. http://hdl.handle.net/10063/8957

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 8*(6), 800–822. https://doi.org/10.1177/0959354318808394

Nordenfelt, L. (2007). The concepts of health and illness revisited. *Medicine, Health Care and Philosophy, 10*(1), 5.

O'Connor, B. (2017). Mental disorder as a practical psychiatric kind. *Philosophy, Psychiatry, & Psychology, 24*(4), E-1-E-13.

Ossorio, P. (1985). Pathology. In *Advances in descriptive psychology* (Vol. 4, pp. 151–202). JAI Press.

Rietschel, M. (2014). Mental disorders are somatic disorders, a comment on M. Stier and T. Schramme. *Frontiers in Psychology, 5*, 53.

Sadler, J. Z. (1999). Horsefeathers: A commentary on "Evolutionary versus prototype analyses of the concept of disorder". *Journal of Abnormal Psychology, 108*(3), 433–437.

Sadler, J. Z. (2005). *Values and psychiatric diagnosis* (Vol. 2). Oxford University Press.

Sadler, J. Z., & Agich, G. J. (1995). Diseases, functions, values, and psychiatric classification. *Philosophy, Psychiatry, & Psychology, 2*(3), 219–231.

Stier, M. (2013). Normative preconditions for the assessment of mental disorder. *Frontiers in Psychology, 4*, 611.

Szasz, T. S. (1960). The myth of mental illness. *American Psychologist, 15*(2), 113.

Telles-Correia, D., Saraiva, S., & Gonçalves, J. (2018). Mental disorder – The need for an accurate definition. *Frontiers in Psychiatry, 9*, 64.

Thornton, T. (2000). Mental illness and reductionism: Can functions be naturalized? *Philosophy, Psychiatry, & Psychology, 7*(1), 67–76.

Varga, S. (2011). Defining mental disorder. Exploring the 'natural function' approach. *Philosophy, Ethics, and Humanities in Medicine, 6*(1), 1.

Wakefield, J. C. (1992a). Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder. *Psychological Review, 99*(2), 232.

Wakefield, J. C. (1992b). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47*(3), 373.
Wakefield, J. C. (1997a). Diagnosing DSM-IV – Part I: DSM-IV and the concept of disorder. *Behaviour Research and Therapy, 35*(7), 633–649.
Wakefield, J. C. (1997b). Normal inability versus pathological disability: Why Ossorio's definition of mental disorder is not sufficient. *Clinical Psychology: Science and Practice, 4*(3), 249–258.
Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry, 6*(3), 149.
Wakefield, J. C. (2013). The DSM-5 debate over the bereavement exclusion: Psychiatric diagnosis and the future of empirically supported treatment. *Clinical Psychology Review, 33*(7), 825–845.
Wakefield, J. C. (2014a). *The biostatistical theory versus the harmful dysfunction analysis, part 1: Is part-dysfunction a sufficient condition for medical disorder?, 39*, 648–682.
Wakefield, J. C. (2014b). Wittgenstein's nightmare: Why the RDoC grid needs a conceptual dimension. *World Psychiatry, 13*(1), 38–40.
Walker, M. J., & Rogers, W. A. (2018). A new approach to defining disease. *The Journal of Medicine and Philosophy, 43*, 402–420.
Ward, T., & Fischer, R. (2019). Behavioral and brain sciences commentary on Borsboom, Cramer and Kalis: Families of network structures – We need both phenomenal and explanatory models. *Behavioral and Brain Sciences, 42*(E31) https://doi.org/10.1017/S0140525X1800122X
Zachar, P. (2014). *A metaphysics of psychopathology*. MIT Press.
Zachar, P. (2015). Psychiatric disorders: Natural kinds made by the world or practical kinds made by us? *World Psychiatry, 14*(3), 288.
Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164*(4), 557–565.

3

3e Cognition and Existing Enactive Frameworks

This chapter presents and summarizes the philosophical orientation referred to as *3e cognition*—seeing the mind as embodied, embedded, and enactive—alternatively labeled simply *enactivism*. The underlying thesis of the current project is that 3e cognition has the potential to provide a more integrative and richer framework of human functioning within which to study and attempt to treat mental disorders. Here I first introduce 3e cognition before looking at recent attempts to consider the wider concept of mental disorder from this perspective. I then comment on the strengths and weaknesses of these approaches. The perspectives I cover include Fuchs (2017), de Haan (2020), and Maiese (2021). I reserve discussing my own 3e Psychopathology framework for the next chapter (Nielsen, 2020; Nielsen & Ward, 2018, 2020). This will allow clear demonstration of where the continued development of 3e Psychopathology has benefitted from the insights of the authors reviewed here.

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7_3

3.1 3e Cognition

3e cognition refers to a set of overlapping ideas about how the mind works. Specifically, 3e cognition refers to a view of the mind as *embodied*, *embedded*, and *enactive*. In this section I will first unpack these three core ideas, and in the subsection following I will go deeper into some of the ideas underlying enactive philosophy.

Embodiment, Embedment, and Enactivism

Embodiment refers to the idea that the mind is no more than the body acting over time, in context, and as perceived by itself. Cognition emerges from the fully material processes of the situated body over time, and it is constituted by the brain-body system rather than the brain in isolation; we are *embodied* beings. The mind then is not 'a thing' above and beyond the organism, neither in the Cartesian mind-substance sense (i.e., we are not made of 'mind stuff'), nor in an information-theory sense (i.e., our mind cannot be uploaded to a computer without radically changing its nature[1]). Rather, 'mind' is an activity; something organisms do. On this view, cognition is not something done by the brain alone, but by the wider organism in non-linear concert, and situated in context. For example, this view allows us to understand basic adaptive process, such as the way our body reacts to threat, as thoroughly cognitive. Consider for a second the way your body reacts to the flashing lights of a police car behind you as you drive along the highway. In sum, the idea of embodiment somewhat decentralizes the brain in our efforts to understand human behavior, without denying that brains are clearly very important for much of the cognition we engage in and value as humans.

Embedment meanwhile, refers to the idea that interactions with the physical and social environments within which the organism is situated provide necessary conditions for life and cognition over multiple timescales. As biological organisms we have basic needs that our

[1] To do so would only ever produce a copy, and one of a fundamentally different kind given its disembodied nature. In a similar vein see Thompson and Cosmelli (2011) regarding the brain-in-a-vat hypothesis as an argument for embodiment.

environment must meet. We have evolved to meet those needs in accordance with the constraints of our past environments, as well as to shape the environment to serve our continued survival and functioning. We have collectively crafted societies to meet these needs in a continued and complex process of niche shaping, which in turn has shaped us as self-domesticated creatures. Across our individual lifespans too, our behavior and cognition have been shaped by our environment and those around us. We tend toward ways of acting on/changing our environment that have helped us meet our needs in the past. Even in the moment-to-moment timescale of everyday tasks, we are locked in a continual dance of interaction which scaffolds our cognitive processes. For example, I don't need to remember where I left my keys because I left them beside the door, and despite my tendency towards distractibility I can craft paragraphs with some semblance of coherence due to being able to offload my languaging on to my laptop. In sum, we are deeply environmentally and socio-culturally embedded creatures. The epistemological consequence of this is that to explain human behavior, we cannot consider the brain-body system in isolation but must consider the *brain-body-environment system* and how it changes over time.

Finally, we are *enactive* creatures. This idea subsumes and builds on the ideas of embodiment and embedment, hence I use the terms '3e cognition' and 'enactivism' interchangeably. There are generally two ways that enactivism is explained. Potentially the most common way is to say that under enactivism 'the mind' is a continually enacted action-oriented process rather than a mysterious object in our heads. This explanation contrasts enactivism with the so-called sandwich model of cognition where an agent is seen to perceive a stimulus, 'encode' and 'processes' it in their 'mind/brain', and then act back on the world. Enactivism in contrast—through a commitment to embodiment and embedment—places us more directly in the world. Instead of a sandwich model, where our experience is seen to be insulated and computed inside our skull, enactivism perceives a *closed sensorimotor loop* whereby action and motility is not an output of 'cognition' in the head, but a vital part of our *sense-making*—i.e., the embodied process of perceiving, making sense of, and engaging with the world. We thereby perceive, act, and live in the world *directly* rather than as homunculi peering out of our skulls through our limited

senses. This sort of explanation of enactivism is helpful, but I fear it does not do justice to what enactivism brings to the table above and beyond embodiment and embedment.

The second way to explain enactivism, which I favor, is admittedly somewhat more complicated, and I suggest that for those unfamiliar it may be worth rereading the current paragraph after completing section 3.1. This way to explain enactivism focuses more on ideas of *meaning, purpose*, and *normativity*, thus better demonstrating what enactivism has to offer above and beyond ideas of embodiment and embedment. One of the biggest mysteries in philosophy is how purpose, meaning, and a sense that some states of the world are better or worse than others (e.g., normativity), can arise in a world of objects and facts. The answer to this mystery, according to enactivism,[2] is that these things are brought forth—or rather *enacted*—through the recursive structures of life forms and their relationship to the world—i.e., the universe has meaning *for* life forms because of the way they are shaped to keep trying to live in an imperfect world. To unpack this, consider that all life must actively maintain a precarious organizational structure against the entropic flow of the universe around it, requiring the metabolism of energy from the environment to do so. To put this in simpler terms, life forms are really complicated and highly organized, while the surrounding universe tends towards disorganization. All life forms therefore have to do a lot of work in order to keep themselves up and running. If they stop doing this work they will die and decompose back into their environment. In a sense then, all life is constantly creating and defining itself. To do such work requires energy, which has to be sourced from the surrounding environment (e.g., food and oxygen), and the ability to adapt to threats and environmental changes (e.g., temperature fluctuations and other life forms that want to eat you). In order to source energy and avoid threat, organisms must, in either a basic or more complex sense, *make sense of the world* in order to

[2] The specific version of enactivism being described here is sometimes referred to as autopoietic enactivism, due to the central inspiration/metaphor for this position being the autopoietic process observed in cells. This is contrasted against more strictly anti-representational brands of enactivism—i.e., Radically Enactive Cognition or REC (Hutto & Myin, 2012)—as well as theories focused on perception—i.e., Sensorimotor Enactivism (O'Regan & Noë, 2001). For discussion of these labels see Ward et al. (2017).

navigate and respond to changes in it adaptively. This effectively sets up a selective pressure.

We can see then that having an embodied purpose to keep living sets up the emergence of meaning for the organism. For a given life form the world ultimately has meaning because some things help it survive and some things threaten its survival, and it will be of clear benefit to the survival of the organism to evolve and learn to make sense of the world in order to respond adaptively. To return to the same ideas using more technical language, under enactivism, organisms are seen as intrinsically purposive in that they are shaped to maintain their highly organized existence in the face of surrounding entropy. They strive to *self-maintain* and *adapt* to changing circumstance (Di Paolo, 2005; Thompson, 2007). This purposeful striving, inherent to all life, sets up a natural normativity and grounds the development of *meaning for the organism* through the needful relationship with the environment[3] (Colombetti, 2014; Thompson, 2007). This naturalized understanding of normativity and meaning is what enactivism brings to the table above and beyond embodiment and embedment and will play an important role in later chapters. For this reason it is worth delving a little deeper into some of the key ideas and concepts ideas underlying the enactive viewpoint.

Going Deeper with Enactivism

As you can probably tell, enactivism is a broad set of ideas, often described in different ways, and I will continue to flesh out this concept throughout this section, and indeed across the rest of the book. For the purposes of the current work, enactivism is best understood as referring to the central role of *meaning* in our lives—as providing a way to understand this meaning while grounded in our biology and a naturalized worldview. In this section I will attempt to further sketch out the 3e/enactive view of human functioning and lay out some of the underlying philosophical ideas beyond the three 'e's. To be clear, I am here only offering an outline of

[3] To phrase this dialectically, human agency and meaning are seen to emerge through the inherent tension between our need to preserve one's life and identity, and our metabolic need to be open to the world (Di Paolo et al., 2018; García Otero, 2022).

enactivism, not arguing for it. For a presentation and defense of the enactive viewpoint see: Colombetti, (2014); Fuchs, (2017); Gallagher, (2006, 2017); Maiese, (2016); Thompson, (2007); Hutto and Myin (2012, 2017) and Varela et al. (2017/1991).

To start with a classic enactivist example, consider a simple life form such as a bacterium. Bacteria have a tendency to move towards concentrations of sugar (their food source) and away from certain substances that are toxic to them, a process referred to as chemotaxis. This is achieved through a well-understood mechanism by which the motions of their flagella are responsive to the concentrations of sugar and some toxic substances (Wadhams & Armitage, 2004). Enactivism highlights that, parallel to its mechanistic nature, this is an *evaluative* process; one by which the bacterium is acting *in the interest* of its own survival. The claim here is not that simple bacteria are conscious in any way that we would recognize, but that sugar has an embodied *meaning* for the bacterium as 'good/food' and that there is therefore a simple mindedness/concern at play whereby the organism is responsive to the conditions required for its own survival. For the enactivist, this models a dynamic present in, and perhaps definitional of, all life—highlighting that *mind is in life* (Thompson, 2007).

This connection between the structures of life and mind is referred to as the deep continuity thesis (DCT). For now, suffice to recognize that there is a sense in which things can be good or bad for bacteria, trees, tigers, and people, in a way that things can't be good or bad for a pile of rocks. This is because, essentially, it is easy for these life forms to die, and hard for them to keep living; they are *precarious* in that they are far-from-equilibrium systems situated in an entropic universe. The very process of staying alive, i.e., *self-maintenance*, necessitates the metabolism of energy which is sourced from the environment. At the cellular level this process of self-maintenance is referred to as autopoiesis; literally 'self-creation'. For the enactivist, this precariousness and the needful relation it establishes between the organism and its environment is the root from which meaning develops.

Another key idea with enactivist philosophy is that of *organizational autonomy*. This is effectively a way to further specify the idea that life-forms maintain their own identity against the changing world around

them. While organisms must be thermodynamically open, in that they use energy from the world in order to keep themselves living, they also maintain themselves as organized structures distinct from their environment over time. All the various parts of an organism work together in a highly interrelated network of causal interactions, responding to and maintaining the wider organizational structure against external perturbations. In other words organisms can be understood as 'operationally closed' or 'autonomous' systems in that the action of their parts is largely constrained by the network of relations that represent and maintain the whole, thus defining the organism as a definable biological entity:

> "We shall say that autonomous systems are organizationally closed. That is, their organization is characterized by processes such that (1) the processes are related as a network, so that they recursively depend on each other in the generation and realization of the processes themselves, and (2) they constitute the system as a unity recognizable in the space (domain) in which the processes exist." (Varela 1979, p. 55, as quoted in Barandiaran, 2017).

When we speak of the bacteria above trying to 'stay alive' for example, what we are talking about is it managing to maintain its autonomous organizational structure. This idea of organizational autonomy then, serves as an underlying concept on which the enactive analysis of the emergence of meaning and normativity is built. This idea will be relevant later in this chapter when discussing the work of Maiese (2021).

A further vital point to understand is that within enactivist literature 'cognition' and 'perception' are seen as reasonably continuous and are often referred to as *sense-making*. This highlights the *relational* and *active* nature of cognition under enactivism. We can see this in the above example of bacteria; the sugar and toxins have meaning *in relation to* the bacteria's precarious situation as a life form. By acting differentially in relation to these two substances in its environment, and in a way that accords with the preservation of its autonomy (i.e., its self-maintenance), there is a very basic sense in which the bacteria is *making sense* of the world. The enactive view holds that a similar dynamic underlies the experience of meaning for more complex forms of life, only that, as the needs and

context of such organisms become more varied and complex, so too do their forms of sense-making.

As another example of sense-making and relational qualities, take the color red. Redness is what is referred to in philosophy as a 'secondary quality' because redness is not *in* an object the same way something's mass is. Redness rather, is subjective and experiential, phenomenal rather than noumenal (although as we will see, enactivism helps us move beyond this dichotomy). If there were no experiencing agents in the universe then there would still be objects, these objects would have mass, and light-waves would bounce off those objects in certain ways. But there would be no 'red'. So where does red come from? According to traditional cognitivist thought, redness is 'in the mind'; redness is an experience/neural code in our mind/brain, hallucinated in response to a particular pattern of activation in the optic nerve. Something about this seems rather absurd. Enactivism provides a different answer—redness is *relational* (Fuchs, 2017). It exists *between* the agent and the world, generated[4]—or rather *enacted*—by the organism, as a way to *make sense* of the world in accordance with its needs. Those from diverse theoretical orientations, including enactivism, can likely all agree that certain organisms have evolved to experience red (and other colors) because it is useful and helps them survive to pass on their genes. Under enactivism though, redness and other colors exist *for* the organism, not as part of a model or hallucination of the world in the mind/brain, but as a learnt and evolved mode of experiencing the world directly. Under enactivism then, there is a veridical world, but there is also what is called an *umwelt*; that same world *as experienced* from a concerned point of view, or the world *for* the organism. The umwelt then is a world of immanent meaning and valence (Thompson, 2007). Importantly, this is not a dualistic notion whereby these so-called 'worlds' are separate things, rather they are the same world viewed from two different perspectives—an abstracted third person view, and a concerned/lived view of the agent. This is referred to by Thomas Fuchs (2017) as *dual aspectivity* and it ensures an important/privileged

[4] I use the word 'generated' here very tentatively. I do not want to imply that enaction/sense-making is always an active/conscious process.

place for phenomenological/first-person methods within the enactive view.

Through enactivism the mind can be seen, not as a linear symbol processing machine with a defined inputs and outputs, but as *minding*; an emergent process/capacity of the whole organism arising from interactions in the brain-body-environment system, developed across evolution and life-span development, to better serve the organism's *self-maintenance* and *adaption* (Di Paolo, 2005; Maiese, 2016; Thompson, 2007). The mind then is not well modelled by computer metaphors and reference to such things as 'representation' and 'processing', but rather is better modeled by life itself:

> "a natural cognitive agent—an organism, animal or person—does not process information in a context-independent sense. Rather it brings forth or enacts meaning in structural coupling with its environment." (Thompson, 2007, p. 58)

When seeking to understand someone's behavior, the 3e/enactivist perceptive demands that we take as our central unit of analysis the sense-making organism standing in relation to its context; i.e., to consider the whole *brain-body-environment* as a dynamic system. 3e cognition is thereby, in a certain sense, holistically minded and anti-reductionistic. By this I mean that enactivism is incompatible with ideas of theory-reduction, where explanations of 'higher-level' phenomena (such as human behavior) are seen as in-principle reducible to the language of 'lower-level' theory, such as genetics or neural circuitry[5] (Andersen, 2016; Brigandt, 2013). This is because, implicit in the enactive view's talk of 'organization' is a commitment to *downward causation*; the view that phenomena at larger/slower levels of enquiry can influence the behavior of entities at

[5] While theory-reduction is certainly at odds with the 3e world view, I do not see a good reason why the 3e idea is incompatible with ideas of explanatory reduction/causal-mechanistic explanation—where wholes are broken down into parts to try and understand some property of the whole (so long as the holistic perspective is not sacrificed). In fact I think this style of explanation may be complimentary to the traditionally dynamical approach (see; Bechtel, 2009; Brigandt, 2013; Kaplan, 2015). I will return to this in Chap. 6.

smaller/quicker levels of enquiry.[6] The reason I mention this here is to explain a terminological shift that I would like to make at this point. Rather than the traditional ontological 'level', I will from this point refer to 'scale'. This is to highlight that, in a world featuring downward causation, simply because a phenomenon exists at a smaller scale of time and space, this does not make it somehow more fundamental or important. This shift is more in-line with the enactive view in that, if it is centered on any particular scale at all, it is of the whole organism and its experience (for further reasons for this shift in terminology see; Potochnik, 2010; Potochnik & McGill, 2012).

So, to summarize, under 3e cognition/enactivism mental processes are necessarily *embodied* in the brain, nervous system, and all other biological systems of the body—they are not things-as-products but processes that we do (i.e., embodiment). These processes necessarily occur within an environment with which we are richly and bi-directionally causally connected (i.e., embedment). For social-cultural creatures such as ourselves this includes a social environment, which we as a group constitute. Phenomenological experience and meaning emerges (i.e., is enacted) by virtue of the organism making sense of and adapting to the world (Di Paolo, 2005); it *is* the body making sense of itself and the world (Fuchs, 2017). Ultimately the enactive/embodied conception of human functioning is based on a relatively simple idea: our psychological functioning and sense of meaning is shaped in fundamental ways by our nature as biological organisms. As striving organisms we have needs, but further, we have a way of achieving these needs, within our contexts, based on our personal, cultural, and species-level histories (Gallagher, 2006; Thompson, 2007; Varela et al., 2017).

Before moving on, note that I am using the term 3e, when the term 4e is often used. I do so because I do not necessarily subscribe to the fourth 'e'—*extension* (where the mind is seen as partially constituted by the external environment; A. Clark & Chalmers, 1998). My reasons for this are multiple but I will briefly allude to them. Firstly, it is doubtable that

[6] Although it should be said that some enactivists would even reject this label—they fear that 'downward' still implies a hierarchy of importance rather than merely scale. This is a reasonable concern but I still find the imagery useful, hence I have compromised by using the term 'scale' rather than the traditional 'level' as will be discussed later.

full extension is compatible with enactivism and embodiment given that the latter two emphasize the process of continual separation between organism and environment (self-maintenance), while extension de-emphasizes this (Maiese, 2017). Secondly, enactivism holds that meaning is always relational—it is generated by an organism through its needful relation with the world (Fuchs, 2017; Thompson, 2007). The constitutional boundaries of the organism become blurry and ever-changing under extension (Maiese, 2017), and this seems to make the nature of the enactive relation very unclear. Thirdly, for our purposes at least, subscription to embeddedness (rich and necessary *causal* relations between organism and environment), as opposed to extension (*constitutional* expansion), can achieve much of the same conceptual ends while allowing for clearer explanations. For example, it would be very hard to explain the depression of some client 'John' if we spend our time trying to decide where 'John-the-system' ended and his environment began. Fourthly, many brands of extension seem to rely on an information-processing account that I disagree with due to their running clearly afoul of the hard problem of content (for more on this see: Harvey, 2015; Hutto & Myin, 2012). Finally, Thompson and Stapleton (2009) show that once the concept of extension is cut to size in-order to fit with enactivism then genuine extension of the mind becomes a much less remarkable and quite rare phenomenon.

3.2 Previous Conceptual Work in Embodied/Enactive Psychopathology

In this section I will review the work of three authors who have made significant comment as to what the wider concept of mental disorder looks like from a 3e perspective. I will first overview some core points made by Thomas Fuchs who has been instrumental in the shaping of enactive ideas about psychopathology, and has been publishing in this area since the late 2000s. This overview will focus on the recent summary of his perspective presented in his book Ecology of the Brain (2017). I will then review the recent work of Michelle Maiese and Sanneke de

Haan. These authors provide more explicitly complete frameworks for considering mental disorder from an enactive point of view.

It may be useful to note that both Maiese and de Haan have been publishing in this area during the development of the current project. As such, inevitable delays with publishing have meant that, until recently, the work of these two authors has been produced largely in parallel with my own work, and with each other. One exception to this is that I have previously compared my own framework to de Haan's and made suggestions as to how viewing our two frameworks in a complimentary manner highlights ways forward for a wider enactive perspective on mental disorder (Nielsen, 2020, 2021b). De Haan in turn has responded, claiming that our views are not compatible (de Haan, 2021; Nielsen, 2021a). A key intention of this book is to present an up-to-date summary of my '3e Psychopathology' framework, including acknowledgement of how it has been influenced by de Haan's venerable contributions. It is for this reason that I have placed a summary of these extant frameworks at this point in the text, rather than comparing and contrasting our views in a later chapter. A further reason for placing this section at this point is that it should hopefully allow the reader to flesh out their understanding of enactivism beyond the overview presented above.

I should also note that some authors have attempted to generate explanatory or descriptive models of *particular* mental disorders, grounded in or influenced by ideas of embodiment, embedment, or enactivism. Examples include models of the autism spectrum (De Jaegher, 2013; Gallagher & Varga, 2015), substance dependency (Zautra, 2015), schizophrenia and others (De Haan et al., 2013; de Haan & Fuchs, 2010; Fuchs & Röhricht, 2017; Krueger & Colombetti, 2018). I have chosen not to review these here as to keep the focus on the wider concept of mental disorder. A brief selection is reviewed in Nielsen and Ward (2018) if this is of interest.

Fuchs' Circular Causes and Dual Aspectivity

Thomas Fuchs' 'Ecology of the Brain' (2017) presents a vision of the embodied mind, with particular focus on understanding the role of the brain through an enactive lens. The central claim of this work is that the

brain is best thought of as an organ of mediation in the sensory-motor loop between organism and world, allowing for more complex and historically informed structures of meaning and response. Under such a view the brain does not determine what we do *per se* but is an evolutionary and socio-culturally shaped organ of integration, allowing for the condensation of our history and present perceptions into our sense-making and action. In developing this thesis Fuchs develops two core concepts that inform his understanding of psychopathology, which he then presents in his penultimate chapter. These two concepts are *dual aspectivity*, and *circular causality*. In this section I will present these two concepts before summarizing Fuchs' conception of mental disorder in 'Ecology of the Brain' (2017). I will then offer a brief commentary on how this conception relates to the perspective presented across the following chapters. As a passing note for clinicians and others who may be less familiar with philosophical terminology, please be aware that Fuchs is partially grounded in 'phenomenology'—the philosophical/scientific description of experience itself. This can bring with it some complex terminology and ideas which I have done my best to unpack, but please be aware that it is normal to find phenomenological language difficult to understand when one is unfamiliar with it.

Fuchs' notion of *dual aspectivity* is a way of understanding the apparent/traditional divide between the psychological and the physical, infamously referred to as the 'mind-body problem'. It is a way of understanding the relationship between conscious experience (e.g., thoughts/emotions/perceptions) and 'physical' states of the brain and body, as well as the relationship between the related phenomenological concepts of the 'living body' and the 'lived body'. Specifically, dual aspectivity is the idea of viewing these traditionally divided 'realms' as one and the same object observed from different perspectives and thus revealing different aspects. The living body and the lived body then, are two sides of the same coin—two different views on the same complex object as well as its actions and constituent processes. Further, both perspectives/aspects are necessary for a full understanding of the organism system. Under this view conscious experience is not a 'product' of the brain, emergent or otherwise. It is also not simply epiphenomenal—i.e., a useless by-product—because it is not a 'product' at all. Rather, claim's Fuchs, conscious experience is the

'integral' state of the organism-system experiencing itself in the world; the world and self as experienced from the concerned point of view of the self-maintaining organism, integrated through the resonant action of the central nervous system (see his concept of vertical circular causality below).

A second critical concept in Fuchs' analysis is that of *circular causality* (Fuchs, 2017, 2022). Fuchs breaks this down further into *horizontal* and *vertical* circularity. Horizontal circular causality effectively summarizes a previously recognized essential component of the enactive worldview; that of a closed causal loop between organism and world. Under enactivism/embedment the organism affects the world through its actions, but is also affected by it, resulting in a dynamic dance evolving across many timescales to shape behavior, organism, and world. Included under this concept are an organism's relations to others. Fuchs' concept of *vertical circular causality* meanwhile, pertains to part-whole circular causality in the constitution of the organism. Fuchs outlines an understanding of the dynamical constitution of the human organism grounded in acceptance of 'downward' causation via organizational/formal causality. Under this view, similar to that evoked in most enactive perspectives, structures at smaller scales (e.g., molecules) can have their interaction constrained by the organization of emergent structures at larger scales (e.g., cells), and this relationship continues to be present across larger and more complex scales of organization (e.g., organs, organism).[7]

This sort of ontology, where there is no hierarchy of fundamentality but only of scale, is a constant within enactive perspectives. Within it, Fuchs' concept of vertical circular causality presents a view where the brain is in constant 'resonance' with the rest of the body and plays a high-level integrative role. The brain is seen to integrate influences from sensory organs (both those responsive to the world, and those that are proprioceptive and/or interceptive), in a way shaped by what has worked in past experience, to produce approximately sensible action from

[7] A similar use of organizational causality to account for 'downward' causation can be seen in Bolton and Gilllet's rework of the biopsychosocial model, although theirs more explicitly incorporates notions of information (2019). In this way their model appears somewhat 'proto-enactive' however their particular use of 'information' seems to incorporate semantic content which seems in tension with an enactive perspective (Aftab & Nielsen, 2021).

moment to moment. If speaking to someone about their understanding of how the brain works, general members of the public will often express a linear sort of understanding whereby events in the brain, or parts of the brain, *cause* experience. For example, that perception of a certain stimulus *causes* activation of the amygdala, which then *causes* anxiety/fear and associated responses. In actuality, activation in the amygdala is but one aspect of a wider flurry of neural and bodily activity in response to threatening stimuli that resonates through the organism as it shapes a survival response (and attempts to pre-empt what will happen next). The biological processes that correlate with or constitute an experience of fear/anxiety are much broader than simply activation in the amygdala. They include for example wider bodily processes such as the release of cortisol and norepinephrine from the adrenal glands, vasoconstriction, and the shutting down of digestion. Activation in the amygdala then, is a correlate of our experience of fear (and many other experiences/functions) but we cannot by any means reduce fear to the amygdala. Rather, it is the experienced integral of such processes across the body that appears to constitute what we label 'fear'/'anxiety' (Colombetti, 2014; Fuchs, 2017).

Fuchs' notion of vertical circular causality offers a way to understand that actions of the organism are both caused by proximal/efficient biological events, and by the wider organization of the organism through processes of constraint and facilitation. Fuchs gives an example of producing speech, where the movement of one's larynx and tongue is proximally caused by the release of acetylcholine resulting in muscle contractions. At the same time, we can meaningfully say that these muscle contractions are shaped/constrained by the superordinate organizational processes of the organism, facilitating their organized production in a way that is meaningful to the organism and similarly enculturated others. Under this view it is the situated and historitical organism acting as a self-maintaining and integrated whole that shapes behavior, not the brain in isolation. Of course, the brain remains a remarkable integrative and plastic organ that is a key part of this process of making sense of and acting sensibly in the world. Hence, Fuch's title: *ecology of the brain*. In summary, Fuchs notions of horizontal and vertical circular causality represent an unpacking of the biological constitution of sense-making processes in the human organism. When paired with his notion of dual

aspectivity we can see the full ramifications of our embodiment: We are the emergent and dynamic organization of our body over time—I am my body in action.

From this view of human functioning Fuchs briefly presents a conception of mental disorders as disruptions in normal processes of both vertical and horizontal circular causality, or more simply as "unfavorable patterns of perception and reaction" (2017, p. 277). He firstly outlines how dis-integration of normal vertical circularity between neural and molecular process, the brain, and the wider organism, result in the dysregulation of certain functions and the correlated aspects of our experience, such as persistent feelings of heaviness and low motivation, or sudden feelings of panic. He emphasizes how these disruptions are primarily to/part of a person's self-experience (i.e., the experience of our own emotions, perceptions, thoughts, actions, and body) and as such often affect a person's understanding of themselves: "To a certain extent, a kind of 'self-division' or self-alienation is always implied. Something *inside me* confronts me, and yet is removed from my influence or otherwise manipulates me, be it a panic attack, a depressive mood, a compulsion, or a hallucination, while I vainly attempt to reinstate my sovereign control" (2017, p. 257). This is an important phenomenological observation that Fuchs makes, and such observations are a clear strength of his approach.

Secondly Fuchs describes how these dis-integrated experiences are barriers to the normal flow of functioning and attunement to our environment, including others around us. In other words, they are disruptions to our agency. In Fuchs' language there is a disruption in the normal horizontal circular causality that keeps the organism in lockstep with the world around it. In particular, Fuchs' emphasizes that these horizontal disruptions often take the form of feedback loops which support the maintenance or exacerbation of the dysregulations in experience and functioning. For example, a lowering in motivation results in less social contact, meaning less positive interaction further lowering of mood and motivation.

In essence, Fuchs rides a line between biological psychiatry on the one hand, and an intentional psychiatry on the other, as under embodiment these two positions can be understood as dual perspectives on the same

complex organism. Fuchs goes further than this however and expresses some degree of favor toward the intentional side of this dichotomy of understanding. He claims that sub-personal things such as brain states, neural mechanisms, and pharmacological agents (such as psychoactive medication) cannot be said to *cause* experiences on this view. This appears to be for two reasons. Firstly, to label such micro-scale and proximal events as causal is to abstract away from the historical context (of meaning) that shaped the very neural architecture supporting the experience/behavior produced. Secondly, Fuchs argues, if experience is the integral of the entire organism-system, then it is always 'caused'/enacted by the system-as-a-whole. A single perturbation in the system cannot be singled out as a privileged 'cause' because the way this perturbation cascades through the rest of the system (or not) depends on the nature and dynamics of the system itself. This seems to lead Fuchs to privilege analysis of mental disorders through the lens of experience and intentionality, and endorse in particular, methodologies such as phenomenology and hermeneutics in the study of mental disorders.

Grounded in this analysis of mental disorder, Fuchs pays some attention to what scientific understandings of mental disorders will require (he also offers an analysis of different forms of therapy which I have not unpacked here and is well worth the read). His primary claim is that this view of mental disorder demands "…a polyperspectival approach" (2017, p. 276). While he does not unpack this much further, what Fuchs appears to be calling for here is a *plurality of methodological approaches*, equating to rich *multi-scale and multi-perspectival* analysis distributed across different researchers. He also calls for a focus on circular causality operating across all scales of organization. However, based on his argument for the primacy of experience and the intentional scale of understanding, he at times appears to suggest the primacy of phenomenological and hermeneutical research methods: "This established the fundamental importance of subject-oriented psycho-pathology…" (2017, p. 277).

By and large, the analysis of mental disorder presented by Fuchs in his penultimate chapter of 'Ecology of the Brain' aligns with my own view which I will begin presenting in the next chapter. There are of course subtle differences. Where Fuchs uses the language of 'vertical circular causality' I prefer the language of 'organizational causality' and

'diachronic constitution'. I believe this is somewhat plainer and more established language that will more readily facilitate communication across different academic domains. Essentially though, I believe these terms are near synonymous. There are perhaps other points of subtle tension. For example, Fuchs speaks of the *fundamental* role of experience-focused methodologies, where I see these as *essential* yet no more fundamental in an explanatory sense than other perspectives. Similarly, he takes issue with the labelling of effects of psychotropic medication as 'causal' in that it seems to ignore the complex context and intentionality of the organism. In short, I think this says more about what we want to mean by the word 'cause' than anything of practical import in the analysis. With these very minor exceptions aside, I have very little 'critique' of Fuchs' work to make. The main challenge that his view faces is simply that it is presented in a very compact form, and often using reasonably complex language. Fuchs' (2017) analysis is focused more on the wider human organism and what mental disorder represents within it, and as such he does not fully explicate a conceptual model of what he takes mental disorder to be. The model of mental disorder that I will present in later chapters is therefore not intended to be in conflict with Fuchs' view, but rather in parallel compliment to it. I will present my alternative perspective on mental disorder through the enactive lens, which aims to provide further insight as to the likely structure and normative basis of disorder under an enactive worldview.

De Haan's Enactive Psychiatry

In her book 'Enactive Psychiatry', Sanneke de Haan presents a fascinating exploration of mental disorder from an enactive perspective, in particular arguing that enactivism has potential in helping to address 'the integration problem' in psychiatry (de Haan, 2020). By 'the integration problem' de Haan refers to an observation that we know of so many different causal factors at play in psychopathology—from social stressors to genetics—but, she claims, we have no clear way to consider how these factors come together to produce the disorders we recognize. In this section I will present a summary of some of de Haan's central arguments and the model

of mental disorder that emerges. I will then offer a brief critique of her view, highlighting areas of potential development. The reader may note many parallels between de Haan's work and the work of Fuchs previously explicated. De Haan's work however presents a much more extended analysis, with her entire book focused on the concept of mental disorder. She thereby presents a fuller and more specified framework of understanding, allowing greater room for critique. Some parts of this section are parallel to Nielsen (2021a, 2021b).

As noted, de Haan's concept of disorder is presented as a possible way to integrate the many kinds of causes at play in psychiatry—or what she refers to as 'the integration problem'. In structuring this argument she proposes that mental disorders can be considered to be composed of four dimensions. These are the physiological, experiential, socio-cultural, and existential. The first three of these are essentially the recognized domains within the biopsychosocial model. The existential dimension meanwhile is something unique and central to her framework (de Haan, 2017). Put simply 'existential' in this context refers to self-reflection—our capacity to evaluate ourselves and how we live our lives. The existential dimension then, refers to the aspects of mental disorder concerning self-understanding, values, and self-judgement. Later in this section I will unpack de Haan's concept of the existential dimension further as it seems to mark a point of difference between de Haan's perspective and my own. For now, as a first pass over de Haan's framework, her wider argument is that an enactive approach to psychiatry has the potential to integrate all four of these dimensions/aspects of human functioning and disorder. This is essentially because under the enactive perspective they are already interwoven, as different aspects of the same complex whole—i.e., a person standing in relation to their environment.

Outside of this overarching argument for the ability of an enactive psychiatry to integrate physiological, experiential, socio-cultural, and existential aspects of mental disorder, there are three key features of de Haan's framework that are important to understand. These are: (1) the central description of the structure of mental disorder, (2) the normative basis for demarcating disorder, and (3) her 'existentialised' understanding of enactivism. I will now unpack these in turn.

Regarding her central description of the structure of mental disorder, one of de Haan's core claims is that mental disorders are best understood as *biases in sense-making*: "…[Mental disorders] refer to cases in which the evaluative interactions of a person with her world go astray. These interactions may include the person's thoughts, feelings, and/or behavior—towards the world and/or to herself" (de Haan, 2020, p. 234). De Haan further specifies that within her framework the observed bias in sense-making has to be 'stable' in order to count as genuine disorder: "…a single instance of inadequate sense-making does not yet amount to a disorder. Psychiatric disorders refer to a more or less stable pattern in how someone's sense-making goes astray over time" (de Haan, 2020, p. 234). De Haan's central conceptualization of 'mental disorder' then is that it represents a global condition in which a person's understanding of and engagement with the world is significantly "…*biased in a specific direction:* the world appears overly threatening, or meaningless, or meaningful, or chaotic" (2020, p. 234).

In specifying the normative basis of her account, de Hann leans away from the idea that there is one meaningful difference that defines disorder or that makes one's sense-making suddenly 'biased enough' to be considered disordered. Instead, she uses a Roschian/Wittgensteinian approach where disorder is distinguished by a cluster of characteristics. She lists four characteristics of pathological sense-making with none being seen as necessary or sufficient (2020). These four characteristics are: (1) pathological sense-making is often 'inappropriate' in the context ('appropriateness' is described as being assessable in contrast to current socio-cultural norms—i.e., does it conflict with 'common' sense?); (2) pathological sense-making is often 'inflexible', i.e., the person acts the same way even in contexts when the action is not adaptive; (3) pathological sense-making often involves inflexible stance-taking, i.e., the person finds it very difficult to see/imagine things another way; (4) pathological sense-making often results in suffering.

One final feature of de Haan's framework that is important to understand is her underlying view of enactivism. De Haan argues that a standard enactive approach fails to capture the existential nature of the human experience. By this she means that as reflexive creatures we not only make sense of the world but make sense of ourselves. In other words, we have

the ability to recognize and evaluate ourselves: "The 'existential dimension' refers to the dimension that opens up due to the capacity to relate to our experiences. That is, we do not just experience things but we can also take stances on these experiences, on our ourselves and on our situation" (de Haan, 2017, p. 528). De Haan sees this capacity to think existentially as a vital part of the human experience and uses it as a basis for her understanding of the emergence of more complex and less immediately metabolic/functional value structures which she refers to as *existential values*. By existential values de Haan refers "…to what motivates certain actions: actions that are not motivated by the drive to stay alive, but rather have to do with living a good, meaningful, or dignified life" (2020, p. 193). The 'existential dimension of psychiatry' therefore refers to the way that sufferers of mental disorder understand and perceive themselves—how they consider their own existence, and the role this can play in their distress. De Haan is clear that she sees psychiatry as permeated with existential features, both in the constitution of mental disorders, and in the sorts of impacts that may result from and indicate disorder. We can consider many different examples of this, such as the deep feelings of guilt and perception of the self as a burden often seen in and maintaining depression, someone who comes to identify with being anxious as a part of their identity, or how the egodystonic thoughts and compulsions of OCD may hinder a person's expression of the kind of person they want to be.

In sum, de Haan's framework presents mental disorders under the enactive view as 'structurally disordered patterns of sense-making'. In other words, as biased ways of viewing and engaging with the world, including how we view and respond to ourselves. Under this view disorder is variably attributed when someone's sense-making is overly rigid and inflexible to the needs of their context, deviates from others of a similar cultural background, and/or results in suffering. She further emphasizes that these patterns of biased engagement constitute significant alterations to our experience, necessarily involve our physiology, are deeply shaped by our socio-cultural histories and contexts, and are interwoven with our existential concerns. In other words, they are dynamic and developmentally informed patterns in our experience and behavior that keep us stuck acting in ways that often go against our needs, often

result in suffering, and often conflict with who we want to be. When summarized like this I maintain that there are deep parallels with the view I will present across subsequent chapters, although de Haan may disagree (2021).

Considering de Haan's framework critically, there is a clear strength in terms of face-validity. When properly understood, her central description of mental disorder as biased sense-making is hard to take issue with. On first reading, one may be tempted to doubt the seeming simplicity of this description. Questions may arise such as: 'a bias from what exactly?', 'aren't cognitive biases part of normal human functioning?', 'is this description overly focused on perception and experience versus action/behavior?', and 'how similar is this this to cognitive essentialism *a la* Cognitive Behavioral Therapy?'. Such critical thoughts are understandable, but are ultimately going to fall flat as they are grounded in an impoverished view of the concept of sense-making. As explored earlier in the chapter, 'sense-making' refers to an organism responding differentially to the world in accordance with its purposes. If an organism can respond differentially to the world, in accordance with its purposes, then it is no longer simply a passive object to be affected by the world in a linear fashion. Rather it demonstrates an ability to enact meaning, to strive to survive and adapt to the world. For the enactivist then, sense-making is the defining act of 'cognition' or of 'the mental'; to be a 'cognitive' system *is to be* a sense-making system. Further than this though, from the enactivist view it is through sense-making that we see the emergence of *agentiality*. There is, therefore, something importantly right about de Haan's claim here. Insofar as mental disorders can be understood as something about 'the mental' going wrong or 'astray', then *by definition* they can certainly be understood as biases in our sense-making processes.

Further, this description highlights that mental disorders are primarily disruptions to our agency—to our ability to act appropriately in accordance with our needs and in accordance with the kind of person we want to be in the world. Given the relational and embodied nature of sense-making, this central description also prompts us to begin considering how the constitution of mental disorders is distributed across brain, body, and environment. This is something I have also argued and will continue

to develop across the following chapters (Nielsen, 2020; Nielsen & Ward, 2018). De Haan's central description of mental disorder then, has clear face validity, and is useful in that if we are seeking to understand or explain mental disorder, it prompts us to think holistically.

There is however, room for improvement here. Imagine someone experiencing a disorder such as melancholic depression. De Haan's description would certainly align with their experience and the kinds of behaviors and presentation we might observe. If we were in a therapeutic context, de Haan's concept would also prompt us to ask this person about their experiences and existential concerns, how they view themselves and so on, and this would undoubtedly be useful. Further, if seeking to understand what maintains this person's depression, and what brought it about in the first place, de Haan's view itself points us to think broadly. There are likely a multitude of different influences across this person's brain, body, environment and history that we would need to consider. However, while de Haan's view encourages us to recognize this, and lays an ontological groundwork through enactivism for these factors to all sit alongside each other, her view does not seem to provide sufficient guidance as to how to go about discovering and conceptualizing the interrelation between such influences and the dynamic emergence and maintenance of the dysfunctional sense-making we observe. In short, her description seems to capture the surface and experience of many mental disorders, but does not provide much guidance as to how we could seek to understand the dynamic constitution of such disorders.

Another key strength of de Haan's framework lies in her Roschian/Wittgensteinian approach to demarcating disordered sense-making from normal human functioning. The core strength here is one of practicality. By assessing for de Haan's four characteristics a clinician could conceivably make a reasonable estimate as to whether mental disorder is present, and I myself have found her characteristics a useful guide in this way. Again though, there is clear room for improvement here. Firstly, I disagree deeply with the idea that 'common sense' or social-cultural statistical normativity should have anything to do with defining behavioral pathology. This boils down to something akin to statistical functionalism, which I argued against in Chap. 2. In brief reminder, it is unclear why oddities or differences in our behavior should be used in any way to mark

us as disordered. As such, this characteristic seems unjustified. Secondly, while very practical, this characteristic-based approach where no particular characteristic is necessary or sufficient to define dysfunction leaves unspecified the question of where these characteristics come from. It is not clear why these characteristics should be used above others. Another way to phrase this concern would be to ask: are these characteristics simply observed features of what we currently label as mental disorders, or are they principled in some way? If we are to be labelling people and their behavior as dysfunctional or disordered, then having some underlying theoretically justified principle of guidance seems ethically necessary.

Finally, another strength of de Haan's framework is her incorporation of existential concerns. From the examples given throughout this section it should be clear that this feature of de Haan's framework captures something important about mental disorders, in that they often concern how we view and evaluate ourselves. This adds richness to de Haan's description and adds to its face validity. Elsewhere, I have argued against the idea that this requires a fundamental change to the underlying enactive framework (Nielsen, 2021a, b), but this is a detail of less import here.

In summary then, there are clear strengths to de Haan's view, but criticisms and areas for potential improvement also emerge. These are: (1) de Haan's framework presents perfectly good description of what mental disorders are, but provides little sense of how they might work (i.e., emerge and maintain) and how we should seek to investigate this. (2) The normative basis for defining disorder she provides is practical, but underjustified, and the inclusion of the 'common-sense' characteristic is problematic. Across subsequent chapters I will attempt to show how my own view parallels much of de Haan's work, but also attempts to address these lacunae.

Maiese's Enactive Medical Model (and One Tangent)

Maiese (2021) also presents an understanding of mental disorder from an enactivist worldview. Similar to de Haan she labels mental disorder as a kind of 'disrupted sense-making'. Parallel to Fuchs, de Haan, and myself, Maiese argues that enactivism affords a middle way between complete

medicalization on the one hand and seeing mental disorders as simply problems in living (in a deflationary sense) on the other—i.e., a collapse of the division between naturalism and normativism through the natural normativity of a living and socio-cultural system. Maiese also supports the broader claim being made, that enactivism provides an ontological framework that allows for the integration of physiological, experiential, environmental, and socio-cultural factors into the concept of mental disorder. In sum therefore, Maiese's (2021) framework seems much more alike the other enactive frameworks here reviewed than it is different. There are however two key points of difference to the positions so far considered. These are (1) Maiese's framework more explicitly engages with *autonomist sensorimotor enactivism*, and (2) Maiese's framework attempts to use the concepts of affordances, identities, and social expectations to normatively demarcate disorder. I will soon unpack these differences in order to help explicate Maiese's view, however, this will first require a brief overview of autonomist sensorimotor enactivism and related ideas. This foray will be of use in later chapters, hence I have separated it off as its own subsection. I will then return to offer a brief critique of Maiese's framework and highlight some useful moves she makes.

A Brief Foray into Autonomist Sensorimotor Enactivism. Autonomist sensorimotor enactivism[8] is built upon the idea that higher levels of autonomy/identity can emerge within the functioning of a biological organism in context, above and beyond that of self-maintenance. As a school of thought it appears to be not in conflict with, but rather a further development upon, the autopoietic core of enactivism that I have so far focused on. Authors such as Di Paolo (2005, 2010) convincingly argue that it is a necessary development upon an enactivism grounded solely in the concept of autopoiesis/self-maintenance. Essentially it is the idea that organisms do not only feature a self-maintaining dynamic in terms of their material organization or form, but that similar

[8] Note that I am here describing *autonomist* sensorimotor enactivism defended and defined by Baradiaran (2017) where the notion of autonomy remains central to the emergence of sensorimotor patterns/identities, as this is the model that Maiese (2021) appears to be engaging with. This is separate to sensorimotor enactivism more broadly, whereby there is not necessarily a focus on operational closure/autonomy, and instead focus is given to regularities in sensorimotor experience, e.g., O'Regan and Noë (2001).

self-supporting dynamics also present in their *ways of doing so*—i.e., in their ways of perceiving and acting in the world. When exploring the work of Fuchs earlier in this chapter I highlighted how the idea of a closed sensorimotor loop between organism and world is a central idea for all enactivists. An autonomist sensorimotor enactivist takes this closed loop and claims that this represents a further opportunity for operational closure (or recursivity), marking the emergence of a new level of at least partial autonomy/identity beyond the organismic/biological—i.e., the emergence of sensorimotor or cognitive *patterns*, or even '*selves*'. This can sound somewhat outlandish to those unfamiliar with it, but there is a certain face validity to the idea. Consider a case of 'brain-death' for example, where sufficient damage to the central nervous system results in the collapse of coordinated thought and action despite the continued (medically assisted) self-maintenance of the wider organism. Without decaying into debates about the metaphysics of identity, there is a reasonably agreed upon sense in which, whilst much of the organism continues, the 'person' is no longer present in this case. Autonomist sensorimotor enactivism can help us make sense of such an example without recourse to dualistic thinking, in that, what is seen to be lost in such a case is not a 'spirit' or other substance, but a coordinated pattern of activity synonymous with personhood.

For the sake of clarity here, there is a distinction I wish to introduce. It will be helpful to make a loose distinction between a claim for the existence of *partial* and/or *temporarily* operationally closed processes emerging *within* embodied functioning, and a bolder claim for the existence of cognitive or sensorimotor 'selves'/'identities' more distinct and independent from our biological existence. Henceforth I will refer to 'conservatively autonomous sensorimotor enactivism' [CASE] and 'radically autonomous sensorimotor enactivism' [RASE] respectively. Certainly, under both RASE and CASE variants, the claim is not for the emergence of a new ontological substance, but rather for the emergence of identifiable processes. As I have understood it, the claim of most sensorimotor enactivists is more in line with the CASE interpretation. The claim does not generally seem to be for emergence of a wholesale new 'identity'/'identities' completely independent from the biological organism and the normativity that arises from its autopoietic structure.

However, some passages from sensorimotor enactive authors do seem ambiguous in this sense and seem to lean toward the RASE interpretation. For example: "Within this independence, the new form of life will be able to generate, via a process of adaptive closure analogous to metabolism, its own set of values, thus making the process irrevocable and resulting in the *coexistence of different identities in a same organism*" (Di Paolo, 2005, p. 446).

For clarities sake, I would like to distance myself from this RASE interpretation from the outset. I am not a philosopher, and my intention in the current project is to utilize enactive thought to answer questions within my own field of psychopathology, not to critique established enactive views. However, some degree of critique must be present in the selection of conceptual tools available, and my intuition is that the RASE claim drifts too far from the principle of embodiment and appears to ignore that 'higher level' autonomous processes must themselves exist within the situated sense-making processes of the biological organism. Briefly, there also appears to be a danger in the RASE interpretation of a '*promiscuous normativity*' whereby values and norms populate unchecked and float ungrounded, where we thereby have no way to prioritize one set of norms or values over the other. This will make it very hard to evaluate human functioning on its own terms. For the purposes of the current project this could conceivably lead to some bizarre places, such as having to consider whether a self-reinforcing pattern of anorexic behavior constitutes a 'self'/'identity' and thereby whether it has some moral claim on us that needs to be considered before seeking to help a person struggling with disordered eating.

The CASE claim however, seems much more reasonable, and as I hope to show, is potentially very helpful. Essentially the CASE claim is that our sensorimotor functioning in the world, through its circular structure, can support the emergence of *self-supporting patterns* identifiable through their operational closure/recursivity. For example, consider a well-rehearsed musician describing the experience of the music 'playing itself', or the way that habits such as nail biting can take on a *metaphorical* 'life of their own'. The claim therefore is in no way dualistic or anti-embodiment, as these patterns exist *within* the embodied sense-making of the organism. The CASE claim is therefore a reasonably modest

expansion beyond an enactivism grounded purely in autopoiesis—if such a pure view was ever actually intended, see Varela (1992, p. 10, as quoted by Barandiaran, 2017). In many ways CASE is simply an acceptance of the messy reality that "biological and cognitive norms and identities are not always coextensive" (Barandiaran, 2017, p. 414). Under CASE then, 'mind' is in, and deeply connected to, but not equivalent to life when considering the normative structures that arise. Put another way, our mental/behavioral lives are filled with habits and hobbies that are not of immediate biological relevance (although I would add, these things often indirectly support our biological functioning, or can be understood as side effects of other adaptive capacities). Such habits can be understood as self-sustaining patterns within embodied functioning, and can even metaphorically be understood as having a 'life of their own' (Barandiaran, 2017; Di Paolo, 2005, 2010; Maiese, 2021). Importantly however, under CASE, this 'life' remains metaphorical. Under CASE the normativity that arises through the self-sustaining/recursive structure of these sensorimotor patterns/habits is *secondary* to the more fundamental normativity of the biological organism striving for survival. This appears to put a stopper on issues of promiscuous normativity, making CASE a much more useful and sustainable tool for our purposes. From this point on then, when I speak of autonomist sensorimotor enactivism I am referring to the CASE interpretation unless otherwise specified.[9]

The utility of autonomist sensorimotor enactivism for the purposes of conceptualizing psychopathology is that it opens up the idea that there may be conflict between different levels of autonomy within the functioning of an organism:

> "…the relation between different self-sustaining processes enabled by a substrate of autopoiesis need not be one of perfect harmony and that, on

[9] One further thing to note is that by accepting the emergence of autonomy as a scale other than the biological, autonomist sensorimotor enactivism does 'open the flood gates' to a wider plurality of 'partial autonomies'. This includes ideas such as that of emergent autonomy within interpersonal interaction (e.g., consider musicians being 'in the groove' with each other, or the experience of a truly engaging conversation that takes on a life of its own), or perhaps linguistic/narrative autonomies, opening up a pathway to an embodied study of memetics. For exploration of this idea see Di Paolo et al. (2018). I am not opposed to these ideas but see them as reasonably speculative at this stage.

> the contrary, the inherent regulative tendencies of sophisticated processes of identity generation are likely to enter into conflict even with basic metabolic values….habits, as we know, can be "bad." The question is that as self-sustaining structures, they are never bad for themselves, but for some other identity (typically, in the case of humans, a combination of the metabolic and socio-linguistic self)." (Di Paolo, 2010, p. 54)

As the above quote hopefully highlights, autonomist sensorimotor enactivism leaves room for conflict between the norms of an organism and the habits/patterns of activity it engages in. It offers a way to see these patterns as having recursive organizational process structures of their own, nested within the functioning of the wider organism. This is a strength that Maiese (2021) draws on, and that I will draw on in later chapters.

One further idea presented within autonomist sensorimotor enactivism is that these 'habits' can be understood at multiple levels of constitutional complexity and specificity. We can understand individual habits/actions in this way, e.g., nail-biting. We can also consider bundles of actions/habits that support each other, e.g., the collection of simultaneous and precariously coordinated actions enacted in the masterful playing of a musical instrument (Barandiaran, 2017). Is has also been suggested that loose/flexible bundles of habits may represent one way to understand the way that we behave differently when we are taking on different social roles/expressions of identity. The idea being that we enact a different set of interrelated habits in different contexts to help us meet our needs/goals in different kinds of situations, not just as organisms, but as the particular identity/bundle-of-inter-related-habits we are current enacting. Consider for example how someone may act very differently when they are entering a hospital situation as a patient, as a doctor, or as a parent. These more flexible 'bundles of habits', or perhaps 'bundles of bundles', are referred to as 'regional identities' (Di Paolo, 2010; Maiese, 2021). As we shall now explore, this idea plays a key role in Maiese's (2021) conceptual framework.

Returning to Maiese. As stated, many of Maiese's (2021) claims align with the core theses of the other enactive frameworks here reviewed (as well as my own). To review, these are that: (1) mental disorders can be understood as disruptions to sense-making, (2) this allows for a holistic

and integrated/embodied view of mental disorder, (3) this also allows for a middle way between naturalist and normative views. Beyond these, Maiese's central claim is that in order to count as mental disorder, disruptions to sense-making must be *maladaptive*. This builds upon my own claim that mental disorders can be distinguished by the self-violation of the (physical *and* socio-cultural) functional norms of an organism, which I will expand on in the next chapter (Nielsen & Ward, 2020). However, Maiese understands adaption/maladaption in a particular way. Given the kinds of complex socio-cultural creatures we are, Maiese perceives adaption not only in reference to the continued functioning of the organism-system but "…with reference to "faring well" in a particular sociocultural setting and maintaining the viability of particular regional identities" (2021, p. 973). However, this too Maiese understands in a particular way. For a pattern of behavior to be maladaptive in her view means that it leads to "…the destabilization of particular regional identities or the self as a whole" (2021, p. 974).

Maiese (2021) specifies that what she means by this is best understood in reference to the notion of *affordances* and *attunement*. An affordance is a concept from ecological psychology that refers to a possible action, arising from the relationship between an organism and its environment (Gibson, 1977). This idea highlights that the sorts of actions we perceive as possible and/or salient depend on the kind of bodies, histories, and current purposes that we have (De Haan et al., 2013; Nielsen, 2022). For example, a dainty teacup affords different actions for: (1) a human with fine fingers versus a human with stubby fingers, (2) a human with a history of practicing tea drinking etiquette versus one who has not, and (3) a human who is looking for something to throw at a spider they are terrified of versus one who is looking for something to put tea in. For most people in most circumstances however, they seem likely to perceive the handle of a teacup as 'for holding' because of the way it is designed to (apparently) fit the human body. Compare this for example to a mouse, who may perceive the teacup as something to hide behind or sleep in. Our bodies, histories, and intensions, shape our perception of the world, particularly as it is *perceived for doing* (Nielsen, 2022). Maiese (2021) takes this idea and combines it with the notion of regional identities explored earlier. As we enact different social roles/modes of being, our

perception of and engagement with the environment should shift in a concordant way in order to support the fluid expression and maintenance of this role, and facilitate the achievement of our different goals as we move between contexts. This is what Maiese appears to mean by *attunement*. She states: "…ongoing attunement allows the agent to be in equilibrium with her surroundings and involves a complex dynamical interplay between the subject's living body and the environment" (2021, p. 976). Fluid attunement then, refers to being able to do what you need to do in your relevant contexts to serve your needs and maintain the role/regional identity you are enacting.

Consequently, under Maiese's (2021) framework 'mental disorder' refers to maladaptive patterns of sense-making that result in 'dis-attunement'—i.e., a disruption in the fluid recognition of and response to affordances, so that we do not manage to achieve the demands of the role we are trying to fill and/or the way of being we are trying to maintain. For Maiese then, mental disorder is "a breakdown in the agent's ability to regulate his or her coupling with the surrounding world, including the social world" (2021, pp. 976–977). More specifically, under her framework mental disorder appears to refer to the consistent wayward or 'dis-attuned' enaction of sensorimotor schemes, in ways that either destabilize our mode of functioning in a particular context, or that do not serve our continued functioning and faring well more broadly.

In critique of Maiese's (2021) framework, two clear strengths are apparent relative to the other enactive frameworks presented. Firstly, Maiese's framework begins to put focus on the *constitutional structure* of disorder itself. Through her claim that 'habits can take a life of their own' she invites the study of these habits/patterns as their own entities/structures. This seems advantageous as it encourages exploration of how mental disorders work, rather than simply what they are. Secondly, Maiese puts explicit focus on the idea of maladaption as the normative basis for claims of disorder or dysfunction within our mental/behavioral lives. This seems advantageous in comparison to de Haan's approach based on Rochian characteristics, as it attempts to offer the sort of justification I argued was missing from de Haan's framework in the previous section. I will now explore these strengths, emphasizing where there appears to be further room for improvement.

Considering the structural claims made by Maiese (2021), we can see that she develops upon the ideas of de Haan in a subtle way. She agrees with de Haan that mental disorders represent consistently biased sense-making. However, rather than speaking of sense-making as a whole, whereby the world is perceived as overly terrifying or overly meaningless, Maiese begins to offer a view from inside the processes sense-making itself. She speaks of habitual patterns of engagement conflicting with others, resulting in a collapse of fluid engagement with the world. There is a sense in which she is seeking to explore, not just *what* mental disorders are, but *how* they work and maintain themselves. This appears to be conceptually related to her talk of habits having a 'life of their own'—i.e., their own norms and processes of self-maintenance. This perspective begins to divide sense-making into smaller units or smaller patterns so that we may understand how consistent biases/maladaptive patterns arise and perpetuate themselves within our embodied functioning. If our aims in seeking to conceptualize mental disorder are roughly scientific as I laid out in Chap. 1—i.e., to better categorize, explain, and treat mental disorder—then this more mechanistic or entitative perspective seems advantageous.

It must be said however, that I cannot claim to be free from bias on this point. As I will present in the following chapter, my own view roughly parallels Maiese's (2021) on this point. One of my central claims has been that the structure of mental disorders can be understood as constituted by *relatively stable dynamic patterns within the brain-body-environment system* (Nielsen & Ward, 2018). Indeed, while I typically avoid the complex language of autonomist sensorimotor enactivism with its references to 'autonomies' and 'identities', I have previously attempted to 'take a dip' into such language, stating that:

> Mental disorders … [can be understood as] parasitic partial 'autonomies' within the process structures of human functioning (i.e., mental disorders themselves partially 'self-maintain' within the context of the brain-body-environment system). This autonomy is dependent on—but in tension with—the biological, sensorimotor, and other adaptive autonomies of the host organism, and conflicts with the normative structures that arise from these, to the detriment of the organism's adaptive agency and likelihood of

> survival. Such a description has value because it emphasizes the partial entitativity of mental disorder. (Nielsen, 2020, p. 117)

Other authors also seem to appreciate the advantage of such a partially entitative view. For example, a similar objectification of the disorder process has been presented by Ramírez-Vizcaya and Froese (2019). These authors first outline an embodied notion of a habit—"[habits are] adaptive, precarious, and self-sustaining network[s] of neural, bodily, and interactive processes that generate dynamical sensorimotor patterns" (Ramírez-Vizcaya & Froese, 2019, p. 1), and suggest that 'bad habits'—including several mental disorders—are those that tend to dominate the expression of other habits, thus reducing flexible attunement and conflicting with the needs of the wider organism and making them no longer adaptive.

The advantage objectifying habits or other such patterns/dysfunctions within sense-making is that it presents us with scientific purchase on the subject matter. Disorders/patterns that look similar in both experience and symptomology may be distinguished within our classifications if the underlying process structures that maintain them (i.e., mechanisms) are found to significantly differ. We can seek to explain mental disorders in reference to observed regularities within their process structures (i.e., mechanisms). And, most importantly, we can then seek to treat mental disorder based on such systematically developed explanations. Such efforts are more conceivable if the processes of psychopathology are already understood as having a constitutional structure rather than simply being an overall bias in our engagement with the world.

While it seems advantageous then to treat mental disorders as separable processes existing within our sense-making, one improvement that could be made to Maiese's (2021) framework would be to remove the language of 'identity'/'selves' in reference to partially autonomous patterns within sense-making. Such language evokes the RASE interpretation of autonomist sensorimotor enactivism, encouraging an understanding of these partially autonomous habits/patterns as more fully independent than they are, and de-emphasizing the fact that these patterns are deeply embedded within our wider organismic functioning. As highlighted in the previous sub-section, this may open up issues of

promiscuous normativity whereby the extinction of some habit or process of behavioral disorder may be considered to have a moral cost, a prospect that seems slightly absurd. The use of words such as 'selves' or 'identities' also seems likely to muddy the water when attempting to consider psychopathology that exists within/overlaps with issues of identity. Variously consider for example, someone who begins to identify as someone who experiences depression in a way that begins to perpetuate their difficulties, someone who through developmental experiences develops an implicit understanding of themselves as disgusting or unlovable, or someone who fails to develop the general sense of 'I'll be ok' that gets many of us through difficult times. Seeking to understand such patterns of sense-making while referring to the patterns themselves as 'selves' or 'identities' seems needlessly confusing. Further confusion develops when we ask the question 'which regional identifies are important?'. Maiese (2021) gives the example of failing to practice one's instrument as a way for a regional identity as a musician to break down, but surely it is not her intention to say that failing to practice one's instrument constitutes a mental disorder. Many of these difficulties can likely be resolved but, in sum, the language of 'identity' and 'selves' carries far too much weight and conceptual confusion in its common usage. I worry this will scare off many non-philosophical researchers from engaging with Maiese's (2021) framework, negatively impacting its ability to influence scientific discourse.

As stated above, a second key strength of Maiese's view is her emphasis on maladaption as the basis for claims of disorder or dysfunction. Maiese presents this as a development of my previous work, and the general thrust of the claim—that mental disorder must be maladaptive for the individual diagnosed—certainly aligns with my perspective (Nielsen & Ward, 2020). This is a strength relative to de Haan's normative conception based of Roschian characteristics which as argued earlier is somewhat under justified. However, in terms of how this is achieved within Maiese's (2021) framework, a reasonable concern would be that too much weight is given to the role of social expectations.

Maiese (2021) rightfully considers that social context plays a huge role in whether a behavior works for someone or not. As such, being able to attune to the needs of the social context is an important part of human

functioning: "A subject who fails to gauge relevant action-possibilities or engages inappropriately in a recurrent, patterned manner exhibits sense-making that is insufficiently grounded in her situation. Because sense-making always is embedded in a particular socio-cultural context, what counts as adaptive functioning needs to take into account this context" (2021, p. 978). At times however, she leans towards pathologizing non-conformity: "Adaptive agency in a complex social world such as ours requires that subjects reliably act in ways that fit with sociocultural practices, customary expectations, and situational demands" (2021, p. 977). While Maiese is clear that she does not wish to label disorder simply on the basis of statistical abnormality, she does describe what could be labeled a form of behavioral statistical functionalism along the lines of Bergner (1997, 2004). This is demonstrated in the following passage:

> "However, behavior is not disordered simply by virtue of not being "normal" or expected (since this might lead us to label homosexuality as a disorder); rather, it needs to be incapacitating in some way. The notion of disordered sense-making and the destabilization of regional identities can capture this notion that persons are unable to do what they are supposed to be doing, even as what they are supposed to be doing is defined partly in relation to social expectations or their own standards. Subjects *falter* in the sense that they cannot adapt to changing circumstances or perform activities that are central to the maintenance of their social roles (regional identities)." (2021, p. 980)

The concern here is that utilizing the meeting of social expectations as a measure of adaption places the boundaries of disorder in inappropriate places. For one, norm breaking can be functional. If a young woman in a small, conservative, and gender-role conforming town begins reading feminist texts and decides to break with the social expectations of her community to the point that she is ostracized, I do not want to label such a woman as disordered. Such a case seems to be an example of expressing effective social agency. We can also consider clear cases where not being attuned to the requirements of social contexts does not count as disorder, such as tourists failing to adapt to local social norms and customs thereby acting in ways that incite negative evaluation.

A less trivial example of non-pathological failure to meet social norms would be that of autism. Current arguments related to the neurodiversity movement hold that many autistic people should not be considered to have a disorder. Under such a view, many instances of autism appear not as inherently disordered but as a different kind of functionality or mode-of-being, with a different set of capacities and challenges (likely opening up a different landscape of potential psychological difficulties). Many of these arguments are convincing (for example, see Chapman, 2021). A common feature of autism however, arguably present for most autistic people, is having difficulties understanding and following the social norms of the wider neurotypical populace. These kinds of views will be very difficult to rectify with frameworks such as presented in Maiese (2021) given the emphasis on following social expectations. Maiese's (2021) emphasis on maladaption seems well justified, and she is right to consider that social context must always be considered. However, that which is weird or unexpected, or harmful only because of unjust stigmatization, cannot fall under the banner of mental disorder. To justifiably claim that an aspect of someone's psychological functioning is disordered then this claim needs to be made in reference to the needs and purposes of the person being diagnosed (Nielsen & Ward, 2020). This is an idea that will be picked up on in the next chapter.

3.3 Summary

This chapter has first introduced and overviewed 3e cognition/enactivism, and then reviewed the conceptual frameworks presented by Fuchs (2017), de Haan (2020), and Maiese (2021). Outside of my own work these are the only authors I am aware of that present complete frameworks for considering the wider concept of mental disorder from an embodied, embedded, and enactive perspective. All three perspectives seem to agree that: (1) mental disorders can in some sense be understood as disruptions to sense-making (i.e., our meaning-oriented perception of and engagement with the world), (2) this allows for a holistic and integrated/embodied view of mental disorder, (3) this also allows for a middle way between naturalist and normative views.

Fuchs' (2017) framework has particular strengths in the incorporation of phenomenological observations and his presentation of a detailed wider framework for considering embodied psychological functioning/dysfunction. He labels mental disorders as unhelpful patterns of engagement it the world, and this lays important groundwork for the other author's frameworks. However, his exploration remains somewhat compact and general. De Haan (2020) builds on Fuchs' seminal ideas and was the first to summarize mental disorders under the enactive worldview as 'biases in sense-making'. She demonstrates how enactivism provides an ontological perspective through which biological, experiential, social, and existential factors can be integrated. I have here presented concerns that: (1) de Haan's framework presents a good description of what mental disorders are but provides little sense of how they emerge and maintain—i.e., of how they work, (2) the normative basis for defining disorder she provides is practical, but under justified, and the inclusion of the 'common-sense' characteristic is problematic. Maiese's (2021) framework in turn builds upon de Haan's ideas. She agrees that mental disorders are disruptions to our sense-making, but takes a somewhat more entitative view whereby mental disorders can be understood as habits (or collections of habits), existing within our sense-making, that disrupt our ability to fluidly attune to the needs of our environment and our ability to fulfill our social roles/regional identities. Through doing so she seems to draw attention to the constitution/process structures of disorder in a fruitful way. However, I have noted a concern with some of the language and conceptual confusion that this engagement with autonomist sensorimotor enactivism has brought with it. Specifically, she engages with confusing talk of patterns of behavior as 'selves' or 'identities' in their own right. Maiese (2021) seems right to shift away from de Haan's Roschian approach to demarcating disorder and toward a concept of maladaption. However, I have here presented a concern that, in doing so, she places too much emphasis on social expectations resulting in some inappropriate boundary lines around the concept of mental disorder.

In the following chapter I will begin outlining my own framework of 3e Psychopathology as it currently stands, beginning with an outline of the conceptual tools present within 3e cognition. I will attempt to show how this framework addresses the concerns raised here.

References

Aftab, A., & Nielsen, K. (2021). From Engel to Enactivism: Contextualizing the biopsychosocial model. *European Journal of Analytic Philosophy, 17*(2), M2–M22.

Andersen, H. (2016). Biomedical sciences. *The Routledge Companion to Philosophy of Medicine, 81*.

Barandiaran, X. E. (2017). Autonomy and enactivism: Towards a theory of sensorimotor autonomous agency. *Topoi, 36*(3), 409–430.

Bechtel, W. (2009). Explanation: Mechanism, modularity, and situated cognition. *The Cambridge Handbook of Situated Cognition*, 155–170.

Bergner, R. M. (1997). What is psychopathology? And so what? *Clinical Psychology: Science and Practice, 4*(3), 235–248.

Bergner, R. M. (2004). An integrative framework for psychopathology and psychotherapy. *New Ideas in Psychology, 22*(2), 127–141.

Bolton, D., & Gillett, G. (2019). *The biopsychosocial model of health and disease: New philosophical and scientific developments*. Springer Nature.

Brigandt, I. (2013). Explanation in biology: Reduction, pluralism, and explanatory aims. *Science & Education, 22*(1), 69–91.

Chapman, R. (2021). Neurodiversity and the social ecology of mental functions. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620959833

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19.

Colombetti, G. (2014). *The feeling body*. The MIT Press.

de Haan, S. (2017). The existential dimension in psychiatry: An enactive framework. *Mental Health, Religion & Culture, 20*(6), 528–535.

de Haan, S. (2020). *Enactive psychiatry*. Cambridge University Press.

de Haan, S. (2021). Two enactive approaches to psychiatry: Two contrasting views on what it means to be human. *Philosophy, Psychiatry, & Psychology, 28*(3), 191–196.

de Haan, S., & Fuchs, T. (2010). The ghost in the machine: Disembodiment in schizophrenia–two case studies. *Psychopathology, 43*(5), 327–333.

De Haan, S., Rietveld, E., Stokhof, M., & Denys, D. (2013). The phenomenology of deep brain stimulation-induced changes in OCD: An enactive affordance-based model. *Frontiers in Human Neuroscience, 7*, 653.

De Jaegher, H. (2013). Embodiment and sense-making in autism. *Frontiers in Integrative Neuroscience, 7*.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences, 4*(4), 429–452. https://doi.org/10.1007/s11097-005-9002-y

Di Paolo, E. (2010). Overcoming autopoiesis: An enactive detour on the way from life to society. In *Advanced series in management*. Emerald Group Publishing Limited.

Di Paolo, E., Cuffari, E. C., & De Jaegher, H. (2018). *Linguistic bodies: The continuity between life and language*. MIT Press.

Fuchs, T. (2017). *Ecology of the Brain: The phenomenology and biology of the embodied mind*. Oxford University Press.

Fuchs, T. (2022). Understanding as explaining: How motives can become causes. *Phenomenology and the Cognitive Sciences*, 1–16.

Fuchs, T., & Röhricht, F. (2017). Schizophrenia and intersubjectivity: An embodied and enactive approach to psychopathology and psychotherapy. *Philosophy, Psychiatry, & Psychology, 24*(2), 127–142.

Gallagher, S. (2006). *How the body shapes the mind*. Clarendon Press.

Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford University Press. https://books.google.co.nz/books?id=Z28sDwAAQBAJ

Gallagher, S., & Varga, S. (2015). Conceptual issues in autism spectrum disorders. *Current Opinion in Psychiatry, 28*(2), 127–132.

García Otero, E. (2022). *Participatory sense-making in psychotherapy*, PhD, University of the Basque Country/Universidad del País Vasco. http://hdl.handle.net/10810/56213

Gibson, J. (1977). The concept of affordances. *Perceiving, Acting, and Knowing*, *1*.

Harvey, M. I. (2015). Content in languaging: Why radical enactivism is incompatible with representational theories of language. *Language Sciences, 48*, 90–129. https://doi.org/10.1016/j.langsci.2014.12.004

Hutto, D. D., & Myin, E. (2012). *Radicalizing enactivism: Basic minds without content*. MIT Press.

Hutto, D. D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. MIT Press.

Kaplan, D. M. (2015). Moving parts: The natural alliance between dynamical and mechanistic modeling approaches. *Biology & Philosophy, 30*(6), 757–786.

Krueger, J., & Colombetti, G. (2018). Affective affordances and psychopathology. In *Philosophical perspectives on affective experience and psychopathology: Vol. XXVIII–2* (pp. 221–247). Quodlibet.

Maiese, M. (2016). *Embodied selves and divided minds*. Oxford University Press. https://books.google.co.nz/books?id=w_quCgAAQBAJ

Maiese, M. (2017). Can the mind be embodied, enactive, affective, and extended? *Phenomenology and the Cognitive Sciences*, 1–19.

Maiese, M. (2021). An enactivist reconceptualization of the medical model. *Philosophical Psychology, 34*, 1–27.

Nielsen, K. (2020). *What is mental disorder? Developing an embodied, embedded, and enactive psychopathology*. PhD thesis, Victoria University of Wellington. http://hdl.handle.net/10063/8957

Nielsen, K. (2021a). Comparing two enactive perspectives. *Philosophy, Psychiatry, & Psychology, 28*(3), 197–200.

Nielsen, K. (2021b). Comparing two enactive perspectives on mental disorder. *Philosophy, Psychiatry, & Psychology, 28*(3), 175–185.

Nielsen, K. (2022). Affordances and 3E psychopathology. In *Affordances in Everyday Life* (pp. 149–156). Springer.

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 8*(6), 800–822. https://doi.org/10.1177/0959354318808394

Nielsen, K., & Ward, T. (2020). Mental disorder as both natural and normative: Developing the normative dimension of the 3e conceptual framework for psychopathology. *Journal of Theoretical and Philosophical Psychology, 40*(2), 107–123. https://doi.org/10.1037/teo0000118

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–973.

Potochnik, A. (2010). Levels of explanation reconceived. *Philosophy of Science, 77*(1), 59–72.

Potochnik, A., & McGill, B. (2012). The limitations of hierarchical organization. *Philosophy of Science, 79*(1), 120–140.

Ramírez-Vizcaya, S., & Froese, T. (2019). The enactive approach to habits: New concepts for the cognitive science of bad habits and addiction. *Frontiers in Psychology, 10*, 301.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press. https://books.google.co.nz/books?id=OVGna4ZEpWwC

Thompson, E., & Cosmelli, D. (2011). Brain in a vat or body in a world?: Brainbound versus enactive views of experience. *Philosophical Topics, 39*(1), 163–180.

Thompson, E., & Stapleton, M. (2009). Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi, 28*(1), 23–30.

Varela, F. J. (1992). *Autopoiesis and a biology of intentionality*, pp. 4–14.

Varela, F. J., Thompson, E., & Rosch, E. (2017). *The embodied mind: Cognitive science and human experience*. MIT Press.

Wadhams, G. H., & Armitage, J. P. (2004). Making sense of it all: Bacterial chemotaxis. *Nature Reviews Molecular Cell Biology, 5*(12), 1024–1037.

Ward, D., Silverman, D., & Villalobos, M. (2017). Introduction: The varieties of enactivism. *Topoi, 36*(3), 365–375.

Zautra, N. (2015). Embodiment, interaction, and experience: Toward a comprehensive model in addiction science. *Philosophy of Science, 82*(5), 1023–1034.

4

The Bones of 3e Psychopathology

In this chapter I present the conceptual bones of my own framework: 3e Psychopathology. I begin with a brief overview of the conceptual tools inherent within the enactive worldview. This should offer further justification for the selection of 3e cognition as a guiding frame, and also offer a flavor of what 3e Psychopathology emphasizes relative to the frameworks reviewed in the last chapter. Following this I echo the format I used in Chap. 2, first unpacking my perspective on the *structural* nature of mental disorder, and then considering the *normative* basis for a label of disorder or dysfunction from an enactive perspective. In the following chapter I will then flesh these conceptual bones out through summarizing them together as a complete conceptual model. Throughout this chapter I will seek to highlight connections to the other enactive frameworks reviewed in the previous chapter and their various strengths and weaknesses.

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7_4

4.1 Conceptual Tools within the Enactive Worldview

In this section I will summarize some of the core conceptual tools that seem likely to make 3e cognition a particularly useful perspective from which to consider, study, explain, and treat mental disorder. The tools I will summarize in turn are: organizational causality and constitution, dual aspectivity, naturalized normativity, cultural embeddedness, thoroughgoing affectivity, a developmental perspective, and finally the fact that enactivism leaves room for—or perhaps even demands—a pluralistic approach.

Organizational Causality, Constitution, and Dual Aspectivity

One vital conceptual tool with the enactive worldview is that of *organizational causality*. Organizational or formal causality is simply the idea that the organization or form of an object can have causal effect on the world. This is strongly related to the concept of *emergence*, the idea that a whole/system may behave differently or gain qualities due to the organization and interaction of its parts, rather than being simply the sum of them. To be clear, such ideas are certainly not exclusive to enactive perspectives, indeed such formal causality was recognized by Aristotle. A simple example of organizational or formal causality would be water. The property of water being a liquid at room temperature is not held by a single H_2O molecule, rather it emerges from the interaction of multiple H_2O molecules repelling each other due to their dipole structure.

Applying such ideas of organizational causality to human functioning, neurotransmitters, cells, and organs, all clearly exist and play vital roles in shaping human behavior, but they are also part of the larger organized and dynamic whole and must be understood as such in order to grasp the nature of their influence. This is the complimentary and related idea of *constitution*. In explaining constitution I have previously given the example of building a tower of Lego (Nielsen & Ward, 2018). Both the *form* of the tower and the individual Lego blocks clearly exist, and an

understanding of both can be useful in an explanation of why the tower falls over or stays standing under different circumstances. During the time that the Lego blocks *constitute* the tower the blocks *are* the tower—yet the blocks still exist. We could similarly say that the blocks and the tower are ontologically 'fused' (de Haan, 2020b). Further, the form of the tower is not a separate 'product' of the blocks' organization, nor an epiphenomenal apparition. Similarly, organisms are made up of many parts (organs, cells, receptors etc.) and derive properties, such as mindedness, from the complex interactions between these parts in context. Both the parts and the wider organism are no less real because of the knowledge we gain about their parts and how they manage to dynamically constitute a minded creature.

Traditionally the dualistic divide between the mental and the physical has dominated our language and thereby our understanding of how 'mental' disorders can relate to physical processes. For example, in Chap. 2 we saw how certain conceptual positions understand mental disorders to be 'products' of biological/neurological abnormalities (e.g., biological essentialists), while others understand mental disorders to exist at a 'psychological level', somehow independent from biology (e.g., cognitive essentialists). Ideas of emergence, constitution, and organizational causality are inherent within the central 3e notion of embodiment—that our psychological functioning *is* the action of our biology in context over time. If our psychological functioning is embodied, then psychological dysfunction, whatever that may mean, seems unlikely to fit neatly into psychological or biological categories. This fits with the empirical evidence that 'mental disorders' are messy sorts of things, with causal factors 'dappled' over the traditional ontological domains of genetics, neurobiology, physiology, environment, as so on (Kendler, 2012). In a sense we might say that mental disorders are *ontologically disrespectful* in that they run amok over such categories. Notions such as organizational causality, dynamical constitution, and embodiment, allow us to make sense of such ontologically disrespectful phenomena because they allow us to conceive of constitutionally complex phenomena and process structures existing across multiple traditional scales of enquiry.

Further than this idea of constitution and a flexible relationship with traditional scales of enquiry, the 3e ontological perspective also

encourages genuine engagement with lived experience and related methodologies. While the subjective-objective divide begins to break down under enactivism in that subjective experience becomes 'real' as opposed to some controlled hallucination, there is a recognized division between the veridical world and the world as perceived from a concerned point of view (Thompson, 2007a). This is best described by Fuchs (2017) through his notion of *dual aspectivity* which I summarized in the previous chapter. As a brief reminder dual aspectivity is the idea the experienced/lived world is the same thing as the veridical world, but *as perceived* by the striving organism. Experience oriented methodologies such as phenomenology and hermeneutics, as well as the experiences of those who experience mental disorder, are often dismissed as 'subjective' and therefore less scientific. Under a 3e worldview however, notions such as dual aspectivity encourage such approaches to be seen as having a vital and unique role in attempting to capture and communicate lived experience. As well as containing the tools to consider constitutionally and causally complex structures and processes then, 3e cognition also contains tools to rectify such a view with lived experience in a plausible and naturalistic way, through the notion of perspective.

Naturalized Normativity

A further conceptual tool inherent to enactivism is that of a *naturalized normativity*. As discussed in the previous chapter, under enactivism life is defined as an autonomous or organizationally closed system that is organized in such a way as to maintain itself against the entopic flow of the surrounding universe. In order to maintain and adapt to the world we need to attain some things (food, shelter, warmth) and are best to avoid others (snakes, cliffs, and rotten food). It is thus through our precarious state and our concern with continuing to live that the world has imminent meaning for us. What this represents is a bridging of the normative gap mentioned in Chap. 2. To break this down a bit further, under enactivism the organism system is shaped by its ontogenetic and phylogenetic past to act in accordance with its needs, in accordance with the constraints of the environments that have shaped it. This allows an organism

to live in a valanced and meaningful world—to have purposes, goals, and even values inherent as tendencies within the organism system—without recourse to an error of teleology (Maiese, 2016; Thompson, 2007). As purposive creatures, many values and norms are inherent in our organization. Under such a worldview both the evaluativists' and the functionalists' views on mental disorder are in a sense correct because functionality is normative and deeply contextual. Later in this chapter I will aim to show how this provides new ground for considering the normative basis for disorder. In short it allows us to move beyond the evaluativist/objectivist divide—to see that mental disorders can be real things in the world *and* that diagnosing them is a deeply evaluative process.

Cultural Embeddedness

To be clear, the above discussion of natural normativity is not to reduce the meaning of human life to mere survival. Psychological functionality is not only normative and contextual in a personal sense, it is also cultural. As cognitively and culturally complex creatures, we have—over our evolutionary history, our cultural histories, and the spans of our individual life times—crafted ever more complex structures of meaning and value, literally building ourselves ways to survive and flourish within the complex social worlds that we continue to collectively construct. A key part of this view is the *constitutional view of culture* or CVC (Durt et al., 2017). The CVC is essentially an extension of the notion of embedment, drawing on the idea of constitution on an interpersonal scale, to present an understanding of culture as a constantly evolving and encompassing structure of shared meaning that groups of individuals dynamically constitute, and as individuals reside within. Under the CVC, groups of individuals, as well as their artifacts and institutions, are seen to constitute a *cultural ontology*—an ever-evolving shared sense of significance and meaning. Such a cultural ontology facilitates intra-group behavior and the transmission of tools, knowledge, and ways of knowing (Durt et al., 2017; Kirmayer & Ramstead, 2017). This 'shared world' is embodied within/enacted through the habits and practices of the group. Such cultural ontologies are then passed on to and continually tailored by younger

generations because they, at least for the most part, represent adaptive ways of understanding, managing, and altering the social and physical environment (Gallagher, 2017; Henrich, 2015; Heyes, 2018). Vitally, such cultural ontologies represent a major reshaping of the environment within which individuals develop, thereby molding the way that individuals perceive, make meaning of, and act in the world (Durt et al., 2017).

Ideas of cultural embeddedness then, are commonly endorsed by enactive authors and provide a way to understand how complex structures of meaning and value can develop in ways that are more distally related to, and can even seem to contradict, our inherent strive to survive (Di Paolo, 2005; Durt et al., 2017; Maiese, 2016). Over evolutionary and life-span time scales, behavioral/evaluative tendencies are selected for and developed, as they allow the organism to *flourish* in accordance with the constraints of the world around them—including the socio-cultural environment (Nielsen, 2021b; Nielsen & Ward, 2020). Under this view norms and values remain thoroughly embodied within the sense-making processes of individuals, the groups they constitute, as well as the artifacts, institutions, and policies they generate. Vitally, all of this is gradually built up from the enactive core of the individual's needful relationship with the environment, thus there remains a deep continuity between life and mind, between biology and normativity, as per the overarching spirit of enactivism[1].

Ideas of cultural embeddedness then, are an important part of the enactive worldview. Under such a view we can speak of complex layers of meaning built upon the root of survival and observe how this is achieved through the group-level enaction of cultural ontologies. This is a vital tool for the study of mental disorder for the simple reason that humans quite commonly do things that, at least at first glance, work against their deeply engrained purpose to survive and adapt. As examples, many of us regularly embalm ourselves with alcohol to the detriment of our health, and many of us risk our safety for political or spiritual causes. These

[1] De Haan (2020a, 2020b) develops a slightly different approach here, based on 'reflexivity'; our ability to see ourselves in the world and reflect on *how* we want to live. She argues that explanations of human values grounded in cultural embeddedness and functionality/flourishing such as I have endorsed here seek to 'explain away' something important about human life (de Haan, 2021). This underlies her commitment to an existential variant of enactivism.

things need to be understood within their cultural and individual context if their functionality for the individual is to be properly evaluated. Through notions of cultural embeddedness, enactivism contains the tools to facilitate such deep consideration of cultural context in a nuanced way.

Thoroughgoing Affectivity

Being able to understand our affective/emotional experience is of clear import for understanding mental disorders, and enactivism provides a useful framework for doing so. Inherent within enactivism is a thoroughgoing understanding of *affectivity*—our experience of emotion and concern in a broad sense. On the enactive view emotions are understood in a way that manages to hold to the constraints of a naturalistic worldview, while also according with our richly affective experience of the world. The 3e understanding of affectivity can be summed up in the following principles: 1) all meaning is *primordially affective*, 2) affectivity is embodied, 3) affectivity is relational and concerns meaning thereby having a functional root, 4) affectivity is embedded. I will now briefly break down these principles before offering a summary example of 'anger' that should implicitly demonstrate the utility of an enactive understanding of affectivity for the study of mental disorders.

Under enactivism emotions are not separate and distinct from cognition, nor something that is added on or somehow adjunct to normal functioning. Rather, affective concern is seen as our default state, and is part and parcel of being a living system. As discussed earlier, living systems enact a concerned point of view on the world. To be concerned is to be affected. In her book 'The Feeling Body', Colombetti (2014) explores this deep connection between affectivity and meaning within the enactive worldview. She presents the idea that life is *primordially affective.* On such a view what we label as emotions (e.g., sadness, anxiety, anger, etc.) are temporary fluctuations of intensity, not against a back drop of neutral rationality, but in the meaningful experience of the organism as it stands in relation to the world (Colombetti, 2014). Affectivity therefore is intrinsic to sense-making, and our emotions are simply socio-cultural labels given to recurring and similar-enough patterns of affectivity. As

with all enaction of meaning under the 3e perspective, emotions are embodied. They are not only subjective experiences but are simultaneously a pattern of activity within our brain and body. This can be understood through the notion of dual aspectivity explained earlier. Consider again from the previous chapter, driving on the motorway when a police car's lights flash behind you and the ensuing flurry of activity across your nervous system and body.

On this view emotions are about meaning. As such they typically arise when an object or situation has strong meaning for the individual. This meaning, as all meaning under the enactive view, is based on the kind of creature we are, the needs we have, and how we have learnt to make sense of and adapt to the world in the past. In this sense they are also deeply functional. Emotions are dynamic modes-of-being that encourage ways of engaging with the world that have in some sense worked for us before (in either our evolutionary or life-span past). Metaphorically then, emotions are echoes of our past experience, both across our life-span and evolutionary development, encouraging ways of making sense of and responding to the world because they have been reinforced, culturally scaffolded, or naturally selected for. Emotions thereby, are not deviations from 'rationality', but rather make us more flexible and adaptive—they are akin to putting on different shoes to help us navigate different terrains (if these shoes spontaneously materialized from our feet). This of course is not to ignore that: 1) we often enact emotions that do not actually serve us well in our current contexts, 2) emotions as physiological processes can be evoked or suppressed by biological or medical means, and 3) that we can seemingly get stuck in patterns of unhelpful emotionality as exampled by mental disorders such as depression and generalized anxiety.

Finally, it is important to consider that under a 3e perspective the relationship between our emotions and the state of the world is a two-way street. Our affective state is dynamic and situational, and thereby needs to be understood through the lens of our embedment. Given their relational and meaningful nature, our emotions are often facilitated by the environment and what it affords us. We in turn often shape the environment in ways that scaffold particular affective states (Krueger & Colombetti, 2018). We craft awe inspiring cathedrals, comforting homes,

neutral/professional offices, and imposing courtrooms. Entire domains of cultural practice (e.g., art, music, film), and industries such as marketing, are built upon the ability of human crafted stimuli to affect us. This is all to highlight that emotions shape how we understand and affect the world, that the world affects how we feel, and that we shape the world to shape how we feel. Here I am primarily drawing on the ideas of Krueger and Colombetti (2018). These authors discuss in detail this two-way relationship between affect and the world from an enactive view, and apply their ideas to mental disorders such as depression and schizophrenia. While I do not have space to unpack this paper here, their analyses are a great example of how the thoroughgoing understanding of affectivity within the enactive perspective can facilitate novel and useful hypotheses regarding the structure of mental disorders.

As a summary example of emotion/affectivity under the enactive worldview, let us consider anger. If I feel angry, this emotion is not simply a state of my brain, but a dynamic collection of processes cascading throughout my brain and body, many of which are automatic or self-propelling. This might include a release of cortisol, a raising of blood pressure, a narrowing of my attention on the object of my anger, a dulling of experiences of pain, and so on. This cascade can be understood as having developed across my evolutionary and life-span development and encourages me to aggressively assert my needs even at the cost of others around me. It thereby represents a temporary biasing of my meaning making, reorganizing the affordances salient to me, encouraging me to be fast and decisive, and to take actions that appeared helpful in similar contexts-of-meaning in my evolutionary or life-span past, such as making ruthless social decisions, yelling, or punching. Across my socio-culturally immersed development I may (or may not) have learnt to recognize the experience of this embodied cascade as 'anger'. Through my experience I have likely also developed second-order layers of meaning about my experience of 'anger' based on whether it has worked for me in the past and how it is understood within my local cultural ontology. For example, I might have variously learnt 'it is not acceptable for me to feel anger', 'anger makes me feel powerful and helps me get what I want', or 'anger is ok, but I am still responsible for my actions when I am angry'. Anger and other emotions then, are not separate to cognition, rather affectivity is

inseparable from how we make sense of the world. Emotions are not deviations from some idealized rationality but are intrinsic to how we experience sense-making. We thereby cannot 'control' or 'manage' anger itself—some things are going to make us angry because of the meaning they hold for us! We can however learn to observe our anger and how we make sense of it, learning to choose how we respond and act when we are angry. Consider briefly in a therapy context how such a perspective allows a compassionate and collaborative perspective on oft problematized emotions such as anger, while still holding a client responsible for their actions. Consider too, from a research/explanatory perspective, how such a framework for understanding our emotional lives provides an appropriately naturalistic frame while integrating developmental, experiential, evolutionary, physiological, and neuroscientific perspectives.

A Developmental Perspective

Another conceptual strength for the consideration of mental disorder found within 3e cognition is the emphasis given to the historicity of the organism and its behavior. From discussion in the previous sections on cultural embeddedness and affectivity, it should be apparent that enactivism encourages consideration of factors across multiple timescales in the explanation of human behavior. The sense-making and behavior of a human organism in any given moment is shaped by our deep evolutionary past, our social-cultural histories, our development and learning across the lifespan, and our current life situations. All of this is funneled through the embodied sense-making of the person in context. As a simple example that captures only some of these influences, consider how holding a snake will mean very different things and evoke different responses for someone who lives in the venomous wonderland of Australia, to someone who has spent years as a snake handler, to someone whose culture or religion reveres snakes in some way. Further, as discussed in the section on constitution, to speak of embodiment is to acknowledge that this historical influence is playing out in the dynamic organization of the organism, through our physiology and behavior, and with this comes recognition of the explanatory power of smaller-faster scales of investigation such as the neurological and chemical.

With recognition of the explanatory relevance of all of these various timescales comes a challenge of practicality. That is, how should we seek to integrate such complexity into our explanations and/or evaluations of human behavior while keeping them simple enough to be useful? One idea that can potentially serve to boil down the developmental complexity somewhat is that of a *mode of functioning* (Nielsen, 2020b; Nielsen & Ward, 2020). In brief a 'mode of functioning' is a model of how a person, in consideration of their body, culture, and development, has learnt to navigate the world and meet their needs—how their sense-making has developed to work for them. This includes how they have learnt to understand themselves, their own emotions, other people and the world around them, and how they have learnt to manage perturbations to their environment or ways of living. I shall return to discuss the relationship between this concept and both the structure and normative nature of mental disorder throughout this chapter. I shall also return to this concept in the later chapter on explanation where I will argue that it has particular utility in the crafting of idiographic explanations of psychological difficulties.

Demand for Pluralism

A final strength of 3e cognition that I wish to emphasize here is that it demands a "…polyperspectival approach" (Fuchs, 2017, p. 276). 3e cognition is best understood as an ontological framework or philosophy of mind—it is a worldview, not a theory. As such it should be seen as open to a range of methods of enquiry and ways of explaining (Donovan & Murphy, 2020). Human behavior is clearly a complex thing, inevitably influenced and shaped by a multitude of factors. Discussion in the previous subsections should highlight that 3e cognition recognizes this and provides a wider frame in which to develop theories of how these factors are integrated within the enaction of human behavior or dysfunction. In short, it seems possible that a 3e framework for considering psychopathology may provide an appropriate ontological framework within which to situate conceptual, classificatory, explanatory, and methodological pluralisms. Further, doing so may provide a common ontology and thereby

a helpful degree of constraint at a conceptual level as to increase commensurability across different approaches.

To break this claim down further, consider the complexity that 3e thinking encourages us to recognize. 3e cognition recognizes the existence and relevance of multiple scales of enquiry in the study of human functioning. These scales extend both spatially and temporally and are understood as different perspectival aspects of the same dynamically constituted and situated organism as well as its development across time. Considering such complexity, it seems doubtful that everything that is useful to know about a particular pattern of disordered behavior (e.g., anorexia) will ever be captured in a single integrated theory. Indeed, under such a complex worldview it seems likely that there will be multiple justifiable ways to think about, classify, study, explain, and treat patterns of psychological and behavioral difficulty. This is reasonably congruent with various pluralist perspectives on psychopathology and related areas, and within philosophy of science more broadly (e.g., Brigandt, 2013; Chang, 2017; Hawkins-Elder & Ward, 2021; Jerotic & Aftab, 2021; Kendler, 2012; Markon, 2013; Veit, 2020; T. Ward & Clack, 2019; Wegerhoff, 2022; Wegerhoff et al., 2020, 2022). As an aside here, it is important to not confuse ideas of pluralism with the claim that 'anything goes'. Some explanations, methods, or ways of classifying are always going to turn out to be more useful for their relative purposes than others, and this is perfectly congruent with the 3e worldview. While 3e thinking pushes us to consider potential causal factors across the brain-body-environment system, it does not require us to see biological factors, social factors, developmental factors, and so on as equipotent in every case. Rather, the relative importance of different factors will depend on what is actually going on, what it is we want to do with our explanations, and what we have the ability to control/influence. But I digress.

To return to the point, through recognizing the factorial complexity inherent in human behavior and its apparent dysfunction, 3e thinking demands a pluralistic approach. On my view this is a clear strength given my wider commitments to pluralism highlighted in Chap. 1. This demand for pluralism extends to all relevant tasks in psychopathology—i.e., conceptualization, classification, explanation, and treatment—as well as the methods used along the way. However, as I have attempted to

outline in the previous chapter and the preceding subsections of the current chapter, 3e thinking clearly makes certain ontological and epistemological commitments. As such, committing to an enactive worldview provides a degree of constraint/guidance in the completion of these tasks. In effect it is the nature of this constraint/guidance that I will seek to explore over the rest of this book. While potentially aspirational, it seems possible that such a degree of constraint may allow a patchwork of different approaches and explanations focused at different scales of enquiry to sit alongside each other in a more commensurable way, thus allowing for a marginally more integrated understanding of these things we call mental disorders.

4.2 Enactivism and the Structure of Disordered Behavior

The previous section has highlighted some useful conceptual tools inherent to the 3e perspective, and in doing so has essentially offered a summary of how human functioning is understood under enactivism. In this section I aim to present my own 3e Psychopathology view regarding how best to consider the ontic structure of mental disorder. To reiterate the distinction made in Chap. 2, by 'structural' I mean that this section is focused on the question of 'what are mental disorders in terms of their causal structures' and not the question of 'why something should or should not count as disorder'. Near the end of this section, I will briefly compare and contrast with the positions reviewed in the previous chapter.

Under 3e Psychopathology mental disorders are most simply understood to be *recurrent patterns in sense-making*. To step back from the enactive language, what I mean by this is that, structurally speaking, mental disorders are simply recurrent patterns and tendencies in thinking, feeling, and behavior—i.e., in how the organism makes sense of the world and engages with it. Importantly, as with all sense-making or behavior, these patterns are *causally influenced and constituted by a host of factors littered across the brain-body-environment system*—to the point that these patterns may often be considered emergent. We are bodily organisms

richly embedded in a physical and social world saturated with meaning. The body, the physical and socio-cultural environment, as well as considerations of evolution, development, and the meaning we enact in the world moment to moment, are all relevant for understanding both why behavior is performed, and why it takes the form that it does. For example, when I get up from my writing, notice that I am hungry, and reach for an apple on the bench nearby, this brief pattern of activity is, to say the least, complicated. It seemingly has to do with mechanical and chemical receptors in my gut, the transmission of relevant signals to my brain, other meal-time associated signals from across the system, the shift in my attention both neurologically and physiologically away from my work, the historically informed embodied-meaning of these signals within my body as 'hunger', the historically informed embodied-meaning I make of the pattern of light hitting my retina as "apple/food/yum", and the dynamic unfolding of relevant neuro-physiological sequences as my body moves through space and grasps for the apple. To be clear, this simple example is not intended to be scientifically accurate, but rather to highlight that all acts of sense-making, big or small, functional or dysfunctional, are complex processes that unfold over time in a bodily involving, situated, and historically informed manner. We have no reason to think that mental disorders are any different in this regard.

One way that mental disorders are different however, is that they are *dysfunctional*. Unpacking exactly what I mean by 'dysfunctional' is the topic of the next section, but for now it is enough to say that mental disorders often cause distress and get in the way of important functions in day-to-day life. Despite this, these repetitive patterns of sense-making continue to unfold within a person's attempts to adapt to the world. In other words, mental disorders continue to work against the functioning of the individual despite the striving efforts of those that are suffering them to survive and flourish. To borrow from de Haan (2020b), mental disorders are in this sense *inflexible*. As patterns of thinking, feeling, and behaving, they are repeatedly enacted despite their negative consequences. On the basis of this observation, and in the absence of evidence for dominant sub-personal causal factors such as predicted by the biological essentialists, we can make a further inference. Namely that these patterns and the wider causal structures supporting them are *locked-in* and thereby

circular in some way; i.e., that they are *reafferent stable dynamic patterns of causal relations within the brain-body-environment system* (Nielsen & Ward, 2018).

For example, consider the differences between Parkinson's disease and Major Depressive Disorder (MDD). In Parkinson's we can observe certain phenomena within functioning—shaking and loss of motor control, among others—but these are tied to a relatively homogenous set of known causal factors within motor areas of the brain, such as atrophy of dopaminergic neurons in the substantia nigra. The reason for the maintenance of these phenomena then is reasonably 'in the brain'[2]. Compare this to depression, where we are not currently aware of any dominant causal factors within the brain. Assuming that we continue to 'fail' in this search for the 'essence' of depression, we do not need to relegate depression (nor other mental disorders) as merely problems in living as per Szaszian approaches. Instead, 3e cognition, with its view of the entire brain-body-environment as a dynamic system, highlights the possibility that the maintenance of dysfunctional sense-making may be emergent from a network of factors and feedback-loops across the system. Rather than representing an undiscovered disease process in the brain, mental disorder syndromes may represent circular multi-scale networks of causal relations. In short, the causal structure supporting repetitive patterns in sense-making starts to look a lot like the fuzzy MPC kinds we explored in Chap. 2 (Kendler et al., 2011).

One key distinction to make between the typical MPC view and the view expressed here however, is that under 3e Psychopathology this MPC causal structure is seen as existing within the adaptive processes of an agent-in-relation-to-the-world. In effect, such MPC structures are understood as embedded within the adaptive striving of individuals as they navigate the world around them. In concordance with the principles of

[2] This is not to say that the wider pattern of difficulties that people with organic diseases experience cannot be fruitfully analyzed though a system-wide lens. Such an analysis would also likely highlight a complex network of causal relations impinging on a sufferer's wellbeing, but we can visualize the network in this instance as being much more centralized around the core pathogenic process in the brain—as being 'denser in the middle'. The claim I am making is that the network supporting depression and other mental illnesses is likely more diffuse (although there may well be hubs of causal connections, within the brain or elsewhere). The structures of mental disorder and physical disorder seem continuous in this way.

enactivism, this highlights the role of meaning and the historicity of our functioning. As I hope to show in the later chapter on explanation, when we look across groups of people with similar problems, we may be able to abstract common phenomena that play a role in maintaining distress/dysfunction (e.g., maintaining mechanisms), and we may even be able to model how these mechanisms typically play out, thus getting a sense of this wider nomothetic MPC structure for a particular pattern of disorder. At an individual level however, the way that these 'mechanisms' unfold is always going to be a little bit different, a little bit messy, and deeply related to how a person has learnt to function.

What I am referring to here can be understood under 3e Psychopathology as a *mode of functioning*, which I briefly referred to in the previous section. Someone's mode of functioning most simply refers to how they have evolved and learnt to live their life. Breaking this down further it refers to the meaning that they give important features of the world, e.g., what meaning does a person give to their own body or sensations therein, to significant others, to strangers, to their own emotions (generally or specifically), to others' emotions (generally or specifically), to their own thoughts (generally or specifically), and to potential threats and challenges? How have they learnt to value or disvalue themselves, to value and disvalue others, and what does it mean to them to live a good life? Within 3e Psychopathology a 'mode of functioning' then, is an intentionally underspecified and flexible concept which serves the vital function of summarizing a complex history of evolution, cultural adaption, and learning, into a rough model of how an individual strives to survive and flourish. Under 3e Psychopathology, mental disorders can be understood as MPC type process structures, but these structures do not play out against a neutral background. Rather, they always exist in the context of the individual's wider mode of functioning. In effect, an individual's mode of functioning can be understood as an ecosystem and a disorder as a particular pattern of dysfunction that dynamically unfolds within it.

In this section I have presented my understanding of the structure of mental disorder when viewed through an embodied, embedded, and enactive lens. Building on previous works I have suggested that mental disorders are *recurrent patterns in sense-making supported by reafferent stable dynamic patterns of causal relations within the brain-body-environment*

system (Nielsen, 2020b; Nielsen & Ward, 2018). Similarly to Maiese (2021) then, 3e Psychopathology understands disorders as patterns in sense-making—as causally and constitutionally complex process structures. As discussed in the previous chapter, this appears to offer advantage in terms of scientific utility relative to de Haan (2020b) and her notion of 'biases in sense making' because, through making claims about the dynamic constitutional structure of disorders themselves, such a view encourages exploration of how mental disorders work rather than simply what they are. In opposition to Maiese (2021), 3e Psychopathology does not endorse notions of seeing patterns of mental disorder (or other patterns within functioning) as different 'selves' or 'identities', no matter the degree of autonomy that a particular pattern may hold. Finally, in this section I have continued to introduce the important notion of a 'mode of functioning' and have stressed that, at an individual level, patterns of disorder are always deeply embedded in how a person has learnt to engage with and make meaning in the world.

4.3 Enactivism and the Normative Basis of Disorder

> "Health, from this perspective, is very different from a statistical species-specific correlation of normality, and there are consequently many ways of being healthy" (Di Paolo, 2010, p. 51)

In this section I return to the question foreshadowed by Thornton (2000) and explored at the end of Chap. 2: Can assuming a richer and non-reductionistic worldview such as 3e cognition allow us to see beyond the evaluativist/objectivist dichotomy and understand the role of values/normativity within the concept of mental disorder in naturalistic terms? In other words, what I am exploring here is whether enactivism affords us a way to see values and normativity as a natural part of the world and thereby collapse the normative gap. I will be re-presenting arguments that it can perform such a function originally presented in Nielsen and Ward (2020), and summarizing the normative basis for claims of disorder under 3e Psychopathology. Throughout this section I am primarily

concerned with what is normatively required for something to be considered a mental disorder under 3e Psychopathology. In line with this purpose I largely set aside structural and epistemological issues.

To refresh on ideas from Chap. 2, objectivists are those that want to see mental disorder as something naturally defined in the world, and typically believe that in order to do so we need to expunge human values from the concept. The functionalists are the dominant position within the objectivist camp. Evaluativists meanwhile, hold that it is impossible to expunge values from the concept of mental disorder and that to do so would miss something vital about what it means to be human. Evaluativism comes in a variety of forms, and in Chap. 2 I presented a loose taxonomy of these. Weak-evaluativism holds that values clearly play a role, but that the relevant values are assumed to be universal and therefore not particularly contentious. Weak-evaluativism does not prescribe the inclusion of socio-culturally specific values in consideration of what counts as disorder. Strong-evaluativism meanwhile holds that socio-culturally specific values should play a role in defining disorder. This makes the disorder label highly contextual, hard to define, and open to charges of unsustainable relativism. Finally, anti-psychiatric evaluativism holds that mental disorder concepts are laden with socio-cultural values in a way that unacceptably constrains individuals and we should give up on the whole enterprise.

As foreshadowed in Chap. 2, I aim to show that there is meaningful middle ground between the weakly and strongly evaluativist positions. This will be a position whereby social and cultural norms sometimes play a role in demarcating disorder, so long as they typically serve the functioning of the individual diagnosed. More so, I aim to show that such a middle ground can be justified under principles of naturalism—thereby collapsing the normative gap between the objectivists and the evaluativists. In short I aim to show that mental disorders can be understood as both natural *and* normative, and that a path to doing so is inherent within the natural normativity found within enactivism (Nielsen & Ward, 2020).

Under 3e cognition, human functioning is normative and value inclusive. As we saw when discussing the DCT in Chap. 3 and naturalized normativity earlier in this chapter, under enactivism meaning emerges through the needful relationship between self-maintaining adaptive

organisms and their environments. Normativity is thereby a natural part of the world, and the meaning we enact is deeply tied to our functioning. As organisms adapt to fill a range of evolutionary and behavioral niches they make more and more specific demarcations in the world, and the meaning they enact thereby becomes more and more complex (Okrent, 2017). When discussing the CVC we also saw how the emergence of more and more complex structures of meaning is further propelled by the relationship between an individual and their wider cultural group (Durt et al., 2017). On this view, 'values' are culturally specific labels for very real evaluative tendencies within our modes of functioning, developed across evolution, culture, and life-span development (Nielsen, 2021a, 2021b). What we can see then, is a thread of continuity from the complex structures of value and meaning we experience as socio-culturally embedded humans, all the way back to a basic organismic strive to survive. This is the true depth of the DCT (Thompson, 2007). While enactivists primarily utilize ideas such as these in an effort to understand the emergence of *meaning* and *experience* for living beings, such ideas also clearly represent an understanding of how *normativity* can emerge in a world of facts (Hume, 1978/1739). Indeed, in many ways 'meaning' is basically the subjective experience of normativity—normativity *for me* as I perceive it. Insofar as an organism *should* act to maintain its own life and flexible adaption to its various environments, there are states, actions, and processes that the organism *should* be in or perform[3]. Vitally, these states, actions, and processes change in accordance with the current needs of the organism and the constraints of the environment.

Based on this deep continuity, and drawing on the work of Okrent (2017), and Christensen (2012), I have previously proposed that enactivism affords a kind of *systems functionalism* for considering the demarcation of disorder in the domain of behavior and sense-making (Nielsen, 2020b; Nielsen & Ward, 2020). Under this view living systems are understood to seek to maintain themselves in a 'non-equilibrium steady state'—i.e., an *organized* state against the backdrop of an entropic universe. For an organism to act against the norms of its own self-maintenance, or to be stuck in a pattern of behavior that works against or disrupts its

[3] De Haan (2020a) uses the term 'relational realities' to refer to this dynamic.

wider mode of functioning represents the system pushing away from its steady-state and towards entropy. The organism is in such a case acting against its very purpose—to strive to survive and adapt—and such actions can thereby be legitimately conceptualized as *pathological*. As a simple example, consider a model bacterium that, for some unknown reason, begins motivating away from sugar or toward toxins. This amounts to the organism making sense of and acting in the world in a way that works against its own survival and adaption. Under 3e Psychopathology this can be understood as analogous to mental disorder (Nielsen, 2020b, 2021b; Nielsen & Ward, 2020).

A key question at this point is, how should we apply this systems functionalism to humans? Human functioning seems, to say the least, somewhat more complex than this simplified model bacterium. One way to operationalize this, previously presented in Nielsen and Ward (2020), is through the concept of 'functional norms'. Most simply, *functional norms* are states of the world necessary for an organism's continued functioning and adaption. This includes states, process, and actions, performed by an organism (or existing within the constitution thereof), so long as the maintenance, existence, or their being within a particular range, supports the organism's self-maintenance and adaption. Examples of such functional norms are reasonably easily identified at the scale of physiology—e.g., blood pressure, the availability of water, heart rate, the respiration process, environmental temperature, caloric intake, the continued attachment of limbs, etc. Part of the reason for these physiological examples being easy to identify is that these norms are roughly similar for everyone, and when such norms are broken the result is often clear illness or death. However, it is important to note that even with these physiological functional norms we can see that they are *graded*, *interrelated,* and *contextual* (Di Paolo, 2005). For example there is a band of temperature within which humans can survive, not a single point, and as the temperature rises or cools away from some adaptive mean, we don't suddenly expire but rather find it harder and harder to survive—i.e., functionality is *graded*. Further, the temperature we can handles depends on other factors such as humidity, wind chill, our knowledge of how to dress appropriately to the climate, and how much water is available. In other words, functional norms are *interrelated* and *contextual*. The important thing is

that all of our functional norms balance out in a way that supports continued self-maintenance and adaption.

Compared to physiological functional norms, potential functional norms of behavior—or perhaps, *functional norms of sense-making*—are much harder to specify. One reason for this is that there is a legitimate diversity of different ways of being when we speak of psychology and behavior. Cultural diversity round the world is a clear example of this—there is simply no one right way to make sense of and act in the world. It appears therefore, that psychological functioning is more deeply contextual that physiological functioning—what works for one person in one context may not work for another. Even behaviors that one would at first assume to be clearly 'pathological' such as repeatedly jumping off cliff faces can be understood as contributing to a functional way of being in the world if, for example, someone was part of a base-jumping community. Similarly, consider how eating a certain amount of food might be perfectly healthy for one person who does a lot of physical activity while being unhealthy for someone who does not. To reiterate then, functional norms are *interrelated* and *contextual*. The upshot of this is that in evaluating the functionality of a particular pattern observed in someone's sense-making, we must consider the various contexts they navigate, as well as the wider *mode of functioning* within which the pattern operates.

Just like physiological functional norms, functional norms of sense-making are *graded*. Behaviors can be *better* or *worse*, without necessarily directly impacting ones self-maintenance (Di Paolo, 2005). Parallel to the temperature example above, I can eat too much chocolate or not get enough sleep and it does not kill me, although it may affect my ability to adapt to a change in context or negatively affect my ability to keep living as my body ages. In other words such factors may negatively affect my ability to *flexibly adapt*. An important point here however is that being sub-optimally adaptive is normal. Eating too much chocolate or not getting enough sleep is not necessarily pathological, rather this is a matter of degree (and context, as highlighted in the previous paragraph). Similarly, I can be somewhat more (or less) anxious than others, even to the point of it being *less* adaptive in my current context, without it being clearly pathological or maladaptive. If however a repeated pattern of heightened anxiety can be seen to negatively impact my wider web of functional

norms to an atypical degree, then my mode of functioning in the world is disrupted and we can begin to speak of pathology or dysfunction. For example, if this anxiety is impacting my ability to eat, to sleep, to interact with others and maintain relationships, to perform an important role I have in my community, or to perform any other tasks that are an important part of my evolutionarily, culturally, and developmentally informed mode of functioning, then my enaction of this anxious pattern of sense-making is clearly working against me. In a sense, I am working against myself through my inflexible enaction of this pattern of sense-making.

To summarize, functional norms are those norms that support the organisms continued self-maintenance and adaption, and by extension, their ability to fare-well in their communities (Di Paolo, 2005; Maiese, 2016). As per the CVC what exactly it means to fare-well and adapt for any individual will subtly change as a function of said individual's tendencies and development, and will co-vary with the culture in which they learned to function. Hence the need to consider the individuals own mode of functioning. From this position, *mental disorder is a recurring pattern in sense-making that runs counter to the individual's functional norms to a significant or atypical degree, disrupting their wider mode of functioning in the world*. The long and the short of this analysis then, is that the normative basis of mental disorder should be based around the question 'is this a problem for this individual within their context and in consideration of their wider mode of functioning?' This is importantly different from physiological medicine where answers to this question seem more easily inferred from statistical norms or empirical data.

I have previously attempted to formalize this systems functionalism into a set of definitions (see: Nielsen, 2020b; Nielsen & Ward, 2020). These works also consider in further detail: the distinction and interrelation between functional norms as I have described them and norms of typicality, evolution, morality, and law; the complex process of distinguishing an individual's functional norms from wider socio-cultural norms; the challenges of altruism and suicide for this perspective; and instances where seemingly 'disordered' behaviors can also be doing functional work (e.g., non-suicidal self-injury serving a function of emotional regulation). In the interest of succinctness I have not delved into such

issues here. However, more fundamentally, what I have been trying to communicate in presenting these ideas is that the process of diagnosing whether disorder/dysfunction is present or not, is, and should be considered, an *evaluation*. It is an evaluation that should *not* be performed in reference to socio-cultural standards, be they implicit cultural values, the dictates of evolutionary psychology, or the 'functional impairments' of a diagnostic manual. Rather this evaluation should be person-centered—performed on the basis of the individual patients own context and mode of functioning in the world. The core question being, is the observed pattern in sense-making working for this person, or is it working against their ability to fair-well and adapt in a significant way?

In this section I have presented my understanding of the normative nature of mental disorder when viewed through an embodied, embedded, and enactive lens. This view shares a certain congruity with Maiese's (2021) notion of maladaption in that mental disorder is defined by a disruption to the adaptive fit between a person's sense-making and their environment. However, I have specified this notion in regards to functional norms and an incompatibility between the observed pattern of sense-making and the individual's wider mode of functioning, while Maiese specifies this notion through reference to "…the destabilization of particular regional identities or the self as a whole" (2021, p. 974). Parallel to my earlier criticism of Maiese's framework, this reference to 'identities' and 'self' seems unnecessary and potentially begets confusion. It therefore seems advantageous that 3e Psychopathology does not reference such concepts. As I have previously argued in Nielsen (Nielsen, 2021a, 2021b) there also appears to be an approximate compatibility between the normative construal presented here and de Haan's four Roschian characteristics of disordered sense-making (de Haan, 2020b). The normative view presented here attempts to provide not only a description of what disordered sense-making looks like, *a la* de Haan (2020b), but rather presents a *justification* for the label of pathology through the concepts of natural normativity, the CVC, modes of functioning, and systems functionalism. As such, while de Haan's characteristics are surely more practical and easier to use, the current view presents a more centralized idea of what it means for sense-making to be

disordered and this appears to offer greater *conceptual* guidance. Hopefully this guidance will be apparent as this book continues, but as an example for now, consider de Haan's characteristics in light of the view presented here. As a reminder, her four characteristics of disordered sense-making were to do with: 1) Appropriateness or typicality-in-context, 2) Inflexibility of sense-making, 3) Inflexibility of stance-taking, and 4) Suffering. Viewing these through the lens of 3e Psychopathology, such characteristics are no longer arbitrary. Rather, suffering is a common outcome of dysfunctional sense-making and inflexibility captures an important feature or dynamic. As such characteristics 2-4 seem reasonably central and justifiable features of dysfunctional sense-making that may support the presence of pathology under the 3e Psychopathology framework. The appropriateness characteristic however appears as less conceptually central or justified. Certainly some behaviors seen in mental disorder can often be 'abnormal', however under the 3e Psychopathology framework presented here this is not an important part of the conceptual machinery.

4.4 Summary

This chapter has reviewed some of the core conceptual tools available within 3e cognition that seem particularly relevant to the study of mental disorder, before presenting the structural and normative understandings of mental disorder under 3e Psychopathology. Some of the important tools covered included organizational causality, thorough going affectivity, a demand for pluralistic approaches, and the concept of an evolutionarily-culturally-developmentally informed mode of functioning. From a structural perspective 3e Psychopathology understands mental disorders as dysfunctional patterns in sense-making that people have a tendency to get stuck in. These patterns are understood to be constitutionally complex, multi-scale, and frequently reafferent causal structures, existing within the brain-body-environment system. From a normative perspective, mental disorders are patterns of sense-making that work significantly against the person enacting them. In other words they are patterns that

get in the way of important functions or otherwise disrupt a person's mode of functioning in the world. In the following chapter I will stich these conceptual bones together and flesh out the 3e conceptual model through comparison other perspectives. Questions of classification will also briefly be considered.

References

Brigandt, I. (2013). Explanation in biology: Reduction, pluralism, and explanatory aims. *Science & Education, 22*(1), 69–91.

Chang, H. (2017). Is pluralism compatible with scientific realism? In *The Routledge handbook of scientific realism* (pp. 176–186). Routledge.

Christensen, W. D. (2012). Natural sources of normativity. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43*(1), 104–112.

Colombetti, G. (2014). *The feeling body*. The MIT Press.

de Haan, S. (2020a). An enactive approach to psychiatry. *Philosophy, Psychiatry and Psychology, 27*(1), 3–25.

de Haan, S. (2020b). *Enactive psychiatry*. Cambridge University Press.

de Haan, S. (2021). Two enactive approaches to psychiatry: Two contrasting views on what it means to be human. *Philosophy, Psychiatry, & Psychology, 28*(3), 191–196.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences, 4*(4), 429–452. https://doi.org/10.1007/s11097-005-9002-y

Di Paolo, E. (2010). Overcoming autopoiesis: An enactive detour on the way from life to society. In *Advanced series in management*. Emerald Group Publishing Limited.

Donovan, C., & Murphy, D. (2020). De Haan on sense-making and psychopathology. *Philosophy, Psychiatry, & Psychology, 27*(1), 29–30.

Durt, C., Fuchs, T., & Tewes, C. (2017). *Embodiment, enaction, and culture: Investigating the constitution of the shared world*. MIT Press. https://books.google.co.nz/books?id=OJakDgAAQBAJ

Fuchs, T. (2017). *Ecology of the Brain: The phenomenology and biology of the embodied mind*. Oxford University Press.

Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford University Press. https://books.google.co.nz/books?id=Z28sDwAAQBAJ

Hawkins-Elder, H., & Ward, T. (2021). From competition to co-operation: Shifting the "one best model" perspective. *Theory & Psychology, 31*, 821–841. https://doi.org/10.1177/0959354321995900

Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.

Hume, D. (1978). A treatise of human nature [1739]. *British Moralists*, 1650–1800.

Jerotic, S., & Aftab, A. (2021). Scientific pluralism is the only way forward for psychiatry. *Acta Psychiatrica Scandinavica, 143*(6), 537–538.

Kendler, K. (2012). The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. *Molecular Psychiatry, 17*(4), 377.

Kendler, K., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*(6), 1143–1150.

Kirmayer, L. J., & Ramstead, M. J. (2017). Embodiment and enactment in cultural psychiatry. In C. Durt, T. Fuchs, & C. Tewes (Eds.), *Embodiment, enaction, and culture: Investigating the constitution of the shared world* (p. 397). MIT Press.

Krueger, J., & Colombetti, G. (2018). Affective affordances and psychopathology. In *Philosophical perspectives on affective experience and psychopathology: Vol. XXVIII–2* (pp. 221–247). Quodlibet.

Maiese, M. (2016). *Embodied selves and divided minds*. Oxford University Press. https://books.google.co.nz/books?id=w_quCgAAQBAJ

Maiese, M. (2021). An enactivist reconceptualization of the medical model. *Philosophical Psychology, 34*, 1–27.

Markon, K. E. (2013). Epistemological pluralism and scientific development: An argument against authoritative nosologies. *Journal of Personality Disorders, 27*(5), 554–579.

Nielsen, K. (2020a). Think of mental disorders as the mind's 'sticky tendencies.' *Aeon*. https://aeon.co/ideas/think-of-mental-disorders-as-the-minds-sticky-tendencies

Nielsen, K. (2020b). *What is mental disorder? Developing an embodied, embedded, and enactive psychopathology*. PhD thesis, Victoria University of Wellington. http://hdl.handle.net/10063/8957

Nielsen, K. (2021a). Comparing two enactive perspectives. *Philosophy, Psychiatry, & Psychology, 28*(3), 197–200.

Nielsen, K. (2021b). Comparing two enactive perspectives on mental disorder. *Philosophy, Psychiatry, & Psychology, 28*(3), 175–185.

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 8*(6), 800–822. https://doi.org/10.1177/0959354318808394

Nielsen, K., & Ward, T. (2020). Mental disorder as both natural and normative: Developing the normative dimension of the 3e conceptual framework for psychopathology. *Journal of Theoretical and Philosophical Psychology, 40*(2), 107–123. https://doi.org/10.1037/teo0000118

Okrent, M. (2017). *Nature and normativity: Biology, teleology, and meaning*. Routledge.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press. https://books.google.co.nz/books?id=OVGna4ZEpWwC

Thornton, T. (2000). Mental illness and reductionism: Can functions be naturalized? *Philosophy, Psychiatry, & Psychology, 7*(1), 67–76.

Veit, W. (2020). Model pluralism. *Philosophy of the Social Sciences, 50*(2), 91–114.

Ward, T., & Clack, S. (2019). From symptoms of psychopathology to the explanation of clinical phenomena. *New Ideas in Psychology, 54*, 40–49. https://doi.org/10.1016/j.newideapsych.2019.01.004

Wegerhoff, D. (2022). *Understanding gangs: Developing an epistemically pluralist framework for gang research*.

Wegerhoff, D., Ward, T., & Dixon, L. (2020). A pluralistic approach to the definition, classification, and explanation of gangs. *Aggression and Violent Behavior, 58*, 101546.

Wegerhoff, D., Ward, T., & Dixon, L. (2022). Epistemic pluralism and the justification of conceptual strategies in science. *Theory & Psychology, 32*(3), 443–466.

5

Fleshing Out the Concept, and Questions of Classification

In the last chapter I considered the nature of mental disorder through an enactive lens while separating the structural and normative 'dimensions', referring to these as the 'the bones' of the conceptual model being proposed. In this chapter, I will put these normative and structural considerations together and attempt to 'flesh out' the fuller conceptual model of mental disorder that emerges. This will be achieved in various ways. This will include describing the concept through some complimentary lenses/perspectives, briefly applying the concept to the example of anxiety disorders, evaluating the concept using Zachar and Kendler's (2007) conceptual taxonomy, and comparing the concept to the most relevant of the conceptual models explored in Chap. 2. I argue that the 3e Psychopathology framework encourages principles of classificatory humility and classificatory pluralism, and also briefly outline a moderate anti-universalist stance. Before bringing the chapter to a close I suggest that some current neurodevelopmental and personality disorder concepts appear underjustified under the developed view and briefly draw links to the neurodiversity movement.

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7_5

5.1 Integrating into a Fuller Concept

In the previous chapter I argued that in a structural sense, the 3e view reveals mental disorders as repetitive patterns in sense-making, with causal structures best thought of as stable dynamic patterns across the brain-body-environment system. I suggested this pattern can be thought of as an MPC-kind structure, spanning the brain, body, and environment. I also highlighted several useful conceptual tools within 3e cognition that are likely to be of benefit when seeking to understand and explain mental disorders. In the second part of the previous chapter I showed that enactivism subscribes to a view of normativity as emergent for self-maintaining complex systems, and thus features the tools required to develop a sophisticated systems-functionalism as a basis for the labeling of certain behaviors as dysfunctional or disordered. On this view mental disorders are patterns in the way that a person makes sense of and engages with the world that run counter to their own mode of functioning to a significant or atypical degree.

Bringing these ideas from the structural and normative domains together, mental disorders can be seen as *dysfunctional patterns in sense-making* (i.e., dysfunctional patterns in behavior, thought, emotion, and all other adaptive/experiential processes). These dysfunctions are constituted by relatively stable dynamic patterns (i.e., networks of phenomena) within the brain-body-environment system of individuals. These patterns in sense-making can be defined as dysfunctional because they run significantly counter to the person's mode of functioning in the world. These dynamic patterns then are dysfunctional process-structures within the adaptive striving of agents, distinguishable by their negative functional effects, inflexibility, and often complex and circular causal structures.

5.2 A View from Some Different Angles

Taking a perspectival approach, where different modes of description can be seen as complimentary models of the same aspect of reality (Chang, 2020), it is interesting to consider this understanding of mental disorder

in different ways, or from different perspectives. One way to approach this concept is through the more dialectical language of autonomist sensorimotor enactivism (Barandiaran, 2017; Di Paolo, 2010; Di Paolo et al., 2018). Describing the concept from this perspective, mental disorders can simultaneously be seen as *parasitic partial autonomies* within the process structures of human functioning (i.e., mental disorders themselves partially 'self-maintain' within the context of the brain-body-environment system). This partial autonomy is dependent on—but in tension with—the biological, sensorimotor, and other adaptive partial autonomies of the host organism, and conflicts with the normative structures that arise from these, to the detriment of the organism's adaptive agency, flourishing, and/or likelihood of survival (Nielsen, 2020b). Such a description highlights the congruity of 3e Psychopathology with Maiese's (2021) perspective. Such an approach has value because it emphasizes the partial entitativity of mental disorder, helping to make sense of how we can meaningfully externalize disorders despite the fact that mental disorders exist as part of a person's own experience and behavior—e.g., to speak of the likes of anorexia as if it is separate to the sufferer. However, it is important to note that I am sticking to what I have in Chap. 3 labeled the CASE interpretation of so-called emergent autonomies within human functioning. While we can meaningfully speak of mental disorders as having emergent autonomy, this is insufficient to grant them any strong sense of separable identity from the wider organism within whose functioning they reside, and any reference to 'identity' in this way should be considered metaphorical. Mental disorders are processes within the agent-world system and share constitutive overlap with normal human functioning. They are therefore not so much embedded in human functioning like a creature in an eco-system, but rather are constituted by/enmeshed with the wider processes of human functioning in which they exist. While we may be able to distinguish mental disorders by their dysfunctional effects and partial autonomy, they remain parts/aspects within the wider functioning of the organism and thereby cannot be seen to give rise to normativity separable 'identities' from that of the wider organism. This alleviates concerns about promiscuous normativity mentioned in Chap. 3. Considering the concept from this perspective however, allows us to see that, in a certain sense,

mental disorders are patterns in sense-making that begin to serve their own continuation rather than the functioning of the individual.

Another perspective from which to consider this concept is through a metaphor to fluid dynamics. While behavior in general has a tendency to 'flow' towards adaption and self-maintenance, mental disorders are process-structures that 'flow' in the opposite dysfunctional and entropic direction. Expanding on this, take a river to represent the processes of human behavior, and the ocean to represent the striven for state of self-maintenance. Stagnancy therefore represents death, and the general tendency of the river to flow towards the ocean—and to carve its own path through the landscape –represents adaption. Occasionally there are bends and rapids that cause the water to tumble as it finds ways around such obstacles. Such challenges to the rivers flow represent the normal trials and tribulations of life. Along the way, in interaction with these obstacles, eddies often form. These are normal back-flows in the fluid-dynamics of the river, representing normal but less-than-ideal behaviors; such as eating too much chocolate or staying up too late. In such eddies the behavior is non-adaptive, but the flow is largely unimpeded (i.e., it is not dysfunctional/disordered). Within this image, mental disorder may start as an eddy, but gets larger and more persistent. Carried by the force of its own adaptive momentum and shaped in interaction with the dynamics of the landscape it flows through, the water cycles back around on itself, dynamically constituting a recurring pattern in the flow of the water and wearing its way into the bank until a pond or deep pool is formed. The water still flows to the ocean, but its progress is significantly slowed-down; it risks stagnation.

5.3 Anxiety as an Illustrative Example

Briefly applying this 3e conceptual model to anxiety disorders as an illustrative example only, I would argue that it provokes a much richer understanding than current approaches. Anxiety disorders are traditionally defined as levels of vigilance and/or fear, disproportional to actual threats, to a degree that is atypical and produces significant harm or impaired functioning (American Psychiatric Association, 2013a). Rather than

assuming this pattern of behavior is caused by an underlying brain lesion (i.e., biological essentialism) or an error or difference in cognition (i.e., psychological essentialism), 3e psychopathology would consider an anxiety disorder as a constitutionally complex network of phenomena within the brain-body-environment system supporting a pattern in sense-making that conflicts with the mode of functioning of the agent. Such a conceptualization also takes us further than the broad strokes holism of the likes of a biopsychosocial approach in that the factors that constitute and contribute to the wider pattern are integrated within the processes of the striving, purposeful, and meaning-experiencing individual.

To begin the example, some of the behavioral and experiential phenomena that together constitute the most obvious aspects of a pattern of anxiety may include: increased affective experiences of worry/fear/doubt, perceptual biases towards potential threat, frequent thoughts about the perceived threat/s, avoidance behaviors, a tendency toward short and shallow breaths, repeated checking behavior, fatigue, sleep disturbance, distraction and irritability, etc. Note how these phenomena relate to differences in how a person acts in and makes sense of the world—i.e., to the differences in sense-making observed. They primarily concern alternations in the *meaning experienced* and *actions taken* by the individual. Heuristically then, we can label these as the *enactive* components of the pattern.

Each of these component behavioral and experiential phenomena are themselves necessarily *embodied*, and therefore the current view immediately provokes questions as to how the observed behavioral and experiential phenomena are themselves constituted or causally influenced at the biological scale[1]. In anxiety these factors likely include but are not limited to phenomena such as:, sympathetic nervous system activation, CO2 and cortisol levels in the blood, the activity and structure of neural circuits and anatomical systems such as the amygdala and HPA axis, changes to circadian rhythms, changes to neurotransmitter and receptor expression

[1] Note here I am lumping together a variety of classically differentiated scales of enquiry, from the physiological to the neurological and the genetic. The same is true for the variety of timescales I have lumped under factors relating to embedment. I have done this simply for ease of communication and to keep this example brief.

in different parts of the brain, genetic polymorphisms, epigenetic factors, gut micro-biota balances, etc.

Further, as an aspect of a person's functioning in the world, it must also be considered how this wider pattern in sense-making is influenced/occasioned by their environment, both currently and across development. Heuristically we can refer to these as the embedded factors. These will include direct causal links from environmental factors to the behavior, but also indirect causal links via the constituent biological factors. Examples include but are not limited to: current stress and access to basic needs, childhood attachment history, the actual threat level of previous environments, modeling of anxious behaviors by others and how this was experienced, dietary nutrients, exposure to drugs including licit ones such as caffeine and alcohol, relationship history (parental, familial, plutonic, and romantic) and whether these relationships supported the development of self-efficacy and confidence, gender norms concerning management of distress, the culturally and family mediated understanding of what it means to be anxious and how to develop a sense of security in the world, history of traumatic experiences, etc.[2]

Vitally, under 3e psychopathology, these various causal and constitutive factors across the brain-body-environment system are not simply free-floating entities which can be summed up in some linear way as to result in the observed pattern of dysfunctional behavior. Rather, they are affixed within the wider functioning of the individual, interacting in dynamic ways. It is when looking at the wider pattern that we may observe *regularities* which seem to play significant roles in the emergence and maintenance of the dysfunctional behavior. As such these regularities can be labelled as *mechanisms*. In anxiety disorders, candidate examples of such mechanisms include the negative reinforcement cycle of avoidance and relief, sleep disruption negatively impacting the ability to reason with one's self, or the way that shallow breathing is often a habitual response to fear/anxiety yet through its impact on blood chemistry and the felt-meaning of associated experiences often compounds the experience of fear/anxiety. As I will aim to show in the following chapter, the aim of

[2] For further example and comparison to extant conceptual models see table two in Chap. 7.

explanatory theory under 3e Psychopathology is to capture and explicate such mechanisms and their wider constitutive structure.

Finally to complete this example, we must consider how this pattern is a problem for the individual diagnosed—how it conflicts with their mode of functioning. In anxiety, common ways this may be the case might include but are to limited to: disruptions to basic functions such as eating or sleeping; engaging with self-injurious behavior in an effort to manage the anxiety; being so overwhelmed by the anxious thoughts and feelings that it is hard to focus on personally important things such as hobbies, work, school, time with friends and family; damaging important relationships through constant seeking of reassurance; disruptions to the ability to serve ones various roles; etc. This notion of conflict with an individual's mode of functioning is congruent with the DSM definition above and its reference to impairment in important areas of functioning. However, the nature of this impairment is not prescribed under 3e Psychopathology, rather it needs to be evaluated in every individual case in reference to how a person lives their life—i.e., is this a problem *for them*.

5.4 Getting More Precise

To add further detail and precision to the sketch outlined so far in this chapter, I will evaluate the concept against a six-factor conceptual taxonomy presented in Zachar and Kendler (2007). I have previously done so in Nielsen and Ward (2018), however the current discussion, parallel to Nielsen (2020b), incorporates further developments regarding questions of classification and is a useful summary of key conceptual elements of 3e Psychopathology.

Causalism/Descriptivism

This factor relates to the question "Should psychiatric disorders be categorized as a function of their causes (causalism) or their clinical characteristics (descriptivism)?" (Zachar & Kendler, 2007, p. 557). 3e Psychopathology conceptualizes mental disorders as recurring and

dysfunctional patterns in sense-making, constituted by relatively stable dynamic patterns of causal relations within the brain-body-environment system (Nielsen, 2020b; Nielsen & Ward, 2018). This position emphasizes both the constitutional and causal complexity of these patterns, and the fact that these patterns are embedded within the processes of human functioning, i.e., the embodied sense-making of individuals. As such, it is assumed that every token/instance of disorder is likely to differ to a reasonable degree to others tokens/instances of the disorder. This includes the presence of differing causal mechanisms across individuals seen to have a similar pattern of disorder (see the discussion regarding entities/agents below). In other words, given the complexity at hand, we cannot reasonably expect that perfectly stable diagnostic entities can be extrapolated, and—to some degree—etiological heterogeneity is to be expected and accounted for by individual clinicians on the ground. However, this is not to mean that *similar enough* patterns across individuals cannot be identified and labeled as diagnostic entities. By similar enough I here mean that successful diagnostic concepts accurately refer to patterns in the world to the degree that they can facilitate: 1) the development of nomothetic explanatory models which highlight common mechanisms and their interaction within a particular pattern of disorder, and 2) the development and selection of targeted treatment approaches. Under 3e Psychopathology then, classifying mental disorders based on their causes is seen as a hugely challenging but conceptually valid endeavor, to be guided by the pragmatic nature of the classificatory task.

Under this view of disorder and of the role of classification, a rough distinction can be made between two kinds of mechanisms at play in mental disorders based on their function. Specifically, those that make some individual more likely to fall into a given pattern in the first place (i.e., etiological mechanisms), and those that contribute to the stability of the disorder pattern despite is dysfunctionality (i.e., maintenance mechanisms). Arguably, maintenance mechanisms, being more causally proximal, seem more likely to be common across different token cases of a disorder, and also more relevant to treatment. As such, maintenance mechanisms seem likely to be far more useful concepts for the demarcation of diagnostic entities. It is very conceivable under this view that mechanisms of both kinds may turn out to overlap across different

patterns of disorder (i.e., this view makes room for transdiagnostic mechanisms). It is likely that, over time, knowledge of such mechanisms across different patterns of disorder will continue to expand. Our understanding then, of causal and maintenance mechanisms, seems likely to shift from more general to more specific until an optimal level for the pragmatic purpose of classification is reached, as specified above (Nielsen & Ward, 2018).

It is also useful to point out that the distinction between causal and descriptivist approaches to classification does not have to be binary. The current discussion highlights the possibility that we may maintain our current descriptive concepts for a time, while slowly incorporating recognized causal mechanisms via a parallel classification system (perhaps in the form of an online database of proposed mechanisms). This would produce an intermediary diagnostic approach while our causal knowledge is still nascent, whereby potential causal mechanisms can be tacked on to individual's diagnoses allowing for the development and application of more targeted treatment approaches (for further discussion see this blog post: Nielsen, 2022b).

Essentialism/Nominalism

This factor relates to the question "Are categories of psychiatric disorder defined by their underlying nature (essentialism), or are they practical categories identified by humans for particular uses (nominalism)?" (Zachar & Kendler, 2007, p. 558). The orientation of 3e Psychopathology within this factor is worth exploring. To begin, consider that despite the complexity of human functioning, we still see recurring patterns of dysfunction emerging across people, as exampled by current diagnostic concepts—imperfect and heterogeneous as they may be. This suggests that there is some tendency within the dynamics of the nomothetic human brain-body-environment system to fall into these similar enough patterns—parallel to the concept of attractor basins in dynamic systems theory. Under 3e Psychopathology it is this tendency for dysfunctions in sense-making to fall into similar enough patterns that allows us to see disorders as 'real'. This is meant in the sense that we can see them as fuzzy

but legitimate kinds/phenomena discovered in the world as opposed to idiosyncratic instances of human suffering (which are also very real but in a different sense).

On this view mental disorders can be understood as *moderately essentialist* in that they are defined by their nature as 'sticky tendencies' in the human brain-body-environment system bound together by similarities in their constitution rather than by sharing some singular stable essence (Nielsen, 2020a; Nielsen & Ward, 2018). They are analogous to biological species in this regard, although they have no causal lineage as species do. The pools of water mentioned in Chap. 2 seem a better analogy then. This type of multiply realized kind is referred to as a "*type-causal*" kind (Magnus, 2014a). The core concept I am describing here could therefore be described as a fuzzy type-causal MPC kind. Note how this brings a certain flexibility; some mental disorders may have tighter hubs of causal relations within the brain (such as, arguably: ADHD, schizophrenia), and some may be more diffuse (such as, arguably: alcohol dependence, depression)[3]. All of these cases can be described as fuzzy type-causal MPC kinds with causal structures spanning multiple scales, but the distribution of causal influences across these scales likely differs (Kendler, 2012b).

The essentialism/nominalism continuum has particular relevance for the task of classification. In turning to the task of classification, we must remember that a classification system is a practical human endeavor, and will therefore always be influenced by pragmatic and concerns and its own historicity (Zachar, 2018). Furthermore, given that the view of mental disorder presented in this book acknowledges the complexity and multi-scale nature of causal structures supporting dysfunctional behavior, it seems very unlikely that a classification system is going to accurately 'carve nature at its (fuzzy) joints' any time soon. Because of these reasons it seems important to make a clear distinction between 'the reference' and 'the referenced' when thinking about classification systems. Even though the current view holds mental disorders (the referenced) to be real,

[3] Some disorders may even have dense hubs of connection in the environment and thus in a sense be 'top-down' disorders. This seems to be the case with dissociative identity disorder (i.e., 'multiple personality disorder'), or other 'culturally bound' syndromes. I put quotes around 'top-down' as I do not wish to imply a hierarchy here, nor to fail to recognize the relational quality of such disorders.

diagnostic entities (the reference) should not be viewed as completely 'real'. They will likely always, or at least for a very long time, remain imperfectly constructed representations of the mental disorders that they are trying to capture. Thus, the current view makes a distinction between mental disorders in nature (seen as moderately essentialist type-causal MPC kinds), and mental disorders as diagnostic concepts (which are probably best described by what Zachar and Kendler call *moderate nominalism*; i.e., they are always partially shaped by our purposes, needs, socio-cultural values, and historical decisions).

Objectivism/Evaluativism

This factor relates to the question "Is deciding whether or not something is a psychiatric disorder a simple factual matter ("something is broken and needs to be fixed") (objectivism), or does it inevitably involve a value laden judgment (evaluativism)?"[4] (Zachar & Kendler, 2007, p. 558). At first pass, the conceptual framework presented here comes down clearly on the side of *evaluativism*. I mean this in the sense that, from a 3e view, values and norms are a vital component in the conceptual fabric of mental disorder. However, I want to partially follow de Haan (2020b) here, in noting that there is a conflict between the very nature of this proposed 'objectivism/evaluativism' continuum and the central tenants of enactivism. As explored in Chaps. 3 and 4, enactivism sees normativity as continuous with the natural world rather than as something that must be expunged to reach some 'objective view'. Norms and values are a part of the natural world when viewed through a non-reductionistic lens, and so it makes no sense to oppose evaluativism with objectivism as Zachar and Kendler do in their taxonomy (Zachar & Kendler, 2007). Instead, the DCT offers us a way to collapse the normative gap and see mental disorders as both objective things in the world *and* as strongly dependent on the negative implication of the sufferer's functional norms. Under this view, norms/values are seen as ubiquitous, and therefore necessary for a comprehensive understanding of human behavior and functioning.

[4] Note that something being broken and needing to be fixed also arguably entails a value judgement and prescription.

When I label the 3e view of psychopathology as *evaluative* it is this thorough-going role of values that I am attempting to highlight, rather than that a behavior being disordered is somehow not objective or less real.

Internalism/Externalism

This factor relates to the question "Should psychiatric disorders be defined solely by processes that occur inside the body (internalism) or can external events also play an important (or exclusive) defining role (externalism)?" (Zachar & Kendler, 2007, p. 558). 3e Psychopathology holds that both internal and external factors, as well as the interactions between them, are vital for a complete understanding of behavior and disorder (as per the principle of embedment). The current concept would therefore fall under what Zachar and Kendler refer to as *moderate externalism.* This position would be flanked on one side by 'internalism' which basically refers to the idea that everything important is happening inside the organism; an overly reductionist view from the 3e perspective[5]. On the other extreme we would find a 'total' or 'radical' externalism which might hold that mental disorder is always caused by socio-cultural factors such as the stresses of capitalism; also an overly reductionist view from the 3e perspective.

The positioning of the current framework on such a 'continuum of externalism' can be further specified using Roberts, Krueger, and Glackin's (2019) taxonomy of externalist positions regarding mental disorder. This taxonomy separates between different classes of externalism regarding mental disorder. Under Roberts, et al.'s taxonomy the current view can be referred to as an example of 'relational externalism'. This is a position that holds mental life—and psychopathology—to be relationally constituted and therefore inherently interactive. At its simplest the current

[5] Internalism could be further separated into those that think everything important is happening a holistic physiological level, the level of 'neuro-circuitry' (such as RDoC), at the level of brain-chemistry, at the genetic level, etc. 3e Psychopathology rejects all such views by its commitment to embedment (the recognition of the contextually dependent nature of behavior), and the taking of the whole brain-body-environment system as its focus of analysis.

perspective holds that mental disorder is a recurring pattern in sense-making that has the normative status of being significantly dysfunctional for the individual. Through the concept of embedment, we can see that the environment (both physical and socio-cultural) plays a vital and likely non-linear causal role in shaping and occasioning any such pattern in sense-making. It also plays a large role in determining the viability and therefore the normative status of resulting behaviors (i.e., different behaviors work in different environments). Both the pattern itself and its normative status are therefore deeply contingent upon organism-environment relations and it would therefore seem reasonable to refer to it as 'relationally constituted'. To clarify however, I do not see environmental factors as a *constitutive* part of the resulting behaviors or of cognition itself. As discussed in Chap. 3, while I am sympathetic to relational externalism of mind, I avoid constitutive expansion of the mind (i.e., 4e/extension).

Entities/Agents

Zachar and Kendler (2007) describe how the entity position generally views "…individuals as *vehicles* for pathological syndromes…", while the agentic position holds that "…each psychiatric disorder as manifest in an individual patient is relatively unique." (p. 559). The current framework would certainly view each manifestation of dysfunctional behavior as unique in important ways. Moreover, the very reason a cluster of phenomena should be seen as a *disorder* at all is because it will ultimately run counter to the functional norms of the agent. This concept of disorder is therefore inextricable from a purposive and agential view. 3e Psychopathology would therefore be seemingly best described as sitting under the *agential* position (Nielsen & Ward, 2018).

This however cannot be the full story. Firstly, consider how, as discussed above, 'entities' in the form of fuzzy but meaningful regularities across agents (e.g., disorders) can be extrapolated from the wider complexities of human functioning. Secondly, consider that under this 3e view, mental disorders are seen to exist as process-structures within our embodied sense-making—to be composed of many component structures both personal/agential and sub-personal/automatic/mechanical.

Lastly consider how, when exploring the current concept through the language of autonomist sensorimotor enactivism earlier, we saw how we can meaningfully speak of a disorder process maintaining itself as an entity. What I am challenging here is the idea that an agential view is necessarily opposed to an entitative one. Rather than a single continuum, entitativity and agentiality are better seen as two separate but non-orthogonal continua (non-orthogonal in that a fully entitative or disease model concept seems incompatible with an agential view). The descriptions of mental disorder at the start of this chapter described mental disorders as process-structures within the agent-world system, discernable in part by their significantly dysfunctional effects. This description highlights both the ultimate dependency of the disorder process on the striving agent, as well as the conceptual separability of disorder from the agent. Overall then, 3e Psychopathology is certainly an agential view, but it also recognizes a certain entitativity to many disorder processes[6]. A great example of this is anorexia nervosa, often spoken of as having an identity separate to that of the suffer.

Categories/Continua

This factor relates to the question "Are psychiatric disorders best understood as illnesses with discrete boundaries (categorical) or the pathological ends of functional dimensions (continuous)?" (Zachar & Kendler, 2007, p. 559). The current view presents mental disorders as dysfunctional patterns in an agent's striving, constituted by many interrelating causal factors across the brain, body, environment system. Many of these compositional factors will themselves be continuous in nature, and the interaction between them complex and chaotic, therefore, the constituted patterns of behavior seem very unlikely to be definable in a cleanly categorical manner. There may also be compositional overlap between individual disorders, despite their being important differences between them. For example, on this view two different kinds (or perhaps

[6] I wish to acknowledge that de Haan (2020b) has extended my thinking on this issue since publication of Nielsen and Ward (2018).

sub-types) of depression may hypothetically be isolated on the basis of the presence or absence of some important mechanism, while still sharing other important mechanisms and features. A large degree of fuzziness and continuity therefore seems to be predicted by the 3e view, as can be seen by the description of mental disorders as fuzzy type-causal MPC kinds when discussing essentialism/nominalism. The precise degree of fuzziness however, will be different across different disorders and is really an empirical question. A blanket statement committing 3e Psychopathology to either a continuous or reasonably categorical view would therefore be inappropriate. Instead, as argued by Kendler, Zachar and Craver (2011), viewing the constitutional structure of mental disorders as fuzzy MPC kinds affords a large degree of flexibility—some disorder may turn out to be reasonably discrete, and some may turn out to be nigh on continuous. To stress however, this not to say that the differences between different kinds of mental disorder, nor the distinction between non-ideal behavior and dysfunctional behavior, will be meaningless. Rather these distinctions are based on the degree of functional impact, and empirical regularities across kinds respectively[7]. These boundaries are certainly fuzzy, but far from arbitrary (Nielsen & Ward, 2018).

5.5 Comparing Conceptual Models

Now that the central concept has been explored in detail it seems pertinent to make comparison to some of the more relevant extant conceptual models explored in Chap. 2. Not all models will be compared in-depth, rather focus is given to popular views, and those that provide an interesting contrast.

[7] As argued by Zachar (2014), this genuine fuzziness invites pragmatic decision making in the development of diagnostic systems. The problem is of course the divergent purposes that these diagnostic systems are meant to serve. In the service of different purposes (e.g., explanatory efforts, the treatment of individuals, the development of talk therapies, the development of pharmacological treatments, diagnosis as relevant for legal decisions) different degrees of abstraction may be pertinent. How those performing the task of classification should respond to these different needs however is well beyond the scope of the current project.

Structural Models

As stated, within the structural dimension, the model of mental disorder developed here aligns best with a fuzzy MPC kind view. It is also explicitly recognized that the causal structure of mental disorders likely spans brain, body, and environment. The enactive view however, also reminds us that this pattern is not occurring in a vacuum, or as some entity that can ever really be completely abstracted out from the agent concerned. Instead, mental disorder is considered to be residing within the wider process structures of the striving individual (themselves in context), and any abstract model of the 'disorder process' is seen as necessarily an idealization (although hopefully a useful one for explanation and the development of treatments). 3e Psychopathology therefore, is more strongly agential than the MPC view.

The most interesting structural model to contrast the 3e view with however, is the likely the essentialist notion of mental disorder. In particular, I want to focus on two places of near similarity between the essentialist and 3e views, so that it can be seen more clearly what separates them.

Firstly, on the current view and as predicted by essentialism, some disorders may, in time, be discovered to feature an 'underlying' hub of tight causal connections that are central to the disorder process. This causal hub of activity may potentially (but not necessarily) be 'in the brain'. For example, we may discover that some causal sub-type of depression reliably involves some alteration in some neural network 'X'. However, on the current view, to conclude from such a discovery that such a hub represents the 'essence' of a disorder would likely be mistaken. The discovery of such a hub—and coming to understand its mechanistic relation to the wider pattern of dysfunctional behavior—would obviously be hugely useful (hence my argument in the following chapter that reductionistic strategies such as RDoC are likely to be fruitful). However, coming to see such a hub as 'the-disorder-proper' would likely represent a gross decontextualization under the developed view. Rather, the current view would suggest seeing this hypothetical hub of causal connections as an identified biological mechanism among many other mechanisms that can support a depressed presentation. Dysfunction lies in the wider pattern of behavior

and its (lack of) adaptivity for the agent in their environment; it exists in the relation between the organism and their environment. To abstract away from this complexity and instead focus on some apparent 'essence' risks reifying the mental disorder in question as a 'disease' or brain pathology, leading to a hyper-focus on the apparent disease processes at the expense of coming to understand the person and their context. To continue the above example, depression is so much more than simply a brain disease; coming to recognize some important brain mechanism at play should not change our recognition of this. The exception to this is if it can be shown that the disorder in question really is better thought of as a disease entity—that the proposed 'essence' really is the one key constitutional factor at play and can truly be said to cause the dysfunction. In this case however, the disorder in question seems to look more like a brain pathology with behavioral symptoms (e.g., Parkinson's), rather than being a mental disorder[8]. This divide—between a brain/physiological pathology with behavioral symptoms, and mental disorder—is better seen as continuous rather than categorical. Recognizing this continuity doesn't fit with the essentialist view. By taking a fuzzy/MPC kind view, the current position allows for recognition of this continuity because, as mentioned in Chap. 2, the notion of an MPC is flexible; able to capture more, or less, heterogeneous clusters.

Secondly, the 3e view developed here sees mental disorders as real patterns to be found in the world; it is a mode of realism about mental disorder, despite recognizing the fact that our references to mental disorders are near inescapably historical/constructed[9]. We can draw a parallel to essentialism here because the realist commitments of both positions mean that a causalist classification system is seen as a genuine possibility. Contrary to an essentialist view however—where types of disorder would be clustered due to a shared essence—the current view would prescribe

[8] Note the similarity to a Szaszian position here in making a distinction between a medical disease and mental disorder. Contra Szasz, this distinction is seen as fuzzy, and the current position also carves out a distinct conceptual space for mental disorder in a way that Szasz did not.

[9] Realism and essentialism—along with internalism—are often conflated. See Hartner and Theurer (2018) for discussion (although note that I disagree with their ultimate conclusion as it seems to rest on a the assumption that normativity cannot be part of the natural world—hence ruling out mental disorder as a fruitful target for mechanistic explanation).

classification based on the *similarity* of the causal patterns/mechanisms supporting disorder as described in the previous section. Again, we can see similarity to the MPC kind view.

Normative Models

Turning to the normatively focused conceptual models, it should by now be apparent that 3e Psychopathology presents a much richer and justifiable variety of 'functionalism' than either the statistical or evolutionary positions. By its situation within a richer framework of human functioning, the current concept moves beyond considering what the individual 'should' be able to do according to either evolutionary or statistical norms at a species or reference-class level—positions which we saw in Chap. 2 face significant limitations. Instead, the 3e view recognizes that the assessment of somebodies functioning is always, in a certain sense, evaluative.

Those well-versed in the philosophy of psychiatry may take opposition here, for evaluativism is traditionally seen as the antithesis of functionalism. Evaluativism sees mental disorders as irreducibly value-laden, while functionalism attempts to find an objective demarcation between the disordered and the benign. Pervasive in western thought is the idea that values/norms and objectivity don't mix; thus, the observed tension between evaluativism and functionalism. As foreshadowed by Thornton (2000) however, enactivism allows us to plug the normative gap; to collapse this dichotomy and move to a synthesis view. This is because enactivism recognizes that—assuming naturalism—if values and norms exist then they must simply be part of the natural world. Through its commitment to the DCT, enactivism offers an account of norms and values as arising for particular organizational structures in the material world; i.e., purposive and precarious systems, striving to self-maintain and adapt (Di Paolo, 2005; Maiese, 2016). Hence, the current concept of mental disorder is evaluativist, yet no less real because of it, for it is tied to the real functional norms of the individual diagnosed—to the adaptive fit between the behavior of the organism and its environment. To again put this most simply: attempting to meet the individual where they are at and making a considered evaluation as to whether the observed pattern in sense-making is working for the individual.

This move will hopefully go a long way in satisfying the evaluativist because it recognizes a role for norms and values in the concept of disorder; namely those norms that support the individual's self-maintenance, adaption, and wider ways of living. However, on the other side this move also avoids Foucauldian-style critique, because it is not the norms of society that are seen as at issue, only those norms that support the adaption and self-maintenance of the individual. Hence the current concept cannot be seen as a socially constructed label for deviance, rather it provides a conceptual route to offering diagnosis in the interest of the client. Observations such as those of Stier (2013)—that wider normative features such as the values of the clinician often play a role in diagnostic decisions—would therefore be seen as erroneous influences rather than reflective of what a diagnosis should represent; i.e., that the client is acting against their own best interests.

Briefly comparing to pragmatism before moving on, as explored when discussing essentialism vs. nominalism earlier, the practical and political realities of generating a classification system mean that diagnostic entities are never going to be perfect representations of the patterns of dysfunction people experience. Mental disorders as represented in our classification systems will always be biased and distorted by the needs and values of the groups generating those systems, as well as the practical limitations placed upon such groups (Zachar, 2018). 3e Psychopathology is therefore not opposed to the observations of moderate pragmatists who recognize this degree of nominalism regarding diagnostic entities. Through its commitment to realism however, 3e Psychopathology is in direct conflict with the radical pragmatism/near total nominalism of the likes of O'Connor (2017). The rejection of nominalism, and commitment to an ideal concept of what mental disorder is, above and beyond how it is used, means that the current concept offers exponentially more guidance than a radially pragmatic position.

5.6 Questions of Classification

The current view supports a gradual shift towards a causalist diagnostic system as our explanations of mental disorder develop. Mental disorders are causal structures in the world which we can come to understand and

categorize on the basis of similarity. At the same time, under the current framework the task of classification is seen as a partially pragmatic endeavor. This is due to the current paucity of causal understanding regarding mental disorders, the presumed complexity of mental disorders, and the fact that—as per Zachar (2018)—diagnostic systems are inherently political documents. Diagnostic systems will always be subject to social needs and pressures in accordance with this, and as such will likely never represent a 'perfect' natural ontology within their domain (even if such a thing is possible within an ever-changing socio-cultural and technological environment). Researchers and practitioners who wish to take a 3e approach should therefore be ever critical of the nosological system of their time, although this is hardly particular to the 3e view.

Before closing this chapter it is worth considering some further implications of 3e Psychopathology for attempts at the classification of mental disorder. This section will briefly propose that 3e Psychopathology aligns with: humility and pluralism regarding classification, moderate anti-universalism, culturally specific nosologies, and the view that alternative modes of functioning are not mental disorders.

Classificatory Humility and Pluralism

The limitations of current classification systems for mental disorder have not been a focus of this book, and I do not wish to make it so. This said, it is important to briefly acknowledge the significant limitations of dominant efforts such as the Diagnostic and Statistical Manual (DSM). For those not familiar, the DSM assumes a descriptivist and categorical approach to classification, isolating a large number of diagnoses on the basis of signs, symptoms, histories, and (usually) the presence of functional impairment. Diagnosis is made based on an algorithmic set of criteria designed to increase the reliability of diagnoses made across clinicians (i.e., the likelihood that two clinicians diagnosing the same client offer the same diagnosis). This focus on reliability in the development of the DSM is often questioned, as it does not necessarily make it more likely that diagnostic constructs are picking out valid patterns in nature with common causal processes (Lilienfeld & Treadway, 2016; Zachar &

Kendler, 2017). There are in fact many recognized issues with the DSM model that give us reason to doubt the etiopathological validity of its diagnostic constructs (for review of these issues see; Karter & Kamens, 2019; Lilienfeld & Treadway, 2016; Zachar & Kendler, 2017). Some key problems include: artefactual co-morbidity[10] (Andrews et al., 2002), symptomatic and etiological heterogeneity[11] (Lilienfeld, 2014), false positives[12] (Cooper, 2013; Wakefield, 2015), concept creep[13] (Haslam, 2016), and the problem of reification[14] (Hyman, 2010).

As presented across the last two chapters, 3e Psychopathology holds that meaningfully similar patterns of dysfunctional sense-making appear to exist across people, that these patterns seem likely to feature common mechanisms of dysfunction, and that, as such, classification on the basis of causes is a conceptually valid endeavor. However, it also highlights the complex nature of the causes at play, and that the causal structures of disorders are likely to be somewhat different across different token cases of a given diagnosis (even in the case that our diagnostic concepts show improved stability with time). There is therefore a congruity with the transdiagnostic efforts of the Research Domain Criteria (RDoC), a biologically focused effort by the National Institute of Mental Health in the United States which seeks the eventual development of a causalist classification system based on the mechanisms that underlie distress. However,

[10] Co-morbidity refers to when an individual has more than one mental health diagnosis at one time. Under the DSM, this occurs at much higher rates than would be expected if mental disorders were independent phenomena, suggesting that this may be an artefact of how we conceive of and measure our diagnostic concepts. Note that there is continuing debate on this issue.

[11] Heterogeneity refers to diagnostic constructs being too 'large', capturing meaningfully different individuals under the same label. This can include individuals with very different symptom profiles (symptomatic), and/or disorders with very different causes/constitutions (etiological). Under the DSM this occurs frequently (Contractor et al., 2017; Dickinson et al., 2017; Galatzer-Levy & Bryant, 2013; Hawkins-Elder & Ward, 2019; Monroe & Anderson, 2015; Olbert et al., 2014).

[12] False positive refers to when people are diagnosed as having a disorder but probably do not have the disorder/a genuine problem.

[13] Concept creep refers to the observed tendency for our concepts of harm to grow over the last hundred years or so. I include this here as the cited paper by Haslam includes many examples from the DSM. If DSM concepts can expand (or contract) with social mores, this brings into question their objective nature.

[14] The problem of reification concerns the fact that DSM diagnoses are only intended to be draft descriptive concepts yet through their use have come to be seen as real things, often with causal power, to an unwarranted degree.

unlike RDoC which, 1) is focused on mechanisms of normal functioning gone awry, and, 2) holds that mental disorders are fundamentally disorders of brain circuitry (Insel et al., 2010), 3e Psychopathology holds that the fuzzy set of mechanisms that constitute the structure of a given disorder are free to exist across the brain-body-environment system. What exactly a causalist classification system based on this view would look like is an area for further development, however, there are some key principles regarding classification that 3e Psychopathology would encourage in the meantime. These key principles are *classificatory pluralism*, and *classificatory humility*.

Classificatory pluralism refers to the idea that that, given the complexity of mental disorders as conceived under 3e Psychopathology, there are likely many valid and useful ways of classifying them that are worth exploring (Markon, 2013). By this I am referring to the range of scientific classificatory approaches currently well discussed in the wider literature such as the RDoC or the Hierarchical Taxonomy of Pathology—a dimensional factor analytic model of the relationships between various disorders ([HiTOP] Kotov et al., 2017). However, I am also alluding to the idea of *culturally specific nosological systems*. The question of universalism—i.e., whether mental disorder concepts reflect the same set of difficulties across the world—has not been a topic of this chapter. To touch on this briefly however, under 3e Psychopathology genuine causal efficacy is granted to psycho-social factors and these are seen as a vital part of the constitution of the patterns we label as mental disorders (see discussion of cultural embeddedness in the previous chapter). As such, culturally specific syndromes and cultural differences in the experience/presentation of mental disorders cannot be seen simplistically as different 'manifestations' of the same underlying biologically defined phenomena. Rather culturally specific syndromes and presentations should be largely understood as separate concepts/phenomena in their own right, albeit with similarities often occurring (Kirmayer & Ramstead, 2017). Roughly then, 3e Psychopathology is opposed to a strong sense of universalism. However, given that we all share very similar bodies which bring with them similar biological mechanisms, and that there are deep similarities as well as differences across cultures, the 3e view would also predict significant overlap in the patterns of difficulty that emerge across different

cultures and environments. For example, phenomena such as *ataque de nervios* (an acute period of intense emotion and sense of being out of control observed in Latinx and Hispanic populations), and *koro* (an acute period of intense fear in men that one's penis is shrinking/disappearing observed across many cultures but typically associated with Chinese populations), share some similar features to the western psychiatric concept of a panic attack and may share some overlapping mechanisms, but are likely best viewed as their own distinct patterns/concepts (American Psychiatric Association, 2013b; Jilek & Mattelaer, 2006; Lewis-Fernández et al., 2002). The idea of culturally specific 'homegrown' nosological systems that draw on culturally specific concepts and values aligns with 3e Psychopathology and would seem to have potential in addressing issues related to systemic power relations in clinical settings where received western conceptualizations typically dominate to the disadvantage of underrepresented groups (Kirmayer & Jarvis, 2019).

The normative formulation of 3e Psychopathology, in being tied to functionality alone, removes reliance on statistical norms. This is advantageous because reliance on statistical norms can result in a pathologizing of modes of functioning that develop within cultures underrepresented in the sample from which the norms are gathered. For example, NiaNia et al. (2016) report multiple case-studies in an Aotearoa-New Zealand context in which young Māori who displayed 'symptoms' usually taken as indicative of psychosis were successfully 'treated' in a way that understood their 'symptoms' as culturally specific phenomena rather than as indicative of a recognized mental disorder. 3e Psychopathology would perceive such cases as falling into two categories demarcated by functionality: culturally specific disorder, and non-pathological culturally specific phenomena/experiences which are causing distress. For both categories the framework developed here would suggest the possibility of finding a mode of functionality that works for the individual and/or reduces distress. By decoupling from the received view of 'mental disorder', NiaNia et al. take a very similar approach—finding paths to functionality and/or alleviation through a collaboration between traditional and psychiatric approaches. Such an approach would seem inconsistent with an underlying concept of mental disorder based on contrast to statistical normality, underlying brain differences, our limited understanding of evolutionary

normality, or the values and norms of wider society, as such concepts would in practice entail a blindness to the unique socio-cultural milieu in which these young people's development was embedded. The approach does however cohere with 3e Psychopathology, its moderate anti-universalist stance, and its demand for deep consideration of the normative and cultural landscapes in which individuals reside when considering application of a label of disorder or dysfunction.

Classificatory humility refers to the idea that, given that same complexity, as well as the nascent state of classification in mental health, we should hold our current (and future) diagnostic concepts lightly at all levels. To example this at the individual level, no matter how thorough an assessment process, a diagnosis in mental health is a draft attempt at classifying an individual's challenges based on a draft classification system. It should therefore be reviewed regularly and be seen as only the roughest of sketches as to the sorts of difficulties an individual faces, facilitating use of a common language and links to the evidence-base. In short, a 3e Psychopathology orientation should encourage care to not over inflate the importance of diagnoses when developing treatment approaches, communicating with patients or general public, and when developing policy.

Alternative Modes of Functioning Are Not Disorders

A final implication of the 3e Psychopathology perspective is that certain current diagnostic concepts begin to appear suspect in regards to their conceptual validity. A central principle of the framework here developed is that alternative modes of functioning should not be considered inherently disordered. Rather, dysfunction is seen as deeply contextual and individually defined, and we must constantly be asking the question 'is this observed pattern in sense-making a problem for the individual in question?'. When observing patterns that exist across development or for a significant period of time then this question arguably becomes even more relevant, in that the observed patterns would appear to constitute stable parts of an individual's mode of functioning in the world. Current diagnostic concepts such as Autism Spectrum Disorder (ASD), Attention

Deficit Hyper-Activity Disorder (ADHD), as well as some personality disorder concepts such as schizoid and schizotypal personality, are clearly capturing *differences* in sense-making, but on the current view the presence of difference alone does not make a disorder. These diagnostic concepts rather seem to represent *alternative modes of functioning*, i.e., different ways of being in and relating to the world. Unless clear functional impairment is present then the current framework would hold that a label of 'disorder' would be inappropriate. In other words: people can be autistic without necessarily having a disorder; if it is not a clear problem for them then people can have difficulties regulating their attention or activity level without this counting as ADHD; and people can be a little bit odd or socially avoidant without this counting as a personality disorder if it does not clearly impact their functioning. The claim here is not that these longer standing patterns cannot be seen as disordered, but that they are not intrinsically so. The question at hand is, has a person been able to incorporate these differences in their sense-making into a wider mode of functioning that works *for them* and allows them to fair reasonably well in life?

It would appear therefore that the current analysis accords reasonably well with the neurodiversity movement, within which there are also arguments for a more contextualized understanding of behavioral functionality (e.g., see Chapman, 2021). However, the current framework also encourages consideration of the way in which alternative modes of functioning, even when not inherently disordered or dysfunctional, may represent significant alternations to the *landscape of potential pathology*. Just as we see differences in the patterns of distress and behavioral dysfunction observed across cultures, it makes sense that those with an alternative mode of making sense of and engaging with the world would also have a different set of 'sticky tendencies' within which they are vulnerable to getting stuck. For example, schizoid or schizotypal personalities, even if not themselves dysfunctional in a token case, do appear to share a relationship with the psychosis spectrum and thus may change the landscape of potential psychopathology that an individual faces both in expression and overall risk. As another example, while the current framework does not hold autistic persons to necessarily have a disorder (unless their autistic qualities get in the way of their functioning or faring well to a

significant degree), it also makes sense that autistic persons, defined as there are by similarities in their engagement with the world, may share particular vulnerabilities to dysfunctional patterns of sense-making that differ from the vulnerabilities of non-autistic persons. Indeed there may well be particular patterns of difficulty *specific* to autistic persons and other such alternative modes of functioning. However, much like the issue of cultural universalism discussed in the previous section, the investigation and classification of such 'mode-of-functioning specific patterns' should be led by the communities in question.

5.7 Summary

The current chapter has attempted to flesh out the conceptual model of mental disorder presented, and explore a range of questions relating to the task of classification. 3e Psychopathology holds that mental disorders are best understood as patterns within sense-making that work against the functioning of the individual concerned, constituted by a network of factors across the brain-body-environment system. I have explored this concept from multiple angles, both dynamic and metaphorical, and have exampled this concept with an application to anxiety difficulties. It has been shown that this conceptual model is agential, moderately externalist, and does not make a pre-empirical commitment to whether particular disorders are best seen as categorical or continuous, defaulting to a view of most disorders as fuzzy MPC type structures which we can begin to understand and define causally. It has been argued that the current perspective breaks down the natural/normative divide through seeing disorders as very real patterns of sense-making which may facilitate the crossing of one's own functional norms. As such it has been argued that the process of diagnosis should always be considered an evaluation. Despite seeing disorders as very real, the current view also holds that current and future diagnostic concepts are biased and imperfect representations of a messy reality, and thus should always be held lightly. Comparison has been made to essentialist and MPC-kind models, and the difference between a mental disorder and a brain pathology with psycho-behavioral

symptoms was explored. Comparison was also made to positions of evaluativism and other varieties of functionalism, showing how the current functionalist position holds novel ground as both natural and normative. Some further implications of the 3e Psychopathology conceptual framework have also been briefly explored including the endorsement of classificatory pluralism, classificatory humility, and moderate anti-universalism. Finally, links have been briefly drawn to the neurodiversity movement and a non-pathologizing stance on alternative modes of functioning has been outlined. The following chapter will shift in focus to the tasks of explanation and, to a lesser degree, treatment; addressing the question 'if this is what mental disorders are, then how should we best seek to explain them?'.

References

American Psychiatric Association. (2013a). Anxiety disorders. In *Diagnostic and statistical manual of mental disorders*. https://doi.org/10.1176/appi.books.9780890425596.dsm05

American Psychiatric Association. (2013b). *Diagnostic and statistical manual of mental disorders, 5th edn (DSM-5)* (5th ed.).

Andrews, G., Slade, T., & Issakidis, C. (2002). Deconstructing current comorbidity: Data from the Australian National Survey of Mental Health and Well-being. *The British Journal of Psychiatry, 181*(4), 306–314.

Barandiaran, X. E. (2017). Autonomy and enactivism: Towards a theory of sensorimotor autonomous agency. *Topoi, 36*(3), 409–430.

Chang, H. (2020). Pragmatism, perspectivism, and the historicity of science. In M. Massimi & C. D. McCoy (Eds.), *Understanding perspectivism* (pp. 10–27). Taylor & Francis.

Chapman, R. (2021). Neurodiversity and the social ecology of mental functions. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620959833

Contractor, A. A., Roley-Roberts, M. E., Lagdon, S., & Armour, C. (2017). Heterogeneity in patterns of DSM-5 posttraumatic stress disorder and depression symptoms: Latent profile analyses. *Journal of Affective Disorders, 212*, 17–24.

Cooper, R. (2013). Avoiding false positives: Zones of rarity, the threshold problem, and the DSM clinical significance criterion. *The Canadian Journal of Psychiatry, 58*(11), 606–611.

de Haan, S. (2020b). *Enactive psychiatry*. Cambridge University Press.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences, 4*(4), 429–452. https://doi.org/10.1007/s11097-005-9002-y

Di Paolo, E. (2010). Overcoming autopoiesis: An enactive detour on the way from life to society. In *Advanced series in management*. Emerald Group Publishing Limited.

Di Paolo, E., Cuffari, E. C., & De Jaegher, H. (2018). *Linguistic bodies: The continuity between life and language*. MIT Press.

Dickinson, D., Pratt, D. N., Giangrande, E. J., Grunnagle, M., Orel, J., Weinberger, D. R., Callicott, J. H., & Berman, K. F. (2017). Attacking heterogeneity in schizophrenia by deriving clinical subgroups from widely available symptom data. *Schizophrenia Bulletin, 44*(1), 101–113.

Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 ways to have posttraumatic stress disorder. *Perspectives on Psychological Science, 8*(6), 651–662.

Hartner, D. F., & Theurer, K. L. (2018). Psychiatry should not seek mechanisms of disorder. *Journal of Theoretical and Philosophical Psychology, 38*, 189–204.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry, 27*(1), 1–17.

Hawkins-Elder, H., & Ward, T. (2019). Theory construction in the psychopathology domain: A multi-phase approach. *Theory & Psychology*.

Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology, 6*, 155–179.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry, 167*, 748–751.

Jilek, W., & Mattelaer, J. (2006). Koro: The psychological disappearance of the penis. *De Historia Urologiae Europaeae, 13*, 53–73.

Karter, J. M., & Kamens, S. R. (2019). Toward conceptual competence in psychiatric diagnosis: An ecological model for critiques of the DSM. In *Critical Psychiatry* (pp. 17–69). Springer.

Kendler, K. (2012b). The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. *Molecular Psychiatry, 17*(4), 377.

Kendler, K., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*(6), 1143–1150.

Kirmayer, L. J., & Jarvis, G. E. (2019). Culturally responsive services as a path to equity in mental healthcare. *HealthcarePapers, 18*(2), 11–23.

Kirmayer, L. J., & Ramstead, M. J. (2017). Embodiment and enactment in cultural psychiatry. In C. Durt, T. Fuchs, & C. Tewes (Eds.), *Embodiment, enaction, and culture: Investigating the constitution of the shared world* (p. 397). MIT Press.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., & Clark, L. A. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454.

Lewis-Fernández, R., Guarnaccia, P. J., Martínez, I. E., Salmán, E., Schmidt, A., & Liebowitz, M. (2002). Comparative phenomenology of ataques de nervios, panic attacks, and panic disorder. *Culture, Medicine and Psychiatry, 26*(2), 199–223.

Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy, 62*, 129–139.

Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology, 12*, 435–463.

Magnus, P. (2014a). Epistemic categories and causal kinds. *Philosophy Faculty Scholarship*. https://doi.org/10.1016/j.shpsc.2014.10.001

Maiese, M. (2016). *Embodied selves and divided minds*. Oxford University Press. https://books.google.co.nz/books?id=w_quCgAAQBAJ

Maiese, M. (2021). An enactivist reconceptualization of the medical model. *Philosophical Psychology, 34*, 1–27.

Markon, K. E. (2013). Epistemological pluralism and scientific development: An argument against authoritative nosologies. *Journal of Personality Disorders, 27*(5), 554–579.

Monroe, S. M., & Anderson, S. F. (2015). Depression: The shroud of heterogeneity. *Current Directions in Psychological Science, 24*(3), 227–231.

NiaNia, W., Bush, A., & Epston, D. (2016). *Collaborative and indigenous mental health therapy: Tātaihono–stories of Māori healing and psychiatry*. Taylor & Francis.

Nielsen, K. (2020a). Think of mental disorders as the mind's 'sticky tendencies.' *Aeon*. https://aeon.co/ideas/think-of-mental-disorders-as-the-minds-sticky-tendencies

Nielsen, K. (2020b). *What is mental disorder? Developing an embodied, embedded, and enactive psychopathology*. PhD thesis, Victoria University of Wellington. http://hdl.handle.net/10063/8957

Nielsen, K. (2022b). Same diagnosis, different problem: The challenge of heterogeneity in mental disorder. *MIND Foundation*. https://mind-foundation.org/same-diagnosis-different-problem-the-challenge-of-heterogeneity-in-mental-disorder/

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 8*(6), 800–822. https://doi.org/10.1177/0959354318808394

O'Connor, B. (2017). Mental disorder as a practical psychiatric kind. *Philosophy, Psychiatry, & Psychology, 24*(4), E-1-E-13.

Olbert, C. M., Gala, G. J., & Tupler, L. A. (2014). Quantifying heterogeneity attributable to polythetic diagnostic criteria: Theoretical framework and empirical application. *Journal of Abnormal Psychology, 123*(2), 452.

Roberts, T., Krueger, J., & Glackin, S. (2019). Psychiatry beyond the brain: Externalism, mental health, and autistic spectrum disorder. *Philosophy, Psychiatry, & Psychology, 26*(3), E-51.

Stier, M. (2013). Normative preconditions for the assessment of mental disorder. *Frontiers in Psychology, 4*, 611.

Thornton, T. (2000). Mental illness and reductionism: Can functions be naturalized? *Philosophy, Psychiatry, & Psychology, 7*(1), 67–76.

Wakefield, J. C. (2015). DSM-5, psychiatric epidemiology and the false positives problem. *Epidemiology and Psychiatric Sciences, 24*(3), 188–196.

Zachar, P. (2014). *A metaphysics of psychopathology*. MIT Press.

Zachar, P. (2018). Diagnostic nomenclatures in the mental health professions as public policy. *Journal of Humanistic Psychology*. https://doi.org/10.1177/0022167818793002

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164*(4), 557–565.

Zachar, P., & Kendler, K. S. (2017). The philosophy of nosology. *Annual Review of Clinical Psychology, 13*, 49–71.

6

The Task of Explanation (and the Beginnings of Treatment)

3e Psychopathology, as developed across previous chapters, conceptualizes mental disorders as dysfunctional patterns in sense-making. These patterns are seen to be constitutionally complex multi-scale process structures spanning brain, body, and environment. Further, these patterns are deeply entwined within the striving of meaning-driven individuals, turning their very striving against them, and are defined as disorders/dysfunctions on this basis. The intent of the current chapter is to begin extending 3e Psychopathology from a conceptual model of mental disorder alone, to a burgeoning epistemological framework for the study and treatment of psychopathology. This will be achieved by exploring what explanatory strategies consilient with the 3e conceptualization may look like. In other words, the current chapter asks the question 'if this is what mental disorders are, how then should we seek to explain them?'. This question will be explored at both research and clinical levels. First however, it is worth briefly considering what exactly an explanation is and on what basis I am distinguishing between research level and clinical level explanations. Following this, Sect. 6.2 will consider the development of explanatory theory under a 3e

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7_6

Psychopathology perspective. Section 6.3 will then consider the role and development of tailored explanations—i.e., formulations—in clinical practice.

6.1 What Does It Mean to Explain?

Explanation, generally speaking, is a complex and contested epistemological and practical task. In this section I briefly assay some important recent insights about explanation relevant for discussion in later sections. I also justify why I make a distinction between explanation as undertaken by researchers and explanation as undertaken by clinicians. While it is not my intention to address the question 'what does it mean to explain?' in all of the complexity that such a question entails, it is worth briefly expanding on what I mean by explanation in this context.

An explanation, on my view at least, is an explicit or implicit postulation—usually in the form of a set of premises, a model, a theory, a narrative, or a causally informative classification—that *accounts* for the origin or continued existence of a *phenomenon* or set of *phenomena*. This understanding of explanation draws on the philosophical work of Haig (2014) and Thagard (2017). By 'account' I mean 'to make sense of', 'to make less surprising', or 'to offer insight into the workings of'. By 'phenomenon' I mean a regularity or pattern inferred from our data and observations about the state of the world (Bogen & Woodward, 1988). For example, within psychopathology we may consider the phenomena of social anxiety, low mood, or auditory hallucinations. As examples from physics we might consider the phenomena of red-shift in the light coming to earth from stars or the gravitational attraction between objects with mass.

It is also important to note that, in being inferred, phenomena are themselves contested and socially constructed. In other words, deciding how to best delineate and describe a phenomenon so that we might then explain it is an important process in itself (Haig, 2014). For example, the description and construal of gravity—via a complex interplay between available data, technology, theory, and social context—has changed and developed across time from early natural philosophy, through Newton, to Einstein and modern debate (Papaspirou & Moussas, 2013). The things

we seek to explain, across the sciences, are moving and contested targets; we do not seek to explain the world so much as our ever-developing representations of it (Bokulich, 2018). This is important to recognize for our purposes as it again draws attention to the importance of conceptual and theoretical work, including analysis and refinement of the things we are trying to explain, as a vital and normal part of scientific enquiry (Bringmann et al., 2022; Wilshire et al., 2021).

Another important point when considering the task of explanation within psychopathology, is that there is a lot of variation in *how* we can go about explaining human behavior and experience (Bechtel, 2009a; Thagard, 2017). Explanations can take many different forms and be targeted at different scales of the brain-body-environment system. To list multiple examples: We can present a narrative that provides a context and makes it apparent in folk-psychological terms why a behavior is engaged with (Johnstone et al., 2018). We can identify dispositions of character or biology that predispose people to certain patterns of engagement with the world (Vanderbeeken & Weber, 2002). We can map the dynamical relationships between relevant phenomena/symptoms and contextual factors as a network (Borsboom et al., 2018). Or we can delve into the brain and theorize about the causal relationships between neural or genetic correlates of certain behaviors and experiences (Insel & Cuthbert, 2015). There are clearly lots of different ways to explain psychopathology. One of the central aims of this chapter is to consider what guidance 3e Psychopathology has to offer when we are faced with so many options.

A final point that I wish to stress is that explanation is always a pragmatic task, undertaken with particular purposes and with particular intended audiences (Potochnik, 2016, 2017). Such variation in explanatory purpose and context brings with it variation in what makes for a good or bad explanation, and therefore how we should best go about explaining. One vital distinction in terms of our purpose for explaining in psychopathology, is whether we are trying to explain a general across-person pattern of distress (e.g., depression), or whether we are trying to explain one person's particular pattern of difficulty (e.g., one person's depressed experiences and other challenges they may face). The former sort of explanatory enterprise is commonly undertaken in a research context, while the latter is usually undertaken within a clinical context in

order to inform treatment decisions. These different contexts bring with them different access to investigatory tools, different explanatory purposes, and different audiences. As such, how we go about the task of explaining in these different contexts is related, but is, and should be, different.

It is based on this distinction that I have chosen to structure the rest of this chapter, with Sect. 6.2 considering explanation in a research context, and Sect. 6.3 considering explanation in a clinical context. While I will seek to avoid such philosophical language, we can think about this distinction as one of 'token' (the individual case or instance) versus 'kind' (a wider classification to which we can meaningfully say that an individual case of mental disorder 'belongs'). As discussed in the previous chapter though, and as will be further discussed in the current one, 3e Psychopathology views this kindship relation as based on similarity rather than on common cause. On the 3e view, mental disorders are inherently messy and contextual things, resulting in the heterogeneity of individual presentations and often-contestable diagnoses/kindship relations. As will be discussed later, this means that individual clinicians cannot simply apply explanatory theory from research in a wholesale manner when seeking to explain an individual's difficulties. Instead they must craft individualized explanations, drawing on wider explanatory theory in a select and reasoned way. In sum, the explanation of disorders or their component phenomena as typically undertaken by researchers and the clinical explanation of token cases are separable tasks that will require different methods, tools, and styles of explaining. It is for this reason that I have chosen to discuss them separately.

6.2 Explanation for Researchers

In this section I will first outline how 3e Psychopathology necessitates a commitment to explanatory and methodological pluralism. More than any one particular method of enquiry, a diversity of research methods and a principle of inter-methodological respect is required if we are to begin to collectively sketch out the constitutional structures of mental disorders as conceptualized under 3e Psychopathology. Based on an

implicit biological essentialism, decades of disproportional funding channeled towards biologically focused research has left behavioral, experiential, and ecologically focused methodologies underfunded and disempowered. I will therefore argue that greater emphasis must be given to such methodologies. However, I will also argue that there is a clear need for research focused on smaller scales of enquiry such as genetics and neurobiology to continue. Whether such research counts as 'reductionistic' or not does not depend on the methods themselves, but on how we understand and treat the findings. This will lead to a brief discussion on the vital integrative role of theoretical researchers within the sciences of psychopathology under the 3e framework. Following this I will then briefly review some current approaches to targeting explanations and organizing research findings in psychopathology, highlighting where these approaches falter under a 3e Psychopathology perspective. I will then outline the Relational Analysis of Phenomena [the RAP] as one plausible approach to the development of explanatory theory that accords well with the 3e conceptualization of mental disorder (Nielsen & Ward, 2020).

Gradualism, Explanatory Pluralism, and Methodological Pluralism

As discussed across previous chapters, 3e Psychopathology makes a variety of theoretical and conceptual commitments. Some key points of difference between the 3e concept developed and *status quo* approaches, for example, are the open commitments to moderate externalism, anti-essentialism, and a contextual/relational understanding of dysfunction. In other words, under 3e Psychopathology, mental disorder pertains to the functional status of the relationship between the sense-making processes of the organism and its context. Further, the causal structures that support continued engagement with dysfunctional sense-making are seen to span brain, body, and environment. Mental disorders are therefore not merely 'in people' but *between* people and the world they are embedded in, as well as spread across it. The ramifications of such commitments for the task of explanation are hard to overstate because, compared to an

often assumed biologically or psychologically essentialist view, the very nature of what we are seeking to explain is changed. Under 3e Psychopathology there is an anchoring of the explanatory target to the scale of the individual making sense of and engaging with their environment, and the network of causal factors maintaining the dysfunctional behavior is presumed to be dispersed and complex. Moreover, the relationship between individual instances of a disorder and a wider diagnostic category to which it can be said to belong is presumed to be variable—i.e., the kindship of instances of mental disorder to a diagnostic category is defined by *similarity of structure* rather than *sameness of cause* or *similarity of symptoms*. Under such a view, heterogeneity is no longer necessarily a problematic artefact, but should perhaps be considered a feature of the subject matter that needs to be accounted for (Nielsen, 2022b). These various commitments, and others inherent to 3e Psychopathology, dramatically change what the task of explanation will look like for researchers.

For a start, on this view we aren't just looking for one 'nugget of truth' which will explain a mental disorder. Rather than a moment of discovery like the well-known myth of Newton and his apple, we would expect a more gradual process of knowledge gathering and explanatory progress. This view can be referred to as *gradualism*. Under the view provided by 3e Psychopathology the task of explanation will represent a gradual and dispersed process where researchers from across the globe slowly work to reveal the network of mechanisms that constitute the causal structure of a mental disorder. Instead of one paradigm defining discovery then, coming to understand a mental disorder will probably be much more like a family gradually assembling a tabletop puzzle—arguing about which piece goes where and what approaches to take, but ultimately collaborating and working on different areas to slowly reveal the image.

Where this puzzle metaphor potentially falls down, however, is that instead of developing a single theory—e.g., the X theory of depression—we will likely need multiple explanations that each focus on different mechanisms in the wider disorder structure and how they operate. Rather than somebody developing a successful explanation of depression as a whole, we would instead expect smaller scale explanations to be developed, mechanism by mechanism. As hypothetical examples, we might

see theories emerging at a neurological-level that concern how difficulties experiencing pleasure relate to difficulties sleeping, or at a psychological/ecological-level regarding how the changes that depressed people make to their environments may actually contribute to the perpetuation of their low mood.[1] Moreover, not all mechanisms recognized in such a collection of explanatory theories may be relevant for all individual cases of a disorder. Due to the fuzzy nature of the kind concept, certain mechanisms may be playing a greater role in individual instances of disorder than in others.

3e Psychopathology demands a comprehensive multi-scale and constitutionally minded view, consisting of brain, body, and environment. This aligns with the empirical evidence that the causal factors at play in mental disorders are 'dappled' across the various scales of enquiry (Kendler, 2012b, 2019). In the face of this complexity it seems nigh on impossible that any one approach to studying mental disorder will be sufficient to warrant abandoning all other methods. Something similar can be said for any one explanatory model of a particular disorder. The likelihood that a single superior explanatory model of any given disorder will emerge that accounts for all useful and interesting facets of that disorder, to the point that it warrants ignoring all other points of view, seems incredibly low. The complex and emergent nature of mental disorder recognized by 3e Psychopathology therefore commits it, not just to the classificatory pluralism discussed in the previous chapter, but also to methodological and explanatory pluralism as argued in Chap. 4. Put simply this refers to the respective ideas that there are—and likely always will be—many ways to both study and model mental disorders, and that we should seek to explore these many ways rather than unite behind a single approach (for parallel arguments see: Clack & Ward, 2020; Hawkins-Elder & Ward, 2021; Jerotic & Aftab, 2021). Such pluralistic ideas are well discussed in related areas and wider philosophy of science (Brigandt, 2013; Chang, 2017; Mitchell, 2002; Sullivan, 2017; Veit, 2020; Wegerhoff, 2022; Wegerhoff et al., 2020, 2022). The conceptual view afforded by 3e Psychopathology then, strongly suggests that the epistemological health of the wider sciences of psychopathology will be enriched when a

[1] This last example is inspired by Krueger and Colombetti (2018).

diversity of research methods is utilized and genuine inter-methodological respect is normalized.

Efforts such as 'the decade of the brain' (Jones & Mendell, 1999), biologically essentialist public communications by major bodies such as the American National Institute of Mental Health ([NIMH]; Insel et al., 2010), pharmaceutical advertisements endorsing chemical understandings of mental disorder (Leo & Lacasse, 2008), and the general excitement over the successes of neuroscience and its amazing technologies, have all contributed to a social landscape that facilitates the funding of brain-focused research. Mental disorder however, is a *psychological* phenomenon (Miller, 2010). Under the 3e view, such psychological phenomena are body involving, but are also world and history involving. To equate mental disorder with brain disorder from this view is therefore a gross de-contextualization, as discussed in previous chapters. If we are to truly come to understand these complex and heterogeneous structures we call mental disorders, then there is a clear need to study them from multiple angles—i.e., at the scales of brain, body, development, and environment (and any other that can add to our understanding). As discussed in Chap. 3, authors such as Fuchs (2017) have argued well for "…a polyperspectival approach" (2017, p. 276) under an enactive understanding of psychopathology, albeit with particular focus on the importance of lived-experience and phenomenological analysis. The 3e Psychopathology framework developed here accords with such arguments—experience focused methodologies have much to offer our understanding of psychopathology. More than any one kind of research methodology however, 3e Psychopathology endorses pluralism. As I have argued across this book, an embodied, embedded, and enactive view of human functioning can help us makes sense of and conceptualize mental disorders. However, this in no way means that only research methods grounded in such a view of human functioning have anything useful to offer for our understanding of mental disorders. Indeed, if mental disorders are as complex as the current analysis suggests, then we cannot afford to be picky. We should be open to any perspective that has some useful insight to offer. This includes brain-focused methodologies as much as it does experience-focused ones.

A tension may be seen to arise at this point between the perceived holism of the enactive view and the perceived 'reductionism' of

brain-focused research methods. It is not my intention to address this issue in detail but, in short, this concern appears grounded in an overly simplified understanding of what reductionism actually is and how big an epistemological threat it entails. An important distinction is well recognized between *explanatory reductionism* and *theory reductionism* (Brigandt, 2013). Explanatory reductionism refers to the targeted decontextualization and decomposition of an object of study, in order to develop an understanding of how its behavior may in part be caused by elements of its constitution at smaller scales of enquiry. Theory reductionism meanwhile is the simplistic and outdated idea that theories themselves will eventually be able to be completely accounted for by theoretical languages at more 'fundamental' levels, i.e., that psychology will be subsumed by biology, biology by chemistry, chemistry by physics, and physics by mathematics—an idea clearly at odds with enactive thinking. I wish to hazard an educated guess however, that most neuroscientists are not theory reductionists (or at least the good ones). It is common to hear neuroscientists speak of context and complexity, and I suspect many would agree that the totality of human behavior cannot be understood by study of the brain alone—i.e., that human behavior is in a sense 'irreducible'. On my understanding, neuroscientists zoom in on and decontextualize the brain because: (1) the brain is clearly very important to the study of human functioning, (2) doing so helps manage the sheer complexity at hand, and (3) reductionism has clearly been one of the most successful scientific strategies of the last few hundred years! The best neuroscientists however, then attempt to re-contextualize their findings when incorporating it back into their wider theories. Enactivism meanwhile, does not deny the existence of objects and processes at smaller scales of enquiry, or their relevance to human behavior. It simply demands that such objects and processes be understood within the context of the wider dynamic and striving person-in-context (Gauld et al., 2022).

As argued in Chap. 4, enactivism is best seen as a worldview, and as such there is room for a plurality of different methods within it (Donovan & Murphy, 2020). As such, I see no reason why researchers grounded in a wider enactive epistemological framework cannot utilize methods of explanatory reductionism, so long as they remember to put their findings back in context when they are done. I would further argue that this in

turn points to the central role of theoretical researchers within a 3e informed science of psychopathology. As briefly discussed in Chap. 1, such researchers are well placed to integrate and re-contextualize knowledge into manageable and useful theoretical frameworks. This brings us to the question of how theoreticians and other researchers should seek to organize research findings from diverse methodologies in order to develop more integrated explanations/understandings. One possible approach to this challenge is the Relational Analysis of Phenomena, or 'The RAP' (Nielsen & Ward, 2020). I will now review some current approaches to organizing research findings and targeting our attempts at explanation in psychopathology before presenting the RAP as one 3e congruent way to approach the development of explanatory theory.

DSM-ICD, RDoC, and Symptom-Based Approaches to Explanation

There are many differing perspectives on how we should organize and coordinate research findings within the study of mental disorder in order to best facilitate our developing understanding (Sullivan, 2017). One way to think about this is to ask the question 'what are the most appropriate targets of explanation in psychopathology?' (Nielsen & Ward, 2020). The seemingly obvious answer to this question is that we should seek to explain the various mental disorders recognized, i.e., the diagnostic concepts found within major classification systems such as the DSM or ICD (Berenbaum, 2013). There are however significant problems with this answer. These issues relate to wider criticisms of these classification systems as extensively reviewed elsewhere (Karter & Kamens, 2019; Lilienfeld & Treadway, 2016; Zachar & Kendler, 2017). Primary among the difficulties with this approach to defining the targets of explanation in psychopathology is that current diagnostic concepts are very *heterogeneous* (Lilienfeld, 2014). In other words, different instances of the same diagnosed mental disorder are often very different both in appearance and in apparent causes. Such concerns regarding heterogeneity are well evidenced with a diverse range of prototypical mental disorders having been shown to capture large and heterogeneous populations, including

post-traumatic stress disorder (PTSD), eating disorders, schizophrenia, and depression (Contractor et al., 2017; Dickinson et al., 2017; Galatzer-Levy & Bryant, 2013; Hawkins-Elder & Ward, 2020; Monroe & Anderson, 2015). This problem of heterogeneity, alongside other challenges, suggests that our current diagnostic constructs fail to pick out similarly constituted entities with common causal processes/structures and are thereby of questionable etiopathological validity. In short they do not seem like the sort of stable phenomena that make for good explanatory targets (Nielsen & Ward, 2020).

Recognition of such challenges, and a perceived lack of progress in the development of causal explanations for mental disorders, has led to the development of alternative approaches. Such approaches are often labelled as *trans-diagnostic* in that they step away from or seek to cut across extant diagnostic categories. Primary among these approaches is the Research Domain Criteria ([RDoC] Cuthbert, 2014; Cuthbert & Insel, 2013; Cuthbert & Kozak, 2013; Insel et al., 2010). Developed by the National Institute of Mental Health [NIMH] in America, RDoC is a funding initiative designed to shift research attention away from the signs and symptoms of mental disorder as per the DSM-ICD, to the underlying causal processes that generate them. Vitally however, in doing so it assumes mental disorders to be disorders of 'brain circuitry' thus taking a biologically essentialist position (Insel et al., 2010; Morris & Cuthbert, 2012).

RDoC is centered around an organizational matrix with a horizontal axis containing seven 'units of analysis' (which specify structural 'levels' of enquiry), and a vertical axis listing basic psychological functions (Cuthbert & Insel, 2013; Cuthbert & Kozak, 2013; Lilienfeld & Treadway, 2016; Morris & Cuthbert, 2012). The seven units of analysis are biologically/neurologically focused including: genes, molecules, cells, neural circuits, physiology, behavior, and self-report. Researchers seeking funding are able to classify their research within this matrix rather than in reference to DSM-ICD diagnostic concepts. The explanatory aim of RDoC then, is to facilitate the study of how phenomena observed at the seven defined levels (e.g., higher levels of striatal dopamine, lower dendritic spine density in brain area X) affect the degree to which the basic functions are achieved (e.g., response to acute threat, approach motivation). The hope

is that this process will uncover *trans-diagnostic mechanisms* relevant to current diagnostic labels (Cuthbert & Insel, 2013; Hoffman & Zachar, 2017). It is important to note however that these trans-diagnostic mechanisms are assumed to be neurobiological in origin, with conceptual priority given to the central unit of analysis within the matrix; 'neural circuitry'. With time, RDoC messaging has evolved to give greater focus to the role of development and the environment, including cultural context and social determinants of mental health. However, the conceptual focus on the nervous system and neural circuitry clearly remains. For example, in a section concerning development and environmental factors the RDoC website states:

> "It is now widely accepted that most mental illnesses result from maladaptive maturation of the nervous system including its interaction with the wide variety of external influences beginning at conception. The social and physical environment comprises sources of both risk and protection for many different disorders occurring at all points along the life span, and methods for studying phenomena such as gene expression, neural plasticity, and various types of learning are rapidly advancing." (NIMH, 2022)

Concerns about the neurocentricism of RDoC have been well discussed elsewhere however, and it is therefore perhaps more useful to focus on what RDoC appears to do well (Berenbaum, 2013; Hershenberg & Goldfried, 2015; Kirmayer & Crafa, 2014; Lilienfeld, 2014; Lilienfeld & Treadway, 2016; Nielsen & Ward, 2018, 2020; Wakefield, 2014). By shifting explanatory focus to the discovery of transdiagnostic neural mechanisms and their relation to specified functions, RDoC represents a shrinking of explanatory targets towards more stable, less heterogeneous, phenomena. This move seems an advisable response to the heterogeneity plaguing DSM-ICD-defined targets. Regarding this move however, authors Hoffman and Zachar (2017) point out an important concern:

> "[t]he worry is that in order to achieve the fineness of grain needed for elucidation of causal mechanisms, we risk losing connection to the "coarse" clinical phenomena of interest." (2017, p. 68)

A vital quality for a target of explanation to hold is that it should maintain its relevance to the reason for seeking an explanation in the first place. The study of psychopathology is fundamentally a pragmatic science—we want to understand mental disorders to better understand how we can help people. By shrinking the explanatory targets of research down to the scale of neural circuitry and its impact on basic cognitive functioning there is a worry that RDoC drifts too far from this purpose. Such an argument would hold that more immedicably useful clinical and therapeutic knowledge is likely to be developed if explanatory centrality was granted to people-in-context rather than to the neural pathways in their heads. A related concern is that the concept of mental disorder is inherently normative, yet outside the specified basic functional domains there is no broader normative element within RDoC with which to give RDoC's findings meaningful conceptual validity (Nielsen & Ward, 2018; Wakefield, 2014b).

In essence then, there is a reasonable argument to be made that RDoC represents an overcorrection in the grain size of the explanatory targets in psychopathology. By this I mean that 'neural circuits' do not maintain sufficient relevance to the wider dysfunction and suffering that motivates our enquiries in the first place. In many ways this concern is probably outweighed by the sheer amount of basic research that RDoC is likely to facilitate, however, we need to be clear about what RDoC is doing. Under the conceptual framework developed within these pages, research within the RDoC framework searches for (largely sub-personal) abnormalities that likely play constitutional and/or causal roles as *components of* psychopathology. Neural circuit abnormalities are not themselves disorders under the current conception. The efforts represented by RDoC thereby represent vital work, discovering and confirming constituent phenomena. Such phenomena can then be utilized, in combination with wider findings, to weave together explanations that more thoroughly and fully sketch out the complex process structures of mental disorders. On the current view such constituent phenomena do not themselves constitute explanations of psychopathology.

RDoC grants greater freedom to researchers, in that under the RDoC framework they no longer have to justify their research interests by linking them to some particular and established problem (i.e., DSM-ICD

syndromes). This freedom will be good for psychopathology as a complete scientific endeavor (Casey et al., 2013), but the vital task of *developing explanations* of psychopathology has different requirements to the larger science within which it sits. On the current view, ideal targets of explanation balance stability and relevance to the larger disorder space (Nielsen & Ward, 2020). By targeting largely sub-personal abnormities and investigating their potential role as transdiagnostic mechanisms, RDoC seems to prioritize the prior at the expense of the latter. In doing so, RDoC seems be performing a different task to that of picking out ideal targets and explaining them. Rather, it seems to be attempting to provide some important sub-personal ingredients for our explanations.

Another emerging approach to explanation in psychopathology is to turn explanatory focus to symptoms or individual phenomena themselves (Wilshire et al., 2021). While less commonly discussed than RDoC or DSM-ICD-based approaches, such a focus is worth briefly exploring. Instead of trying to explain the wider concept of depression as per DSM-ICD-based approaches, or seeking to understand the potential neurobiological ingredients that may underpin various disorders as per RDoC, symptom-based approaches focus in on individual symptoms or phenomena (T. Ward & Clack, 2019a). For example Clack and Ward (2020) example how a multiscale understanding of an important phenomenon within depression—anhedonia—can contribute greatly to our understanding of how depression works. Part of the motivation behind these approaches, similar to the RDoC, is to improve the stability/homogeneity of the targets of our explanations. Another key advantage of this approach is that it highlights that symptoms themselves are complex constructs, partially socially constructed, and thereby themselves in need of critical theoretical analysis (Wilshire et al., 2021).

In focusing on single symptoms or 'clinical phenomena' such as anhedonia however, symptom-based approaches take a large step away from our currently recognized syndromes of mental disorder. Much like RDoC then, in order to achieve greater stability/homogeneity in their explanatory target, such approaches represent a shrinking of explanatory targets away from currently recognized patterns of distress and dysfunction. Such a step is well reasoned and seems likely to be very fruitful. Largely and for the most part however, such approaches represent a method of

decomposition; of breaking the problem down into parts. Typically, methods of decomposition are followed by reassembly; putting the pieces back together with an understanding of how they work (Bechtel, 2009b; Kendler, 2008). What is strung back together in symptom-based approaches is an understanding of an individual component phenomenon—i.e., one important part of a larger disorder—not necessarily an understanding of the wider pattern of distress or dysfunction.

A concern I have with symptom-based approaches is whether or not they will capture well the self-perpetuating structures that result in genuine disorder as highlighted under 3e Psychopathology. For example, experiencing intrusive thoughts is an important part of OCD (Seli et al., 2017), yet having intrusive thoughts is extremely normal across many cultures and by itself is rarely problematic (Radomsky et al., 2014). Rather, this phenomenon of intrusive thoughts needs to be viewed in interaction with other phenomena such as compulsions, distress, and the moderating effect of traits such as thought-action fusion, in order to understand why this collection of phenomena constitute a 'disorder' for the person affected. As another example, there are many instances of the 'symptom' of experiencing recurring hallucinations—often seen as quintessentially pathological—actually having no clear association with dysfunction or significant harm (Fulford & Jackson, 1997; Larøi et al., 2014; NiaNia et al., 2016). Similarly then, hallucinations need to be viewed in relation to other phenomena such as confusion, distress, and cultural-moral conflict, before we can see how they should be considered disordered. This suggests that targeting symptoms themselves, while to the benefit of target *stability* and the thoroughness with which we can analyze said target, such a move still somewhat compromises on the *relevance* of the explanatory target to our wider purposes. This concern is a large part of what inspired the Relational Analysis of Phenomena approach outlined in the following section. For more detailed review of different approaches to explanation, including review of symptom network based approaches, see (Nielsen, 2020b; Nielsen & Ward, 2020).

The RAP

The Relational Analysis of Phenomena [the RAP] is a meta-methodological framework for integrating research findings and developing theories for the explanation of mental disorders. Based upon the 3e Psychopathology conceptualization of mental disorder, the RAP is designed to capture hypothesized circular/recursive causal process structures distributed across brain, body, and environment. In particular, it is designed to produce theories of how disorders maintain as opposed to how they originate, producing theoretical sketches of the dynamical constitution of mental disorders across the brain-body-environment system—i.e., sketches of how unhelpful differences in sense-making tend to maintain themselves over time. The RAP is *not* designed to produce etiological explanations in the sense of linking patterns in sense-making to distal causes such as genetics and developmental factors. It is also *not* designed to produce individually tailored explanations for mental disorders as will be discussed in the second half of this chapter. In this section I will briefly present some of the core ideas within the RAP, before briefly overviewing its various phases. For fuller discussion of the RAP I would direct the reader to Nielsen and Ward (2020) and Nielsen (2020b).

Instead of broad heterogeneous clusters of symptoms such as those featured in the DSM-ICD, or neural circuit abnormalities as focused on within RDoC, the RAP gives explanatory focus to the *relationships between component phenomena* within a given disorder space. By 'component phenomena' I am referring to the reliable tendencies and occurrences that typically occur within a mental disorder, examples might include anhedonia, sleep latency difficulties, or grandiose delusions. This is a similar concept then to 'symptoms', but does not carry assumptions of disease or strict homogeneity for the wider disorder concept, and highlights that phenomena/symptoms are themselves detected/constructed rather than being pre-given and therefore need to be argued for and defended as valid and useful patterns in the world to utilize within our explanations (T. Ward & Clack, 2019b). By 'disorder space' I am referring to the general conceptual space denoted by a mental disorder that one might want to come to understand, e.g., depression, OCD, PTSD. Use of this term highlights, as discussed in earlier chapters, that

the division between mental disorders themselves and therefore their classification is contestable, and that there are likely many reasonable ways to approach this. The central idea of the RAP is that if we can repeatedly come to understand the causal relationships between enough of the component phenomena within a given disorder space, then we will, with time and coordinated effort, develop a sketch of how that disorder works—how it biases the sense-making of individuals and how it maintains itself over time (Nielsen & Ward, 2020). As discussed in the previous section the underlying understanding of scientific progress here is one that is gradual and distributed across different researchers/labs, rather than a view where progress comes in leaps of genius discovery.

Relationships between component phenomena are targeted within the RAP by use of the concept of 'phenomena complexes'. Phenomena complexes are artificially selected units of two to three component phenomena that reliably co-occur within a given disorder space. For example, a phenomena complex within anxiety disorders might include increased irritability, increased sleep latency, and frequent anxious rumination. By 'artificially selected' I mean that phenomena complexes are not meant to represent naturally occurring or isolatable parts of a disorder. It is proposed that, as epistemic units, phenomena complexes represent a balance point between the stability and relevance of explanatory targets within psychopathology (Nielsen & Ward, 2020). This can be seen most clearly in relation to clinical practice. Detailed comprehension of the neurological underpinnings or wider constitutional structure of individual symptoms, such as anhedonia, seems of limited practical value to mental health professionals given currently available clinical tools—such targets are therefore likely stable but not very relevant. As discussed above meanwhile, the diagnostic concepts of the DSM-ICD tell a clinician very little about what is actually causing someone's problems because, while they describe genuine problems, they are overly heterogeneous—i.e., they are relevant but insufficiently stable. Phenomena complexes meanwhile, in capturing the likely relationships *between* common clinical phenomena within a given disorder space, would seem to represent useful epistemic units through which clinicians could draw on the evidence base in a way that was directly relevant to helping their patients. In being composed of only a small number of phenomena, such complexes also seem more

likely to be stable than current DSM-ICD constructs. Imagine briefly three different databases, one listing DSM-ICD diagnostic concepts and linking these to proposed explanatory theories, one listing individual symptoms or basic functional disturbances as per RDoC and linking these to proposed neural mechanisms, and one listing phenomena complexes or allowing you to select multiple phenomena and linking these to hypothesized and established 'micro-theories' about how the clinical phenomena may be influencing and supporting each other. I would argue that the third database would offer the clearest clinical utility. For further discussion of issues of heterogeneity in connection to the RAP, see Nielsen (2022b).

Now that I have outlined the general approach and intention of the RAP—i.e., that it is a gradualist, multi-scale, and collaborative approach to explaining the maintenance/constitution of mental disorders via the targeting of the relationships between component phenomena—I can now shift to outlining the actual phases of the RAP as a method. The phases of the RAP are as follows. Phase 1: Explainers first list out the phenomena that reliably occur within a disorder space and sketch the apparent relationships between them. This is conceptually parallel to the development of a symptom network for a given disorder. Phase 2: Explainers then select out a small number of seemingly related phenomena to form a phenomena complex, and describe these phenomena at multiple scales of enquiry—from genetics to culture—using multiple descriptions drawn from the literature or guided by their own investigations. Phase 3: Finally, explainers utilize their now rich understanding of this small selection of phenomena to abductively infer possible causal relationships within the complex and evaluate these explanations both experimentally and in accordance with accepted epistemic values. An important feature of the RAP is that it is seen as an iterative process, with researchers free to return and cycle back through the various phases on the basis of their evaluations. The central goal of the RAP is to support the gradual and dispersed development of scientific understanding regarding the complex multiscale process structures conceptualized to be at play in the maintenance of mental disorders (Nielsen & Ward, 2020). The three phases of the RAP are represented in Fig. 6.1.

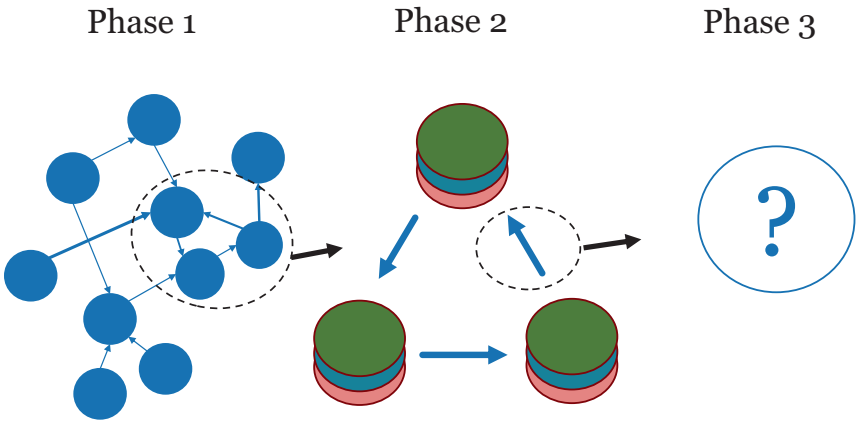

Fig. 6.1 The Stages of the RAP. Phase 1 involves sketching out and refining a map of constituent phenomena and their relationships within the disorder space. Phase 2 involves selecting a phenomena complex and describing its constituent phenomena across multiple scales. Phase 3 involves abductively inferring a mechanism that explains relationships within the phenomena complex. This structure is iterative in that cycling back through the phases should gradually fill out our understanding of the dynamic constitution of the disorder under study

Summarizing 3e Psychopathology and Nomothetic Explanation

In summary, 3e Psychopathology calls for a pluralistic approach to research level explanation and a general principle of inter-methodological respect. Given the complexity highlighted by the 3e Psychopathology conceptualization of mental disorder developed across previous chapters it is held that there are many useful ways to study mental disorders, to organize our findings, and to explain them. In accordance with common criticisms, DSM-ICD diagnostic constructs do not do justice to the complexity and individual variation of mental disorders. RDoC represents one plausible way to organize research efforts within psychopathology, but concerns were highlighted regarding its neuro-centricity and lack of normative consideration. The RAP represents one plausible and 3e

Psychopathology congruent way to organize research findings and develop explanations of psychopathology. It is important to note however that the RAP is designed to develop an understanding of the maintenance of mental disorders, not their origin or development.

6.3 Explanation for Clinicians—i.e., Formulation

If someone is facing mental health related challenges in their life and decides to go and see a clinical psychologist, psychiatrist, or other mental health professional, the first meeting with this clinician will likely involve describing the nature of the challenges being faced. While this clinician may or may not attempt to categorize or 'diagnose' these challenges as a recognized pattern of difficulty, this clinician—whether they realize it or not—will always come to understand and explain the challenges in a particular way. They will make inferences based on the information provided, explanatory theory, and their own training and conceptual assumptions, as to how they understand what is going on. In other words, they will begin to *formulate* an understanding of the difficulties being faced. This act of formulation—and theorizing about it—is given particular attention within clinical psychology as the entire discipline is based upon the idea of assessing, understanding, and helping people utilizing knowledge and ways of understanding grounded in psychological science rather than in medicine and diagnosis (Johnstone, 2018). Formulation is, however, explicitly discussed and utilized within other disciplines. Formulation is defined and understood in many different ways, but it is generally agreed that a formulation can be thought of as a hypothesized explanation/understanding of a patient's difficulties which can then be used to guide clinical decisions and treatment (Bruch, 2015; Eells, 2015; Johnstone, 2018; Johnstone & Dallos, 2013; MacNeil et al., 2012; T. Ward et al., 2016). It is also important to recognize that formulations serve a diverse range of functions, for example they are also often used in report writing and to argue for patient access to further supports within our healthcare systems.

Across the rest of this chapter I will explore some implications of 3e Psychopathology for the act of formulation—i.e., how a 3e informed clinician may go about explaining a token case of mental disorder. My wider intention here is *not* to present a methodology for formulation. Instead, my intention is to begin exploring how 3e Psychopathology can have genuine impact on day-to-day clinical practices through the constraints and guidance it offers for the task of formulation. In other words, my aim here is to explore implications for the tasks of both explanation and, to a lesser extent, treatment. In order to do so I must first give a sense of how approaches to formulation vary. For this reason a rough methodological taxonomy of formulation practices will first be presented. Once this taxonomy has offered an overview of some key ways that understandings and approaches to formulation differ, I will then shift to discussing the general function of formulation under a 3e conceptualization. A methodologically pluralistic understanding of formulation will be endorsed, alongside discussion of how a 3e Psychopathology orientation can provide constraint and guidance for clinicians. Toward the end of this discussion two tools will be presented for use by clinicians: a summary list of principles and values for the evaluation of formulation processes, and a visual schematic for guiding the exploration of moment-to-moment sense-making during assessment and treatment. My primary hope with these tools is that they may be of benefit for fellow clinicians interested in exploring how 3e Psychopathology can supplement their practice. However, these tools may also be of interest to philosophers of psychiatry both professional and amateur who have an interest in formulation.

As an aside, throughout this section I will variously use the words 'person', 'client', and 'patient' to refer to a person seeking help from a clinician. This is often a source of surprising controversy, with much debate made over whether 'patient' is too medical and grants too much power to the clinician, or whether 'client' is too business-like and therefore inappropriately transactional for use within care-based professions. I use these various terms intentionally, in defiance of such debates. Words are obviously important, but in this case the particular words are not as important as what we mean by them and how clinicians actually relate to the people they see.

A Rough Methodological Taxonomy of Formulation Practices

There are many different ways to think about formulation, including in terms of what exactly it is, what it aims to do, and how best to do it. In this section I will present a rough methodological taxonomy of formulation, in the style of Zachar and Kendler's (2007) conceptual taxonomy of mental disorder utilized in earlier chapters. This taxonomy is intended to capture some of the important ways that understandings of formulation differ, and is not intended to be comprehensive (for further discussion of formulation see: Eells, 2015; Johnstone, 2018; Johnstone & Dallos, 2013; T. Ward et al., 2016). Rather than remaining entirely impartial in presenting this taxonomy, I will at a few points comment when a particular approach has a clear epistemological or pragmatic advantage. While my intention here is to demonstrate how approaches to and understandings of formulation differ, it would be dishonest to pretend that all approaches are made equal. Particular attention will at times be given to the Abductive Theory of Method (ATOM) developed by Ward et al. (1999) and Ward et al. (2016). The first reason for this is that ATOM is the approach to formulation in which I was trained and draw on in my own practice, and I feel it is important to recognize how it has shaped my thinking. The second reason is that, as I will seek to demonstrate, this approach to formulation features several epistemic advantages and accords reasonably well with a 3e informed approach. I will now overview the 8 factors of this taxonomy in turn. They are: theoretical loyalty, tailoring, degree of collaboration, selection of the explanans, the understanding of truth, the explanatory purpose, the explanatory target, and the explanatory style.

Theoretical Loyalty. This factor refers to whether formulation is understood to be engaged in within the confines of a particular theoretical orientation/treatment modality, or whether formulation is understood as a process more or less independent of such theoretical commitments. For example, someone who understands formulation in the former way and who is committed to an Acceptance and Commitment Therapy (ACT) approach would attempt to understand a patients difficulties utilizing the various concepts/tools that ACT contains—i.e., in

reference to how fused they are with their thoughts, how attached they are to self-concepts, how connected they are with the present moment, how accepting they are of their present experience, how aware of their values they are, and whether they are acting in accordance with their values (Harris, 2009). Someone who understands formulation as a process independent of a particular theoretical orientation, meanwhile, is offered less clear guidance as to the theoretical objects and tools they can draw on. This brings greater freedom to pick and choose theoretical tools as their needs require, but also greater cognitive demand on the clinician to somehow tie together insights from diverse theoretical perspectives in a coherent way. This factor can be considered as somewhat of a continuum from someone with a strict understanding of formulation existing within a theoretical framework (e.g., the devout ACT practitioner above), to someone who draws on or integrates multiple theoretical perspectives as they feel is useful (e.g., someone who integrates or otherwise combines ACT with ideas from somatic therapy and the psychodynamic tradition), through to someone who understands formulation as a common explanatory process across different theoretical orientations—i.e., a trans-theoretical approach. Integrative but non-trans-theoretical approaches are often esoteric in nature, however some more formalized integrative approaches do exist (Dallos, 2006). The latter trans-theoretical approach is endorsed by Eells (2015) who distinguishes between the process of formulating, and the content of formulation, presenting an integrative view whereby the process is similar no matter the theoretical content one draws on. Such an approach is also endorsed within ATOM, under which—no matter one's particular theoretical commitments—the act of formulation is understood as an abductive explanatory process analogous to the methods of science—i.e., as an inference to the best explanation (T. Ward et al., 1999, 2016).

Tailoring. This factor relates to how novel or personalized formulations are seen to be and can again be thought of as existing on a continuum. At the one end we may have professionals who take an almost *diagnostic* or *off the shelf* approach, identifying patterns or syndromes and importing relevant models or theories wholesale to inform their understanding. Somewhere in the middle would sit *template-based* models of formulation whereby a set structure is utilized but individually tailored to

the patient based on the findings of assessment. An example of this would be a classic Beckian CBT formulation utilizing concepts of core beliefs, assumptions, and the flow-on effect of these through someone's thoughts, emotions, actions and bodily sensations (J. S. Beck, 2020). Other examples of this template-based approach would include the 5-P model (a common approach to formulation involving describing the problem and listing out apparent predisposing, precipitating, perpetuating, and protective factors), or the various triangle models utilized in some psychodynamic approaches (Malan & Parker, 1995). A flexible use of various theoretical templates for formulation is endorsed and further discussed by Eells (2015). At the other end of the continuum we see completely *bespoke* explanations with no predetermined theoretical structure. An example of this would again be the ATOM approach which describes the inferential processes of developing a good formulation rather than prescribing any particular way of structuring ones understanding of the problem itself, other than that it should highlight the core 'working parts' or mechanisms so as to best target attempts at intervention (T. Ward et al., 1999, 2016).

Degree of Collaboration. This factor refers to the question 'is a formulation something that is developed by the professional as an expert, or is it something that the professional and client develop together in a collaborative fashion?' Answers to this question can be understood as existing on a continuum. On the extreme *collaborative* end we can arguably consider the person-centered therapy of Carl Rogers which aims forgo any sense of the therapist as expert, and instead seek to facilitate the persons own growth and potential through human connection and collaborative observations. On the opposing *didactic* end, we may consider a professional who views a formulation as something they themselves complete, on the basis of their assessment of the client, which they may then choose to disclose to the patient. Somewhere in the middle we may consider a moderately collaborative approach whereby both parties are seen to have relevant knowledge and skills, and come together to produce a shared understanding of the problem at hand in order to help them decide what to do about it, such as the approach described by Johnstone (2018).

Selection of the Explanans. This factor is similar to the notion of internalism vs externalism within Zachar and Kendler's (2007) conceptual taxonomy discussed in earlier chapters. In philosophy 'the explanans' refers to the collection of premises used to explain something. In other words, this factor concerns what sort of causal factors are incorporated into a formulation. In particular this factor refers to whether formulations simply focus on psychological factors in a restricted sense (i.e., how the client is understanding and responding to the world) or whether they incorporate—or even focus on—causal influences and contextual factors from across the wider system (e.g., biological, environmental, and social factors). An example of a psychologically focused approach would be the formulaic CBT-style formulations mentioned earlier. While in practice contextual factors will often be acknowledged by CBT practitioners, there is a clear focus on cognition and behavior. This is useful in many ways as it is the client's ways of thinking and behaving that they have the most control over, and is somethings that the practitioner and patient have access to in the room. However, consider for example someone experiencing mood disruption and fatigue difficulties secondary to hypothyroidism, or as a result of bullying or racism. It is difficult to see how to acknowledge the role of such systemic factors within traditional CBT formulations. On the other end of the spectrum, we may consider systemic approaches which give focus to interpersonal dynamics in a family unit, or practitioners that attempt to craft understandings that integrate causal factors from across the biopsychosocial spectrum.

The Nature of 'Truth'. This factor relates to the questions 'how true is a formulation?' and 'in what sense is a formulation true?' This factor is not modelled well as a continuum, but rather we can see multiple distinct positions available. Full discussion on the nature of truth is clearly a topic well beyond the scope of the current project, however, suffice to say that there a several established positions. An *empiricist* position for example might hold that a formulation is 'true' insofar as it accords with the evidence and makes accurate predictions about the client's behavior—i.e., it appears to align with 'objective reality'. A *social constructionist* perspective meanwhile might hold that there is no single reality/truth for a formulation to align with and therefore might focus on understanding the various perspectives/narratives that the client and others hold (Harper & Spellman,

2006). On the other hand, a *pragmatic* perspective may circumvent such metaphysical quandaries and hold a formulation to be 'true' insofar as it is useful in helping the client. Drawing on pragmatic understandings of science and explanation, this can be further specified using the notion of *pragmatic values*. Pragmatic values are qualities that a formulation/explanation can have which make it more useful, such as proximity to loci of control/influence, incorporation of factors important to the client, and communicability (Potochnik, 2016). The approach utilized within ATOM meanwhile could be described as *epistemic*. ATOM recognizes that, given its status as an explanation, there are a variety of *epistemic values* that we can use to assess the 'truth' of a formulation (T. Ward et al., 1999, 2016). Epistemic values are qualities that pertain to the likelihood of accuracy, such as parsimony, internal coherence, coherence with other established knowledge, and predictive value (Haig, 2014; Thagard, 2017).

Explanatory Purpose. This factor refers to the underlying perceived purpose of developing a formation, and is again not well modeled as a continuum. As discussed it is generally agreed that formulations are hypothesized explanations for a person's difficulties that can help guide attempts at intervention, however within this there is still significant variation as to the intended purpose. For example, some approaches to formulation place emphasis on explaining someone's difficulties in a *depathologizing* way—as a normal or reasonable response to abnormal or difficult circumstances. A clear example of such an explanatory purpose would be the Power Threat Meaning Framework which seeks to highlight the relationship between someone's difficulties and adverse environments and experiences across their life (Johnstone et al., 2018). Other approaches do not place emphasis on depathologizing distress and may instead see the purpose of formulations as more *purely explanatory*. ATOM would be an example of this in that, as we saw when discussing the 'nature of truth' factor, formulations under ATOM are evaluated on the basis of epistemic rather than moral values (T. Ward et al., 1999, 2016). Another example of variation in this factor would be formulations developed to suit particular *administrative* needs. For example, in practice, formulations may be completed with the intention of highlighting causal links between a person's presenting difficulties and certain life experiences in order to facilitate access to restricted supports such as funded sessions.

The Explanatory Target. This factor relates to how approaches to formulation differ in regard to what is seen as the appropriate explanatory target. This factor is also not modeled well as a continuum and appears to be the least widely discussed of the factors I have highlighted here. Many approaches simply assume that the appropriate target of explanation for a formulation is 'the presenting problem'. While obviously true, this answer is somewhat naïve in that it is not sufficiently specific. By 'the presenting problem' one could just as easily be referring to the specific complaint that led a client to seek help (e.g., 'I am having trouble sleeping and I have lost weight recently without really trying'), the wider set of symptoms or difficulties disclosed during assessment (e.g., flat mood, lack of appetite, early morning waking, anhedonia), or the diagnosis offered by a clinician to categorize and label the difficulty (e.g., depression). Another approach to specifying the explanatory target involves collaboratively developing a list of agreed upon questions to be answered or explored. Such an approach, based on my limited understanding, is often utilized in family or systemic therapies (Dallos & Stedmon, 2013). The ATOM approach to specifying the target of explanation is somewhat technical, but brings clear advantages in terms of specificity, so is worth briefly unpacking here. Within ATOM, distinctions are made between the initial presenting complaint, the focus of enquiry (i.e., a refined version of a presenting problem including core questions that need to be answered), and the clinical phenomena (i.e., reliable clinically relevant patterns inferred from the information sourced during assessment). Within ATOM it is the clinical phenomena which one seeks to explain with a formulation. A key concept that ATOM draws on is that of the data-phenomena distinction outlined by Bogen and Woodward (1988). Under this distinction 'data' constitutes information about the world and is inherently fallible and noisy, while 'phenomena' are inferred based on reliably detected patterns within the data. The intention of a formulation is clearly not to explain every moment-to-moment choice a patient makes or account for every individual item on psychometric measures they complete (i.e., to explain the data). Any attempt to follow up on and explain *every* idiosyncratic data point in a person's report or history is likely to lead a clinician down the proverbial garden path. Rather, as specified in ATOM, a formulation should seek to explain clinically relevant

and reliable patterns within the data available—i.e., the clinical phenomena (T. Ward et al., 1999, 2016). Making this distinction brings clear advantage when synthesizing and making sense of large amounts of complex information, as one does when developing a formulation. As I will argue below, for those aiming to recognize the embodied, embedded, and historically informed nature of psychological problems, utilizing this data-phenomena distinction when specifying their intended targets of explanation is a practical necessity.

The Explanatory Form. It is well recognized in philosophy of science that explanations come in many different forms (Thagard, 2017), and this factor pertains to the style or form of explanation utilized in formulation. Again, this factor is not well modelled as a continuum with a variety of approaches taken. Formulation as endorsed by Johnstone (2018) or as seen within the Power Threat Meaning Framework (Johnstone et al., 2018) for example, appears to take a *narrative* form in that they are focused on tracking important historical factors and developing a sense of interpersonal understanding of the client and the challenges they are facing through telling their story. The collaborative observations made during emotion-focused therapy meanwhile, while not usually understood strictly as formulations, can be understood as *dynamic* style explanations which track relationships between events without specifying any underlying causes (Greenberg, 2004). *Dispositional* style explanations alternatively seek to explain a person's challenges in reference to traits/dispositions that they have (Vanderbeeken & Weber, 2002). For example, explaining someone's difficulties regulating their drug use in reference to their being highly impulsive and having poor problem-solving skills. A limitation of this approach is seen in terms of depth in that it is left unclear where exactly these dispositions come from and what one might be able to do about them. *Mechanistic* approaches meanwhile attempt to isolate key components/factors that, when understood in context and interaction, offer an explanation of a target phenomenon, make it less surprising, and suggest ways that it might be manipulated (Glennan & Illari, 2017). ATOM is a good example of a mechanistic approach to formulation (T. Ward et al., 1999, 2016). The well-known 5-P model meanwhile, as described earlier, attempts to explain a presenting problem by essentially listing putative causal factors. This is in some ways similar to a

mechanistic approach in that it attempts to highlight important causal factors, however it does not seek to understand how the factors integrate to bring about and maintain the presenting problem. This can be referred to as a *list-based* approach to explanation.

Formulation and 3e Psychopathology

Now that some of the important ways that approaches to formulation differ have been outlined, we can begin to consider what formulation may look like under 3e Psychopathology. This discussion will be broad in nature because, as I will argue, 3e Psychopathology does not prescribe any one particular approach to formulation. As such it is *not* my intention to present a method of formulation here. Rather my intention is to explore some implications of 3e Psychopathology for the task of formulation. I will first outline the general function of formulation under 3e Psychopathology, showing that the general intention of developing a tailored explanation for a person's difficulties remains. I will then argue for a methodologically and explanatorily pluralist understanding of the act of formulation based on the multi-scale complexity that 3e Psychopathology conceives is at play in mental disorder. I will aim to show that this is not an unrestrained pluralism whereby 'anything goes' and will highlight some of the core constraints that 3e Psychopathology offers during the formulation process. Following this, I will continue to explore the nature of the methodological guidance offered by 3e Psychopathology by considering the degree and nature of constraint placed upon each factor in the taxonomy presented above. Near the end of this discussion and in the following section some core ideas will be synthesized into two tools which may be helpful in the development of 3e Psychopathology informed formulations. The first such tool is a table of principles underpinning good formulation when mental disorder is understood through a 3e lens. The second tool is a graphical representation of the process of moment-to-moment sense-making which I have found helpful in a clinical context when gathering experiential data for formulation development, as well as for encouraging meta-cognitive reflection and understanding.

The Function of 3e Formulation. As should be clear from the earlier discussion, the act of formulation is generally agreed to concern the generation of a hypothesized explanation for a patient's presenting problems. Essentially then, a formulation is an attempt at explanation at the idiographic level. Under 3e Psychopathology this intention remains. As presented across previous chapters, 3e Psychopathology understands mental disorders to be dysfunctional patterns in sense-making constituted and influenced by complex multi-timescale process structures spanning brain, body, and environment. These structures are deeply interwoven with the sense-making processes of an individual and are defined by how they hold said individual stuck within a pattern of engaging with the world that does not work for them. The basic function of formulation under 3e Psychopathology then, is to capture or grasp a *good enough sketch* of the complex structures keeping the patient stuck in this unhelpful way of engaging with the world. In essence, a formulation as viewed through 3e Psychopathology needs to capture enough of what is going on within the brain-body-environment system to offer insight as to why a person is stuck. Moreover, in doing so it needs to offer guidance as to what might help get them unstuck.

A vital question at this point is, what exactly does it mean for an explanation to be a 'good enough sketch'? I have used this phrase here to highlight that, as a *situated* explanation developed for a *particular purpose*, a formulation is subject to a complex array of values both epistemic and pragmatic. As discussed earlier, epistemic values are qualities that make an explanation more likely to be 'true' or accurate and valuable in terms of knowledge (Haig, 2014; Thagard, 2017), while pragmatic values concern qualities that will make an explanation fit for its intended purpose (Potochnik, 2016, 2017). In recognizing that the act of formulation concerns the development of a situated explanation then, 3e Psychopathology would invite recognition of both epistemic *and* pragmatic values during the construction and evaluation of formulations. Examples of relevant epistemic values include: explanatory breadth/scope, coherence with other established knowledge, internal logical coherence, and (arguably) simplicity. Examples of relevant pragmatic values include communicability, proximity to loci of control, and a balance between explanatory aspiration and limitations on time and resources.

The recognition of pragmatic values is vital because doing so acknowledges that different kinds of explanation/formulation are better suited for different contexts. For example, a psychiatrist who prescribes medications will likely see greater value in the incorporation of biological causal/constitutional factors into their formulations because they want to be able to predict and make sense of the possible impacts of the medications they prescribe. As a psychologist on the other hand, there is little practical value to incorporating detailed biological factors into my formulations because these are not within my loci of influence. Instead, my formulation is of greater practical value if it focuses on capturing the psychobehavioral process at play, because this is what I can target with the development of a psychotherapeutic intervention. Different contexts and purposes thereby call for different explanations, demonstrating the role of practical values in the evaluation of formulation. In particular I have also used the phrase 'good enough sketch' to highlight that clinicians and their clients/patients are themselves situated and thus limited in time and resources, thus placing further practical limitations on their efforts at formulation. A 'good enough sketch' then is one that, on balance, accords with such epistemic and pragmatic values and limitations. Such a 'good enough sketch' can then continue to be refined, or indeed replaced, as understanding develops and further information is uncovered. In the following section I aim to begin showing how the gentle methodological guidance offered by 3e Psychopathology and its particular conception of mental disorder can facilitate the development of such valid and useful formulations in clinical practice.

Before moving on, it is important to recognize that in focusing on formulation what I am describing is a *problem-focused* conception of therapeutic encounters. This accords with the wider intention of this book to begin to develop a 3e inspired approach to *psychopathology*, i.e., the study and treatment of mental disorders. I wish to be clear that this is not to argue against the value of more *person-focused* or *relational* approaches, such as the recent enactive analysis of the therapeutic encounter through the lens of participatory sense-making by García (2022). Such relational and person-focused approaches place less emphasis on understanding 'the problem' or 'psychopathology', and rather place greater emphasis on the person-as-a-whole, as well as the situated relationship with the clinician.

One important point of similarity between relational approaches and 3e Psychopathology however, is in recognition that the clinician themselves are situated and subject to their own historicity. Like it or not, formulation is not some abstract process that can be perfectly represented in methodological discussions such as I am engaged with here—it is rather an embodied and situated activity. It is not after all only clients/patients that are embodied and shaped by their history and context! Recognizing this brings a natural demand for clinicians to be reflective as to their own histories, contexts, the decisions they make, and how they are themselves engaged in a process of sense-making in the therapy room. Formulation (and most other psychotherapeutic techniques) can be understood as sense-making about sense-making (Aftab & Nielsen, 2020). Both clinician and patient are seeking to make sense of the problem at hand—how it emerged and how it is sticking around—but also to find new ways to make sense of the situation that may afford new and more adaptive ways forward.

A Constrained Methodological Pluralism. If all readers of this book somehow miraculously agreed that mental disorders are indeed best conceptualized as they are within these pages, this would still not mean that there was suddenly one best way to approach formulation. Parallel to the discussion of pluralism regarding nomothetic explanation in the first half of this chapter, the complexity highlighted by the 3e view of mental disorders and human functioning strongly suggests that there are likely to be many valid ways to approach formulation. In other words, 3e Psychopathology endorses methodological and explanatory pluralism regarding formulation. This is of course not to say that 'anything goes'. In this section I will outline some *guiding ideas* for formulation under 3e Psychopathology, including exploration of the degree and nature of constraint that 3e Psychopathology offers upon each factor of the conceptual taxonomy outlined earlier.

Under the 3e view, the mind is not 'a thing' locked away from the world and shaping how we act, rather it is a continued process of engagement with the world and with one's self (de Haan, 2020b; Fuchs, 2017; Nielsen, 2020b; Thompson, 2007). This is a fundamental difference when compared to many other psychotherapeutic approaches. Cognitive Behavioral Therapy [CBT], for example, is ultimately based on a linear

model of cognition whereby we have certain 'core-beliefs', which then bias the way we interpret situations, shaping how we think, feel, and act (J. S. Beck, 2020). The therapeutic goal in Beckian CBT then, is to explore ones thinking enough to recognize how these 'core-beliefs' are encouraging biased or irrational ways of thinking, and thus feeling and acting. Behavioral strategies are utilized within CBT, but they are typically understood as 'experiments'; in which the person gathers counter-evidence to their core beliefs, and which thus ultimately work by changing our thinking. Within this model, core-beliefs and thoughts are given a certain priority/centrality, and it is from these and other 'cognitive errors' that mental disorders are ultimately seen to flow. A similar focus on cognition or 'the mind' and the way it shapes emotion and behavior can also be seen in psychoanalytic approaches which typically seek to understand or resolve hidden conflicts, drives, or other unconscious motivations that are seen to support dysfunctional behavior. Under the 3e view however, 'the mind' is understood in a *circular* rather than *linear* manner, and thus not locked away from the world (Fuchs, 2017). Dysfunctional patterns of thinking, feeling, and acting, are not seen to *stem from* dysfunctional core-beliefs locked away in the mind, but rather to *emerge within* the circular engagement between organism and world over time. Moreover, thinking and languaging are seen as important aspects of human meaning-making, but are not given *a priori* explanatory privilege in shaping human distress and dysfunction. Thinking, languaging, attending, remembering, imagining, these are all understood as *things we do* as organisms. Just like every other behavior, they are shaped by our evolutionary history, cultural history, individual history, context, and embodiment; but as part of our history they continue to unfold as we walk into the future. No matter what form sense-making takes, it remains an historically shaped, active, and world-involving process.

Given this underlying model of cognition as sense-making, the emphasis in clinical formulation under 3e Psychopathology should not be on inferring mental/static/theoretical/dualistic constructs that seem to explain an unhelpful pattern of behavior (e.g., core beliefs, drives/conflicts, personality types, or 'trauma' as an energy lodged in the body). Rather, it should be on breaking down and identifying the constituent *tendencies* or *recurring patterns* in sense-making that constitute or

contribute to the problem at hand and explaining these tendencies/patterns in reference to the person's history, context, and embodiment. This explanatory step is undoubtedly an abductive inference—i.e., an inference to the best explanation—thereby according with ATOM as described across earlier sections, however, the explanation inferred is constrained by the wider commitment to embodiment and the process-orientation of enactivism, as well as a wider recognition of pragmatic values (Haig, 2014; T. Ward et al., 1999, 2016). For example, in a PTSD type presentation, a 3e informed formulation of the phenomenon of hypervigilance developed for the purpose of targeting in psychotherapy would not postulate a hidden belief that 'the world is a dangerous place' from which a 'biased' emotional and behavioral response stems. Rather it would first describe the recurring hypervigilant pattern in the person's engagement with the world as an unfolding process (e.g., constant scanning of the environment for potential threat, the embodied experience of increased fear and anxiety when ambiguous stimuli are noticed, the subtle relief and sense of safety felt when the ambiguity is resolved, the reengagement with scanning behavior). It would then seek to explain this pattern in reference to the person's history with traumatic events, the resulting upregulation of their various embodied threat responses, the myriad of other 'symptoms'/tendencies that maintain a high baseline level of stress/threat (e.g., sleep disruption, intrusive memories), current contextual stressors and threats, the mammalian tendency toward inflexibility and short-term reward when under conditions of stress, and how this behavior is reinforced through the short-term sense of safety it provides. It would also then consider how this behavior itself maintains stress, thus contributing to the maintenance of the wider set of observed tendencies in sense-making that constitute the person's 'PTSD'.

Instead of postulating 'core-beliefs' or other purely theoretical concepts then, formulation under a 3e conceptualization should seek to explain the presenting problem as *a process of engagement and meaning-making which continues to be enacted*. The explanatory tools one utilizes should generally be more thoroughly embodied, developmentally minded, and process-oriented concepts. Examples include but are not limited to: behavioral reinforcement or punishment, behaviors having a previously adaptive function (e.g., hypervigilance and emotional avoidance may

have been useful in a previous abusive context), developmental disruptions to skill development and resulting lack of know-how (e.g., someone never learning to recognize and label certain emotions, accept their own emotions, assert their needs in certain kinds of relationships, or preserve their sense of self-value in the face of others achievement), 'normal' human or mammalian wide regularities in sense-making (e.g., decreased flexibility/creativity when stressed, state-dependent memory, mood lowering with lack of positive activities or purpose, delay-discounting), and the effects of differences in embodiment both directly and in terms of how they may have affected development (e.g., medical conditions, physical or sensory disabilities, neurological differences such as with Autistic persons or those with ADHD, and how these differences may have affected the development of emotional or interpersonal skills). Utilizing such explanatory tools means that the presenting problem is formulated as a historically informed yet actively unfolding process. Biological differences such as genetic predispositions, and social realities should absolutely be recognized, particularly when doing so brings practical advantage, but they should be understood as part of the wider dynamical constitution of the problem rather than as ultimate causes. As well as aligning with the wider enactive view of human functioning, the practical benefit of this is that it highlights the patient's agency—i.e., their active role in the problems they are facing. Rather than being subject to an 'underlying' difference in their cognition or biology, formulating in this way unpacks how a person is continuing to enact a pattern of engagement with the world that is not working for them. Vitally, it offers a sense of why this may be the case, thus highlighting ways that they may have power to change the process, try something new, and perhaps get unstuck.

At this point a general sketch of case formulation under 3e Psychopathology has been offered. In order to continue exploring and specifying some general guidelines for formulation under a 3e conceptualization, I will now return to the taxonomy of approaches to formulation presented earlier. Considering this taxonomy, the various factors vary in the degree to which ones positioning upon them would be constrained by adherence to the theoretical and conceptual assumptions outlined across the earlier chapters of this book. For example, it should be clear from earlier discussion that 3e Psychopathology with its view of

disorders as having a constitution dispersed across the brain-body-environment system would accord with a system-wide approach to formulation over a purely 'psychologically' focused one in a restricted sense. Meanwhile it is not clear that 3e Psychopathology as here developed would immediately point to the selection of any one explanatory style. I will now consider the degree and nature of the methodological guidance offered by 3e Psychopathology for each factor in turn. While I by no means claim to be an experienced therapist, I will at times draw on clinical examples in order to demonstrate the real world impact of these sometimes overly technical and philosophical discussions. I will not cover the 'explanatory purpose' factor as this has already been sufficiently discussed in this current section.

Theoretical Loyalty. The 3e Psychopathology conceptual model of mental disorders recognizes the interplay of causal factors from across the brain-body-environment system, and their complex integration through the agentic sense-making processes of the situated individual. Given the complex and multi-scale nature of this conceptualization it seems unlikely that a formulation grounded in a single theoretical lens will outperform integrative or trans-theoretical approaches that offer an opportunity to draw on multiple perspectives when seeking to explain someone's challenges. As such a 3e Psychopathology orientation should encourage consideration of multiple theoretical perspectives when developing a formulation. In integrating the various perspectives into an understanding of the problem at hand however, a clinician should consider the core principles of: embodiment (i.e., the different perspectives used should not utilize dualistic or overly abstract concepts and should accord with wider knowledge of human functioning), embedment (i.e., the different perspectives used should offer genuine consideration of historical and current context in shaping the problem at hand), enactivism (i.e., the different perspectives used should capture the role of meaning and experience and present the problem as a continually unfolding process) and situated normativity (i.e., highlight/question how the problem is a problem for the client on their own terms). Genuine consideration of these principles, and others—both pragmatic and epistemic—considered throughout this section and summarized in Table 6.1, should sufficiently constrain the explanatory pluralism being prescribed here. In recognizing

that formulations are situated explanations and thereby subject to epistemic and pragmatic values, 3e Psychopathology also accords with a trans-theoretical view of formulation. As described earlier, such perspectives understand formulation as a method or inferential process separate to the particular theoretical tools one brings to bare.

It is also worth noting that the degree of constraint/guidance offered by 3e Psychopathology upon this factor is really only moderate in nature. As discussed, formulation is a situated and practical endeavor as well as an explanatory one and, as discussed, pragmatic values are relevant when evaluating one's approach to formulation. There is therefore nothing directly impeding a 3e-oriented clinician from utilizing a single theoretical lens in the development of a particular case formulation, so long as they have considered reasons for doing so. While it seems unlikely that, compared to an approach utilizing multiple theoretical lenses, a single theoretical lens, such as CBT, will be able to capture as many aspects of the disorder process as it is conceived under 3e Psychopathology, a clinician may still choose to stick to a single theoretical approach for practical reasons. For example, they may have limited time available for the assessment, they may work in a team that likes to utilize a certain shared theoretical language, they may reason that a particular theoretical approach may facilitate engagement by the client, or perhaps the client has requested a particular treatment modality. In short, while a 3e approach may push a clinician towards the integrative or trans-theoretical end of this spectrum, the relative explanatory and pragmatic costs and benefits still have to be considered when choosing which, and how many, theoretical lenses to utilize. The principles of formulation outlined in Table 6.1 represent one way to do so.

Tailoring. As outlined across previous chapters the 3e conceptualization of mental disorder is deeply agential. Even if we come to discover that some currently recognized mental disorders or subtypes thereof have tightly knit essence-like causal structures buried somewhere in the brain, the 3e perspective would push us to recognize that the functional impact of such structures plays out through the agent's sense-making in the world. As such, every presenting problem is assumed to be unique—my depression is not your depression and so on. In short, under 3e Psychopathology, mental disorders are acknowledged as messy, contextual, agential, and world-involving. Under such a view it seems unlikely

that a purely formulaic approach to formulation will ever sufficiently compete with a more flexible and bespoke approach whereby one can shift and adapt to capture the unique challenges of different clients and to respond to the specific needs of the assessment context.

In general then, 3e Psychopathology aligns with a bespoke approach to formulation and holds this an ideal. One caveat to this however, concerns the relationship between formulation and the wider sciences of psychopathology. Earlier, in the nomothetic section of this chapter, I discussed the idea of epistemic units and how their clinical utility is in being able to serve as knowledge bridges between clinicians and the wider research-base. It was discussed, for example, how current DSM-ICD constructs are much too large and unstable to offer clear understanding of a client's problems simply on the basis of diagnoses. The idea of a *phenomena complex* was discussed as a more appropriately sized epistemic unit that, if developed and utilized within explanatory theories, may improve the ability of clinicians to be able to draw on scientific knowledge when formulating. Under a 3e approach to formulation the intention is to develop a bespoke explanation for a client's difficulties—one that is particular to their history and situation—however, this does not mean that clinicians are not free to import such 'phenomena complexes' or other epistemic units into their formulations (for further discussion see: Nielsen, 2022b). There is clearly great practical and epistemic utility in being able to recognize a 'phenomena complex' or other such patterns within a client's presentation that are also recognized in scientific literature (or alternatively that are common within the clinician's regular client base). The process of noting such a pattern is almost pseudo-diagnostic in that the logic is more about pattern recognition than generative abduction. When aspects of the formulation are inferred in this way then, there is a sense in which the process becomes somewhat more formulaic rather than completely bespoke. This does not seem a bad thing however. Utilizing such epistemic units in formulation is efficient and therefore practical, encourages common language, and further allows for formulations to improve as scientific knowledge develops. In instances where this more diagnostic style of inference is utilized however, it is still vital to situate the inferred mechanism within the wider formulation and one's understanding of the client's mode of functioning.

As an example of this, I see many clients who are dealing with difficulties recovering from a concussion/mild traumatic brain injury (mTBI). Most people who experience an mTBI recover within a matter of weeks to a few months, however a significant proportion of people, for a variety of reasons, experience delays in their recovery. Within this population there is a very common pattern that emerges, often referred to as 'boom and bust', whereby someone will attempt to push through their symptoms of headache, fatigue, concentration difficulties and so-on, in order to try and return to important life activities. This pattern often plays a role in slowing recovery as it disrupts the ability to rest and recover, to gradually re-build tolerance to everyday stressors, and to develop a sense of agency over the recovery process itself. This pattern is therefore extremely useful to notice as a clinician. It provides me with a common language to communicate with other professionals and to engage with literature. 'Boom and bust' and its relationship with other difficulties such as sleep and mood disruptions will therefore often act as an important epistemic unit within my formulations with these patients. However, 'boom and bust' is still quite a causally heterogeneous phenomenon. Some of the patients I meet 'boom and bust' due to a difficulty tolerating their anxiety about recovery if they pause and rest, some 'boom and bust' because they have a history with chronic pain and have learnt to habitually ignore interoceptive signals of distress from their body, and some 'boom and bust' because their disrupted sleep-wake cycle has left them so exhausted that they just continue to work by force of habit. 'Boom and bust' then is a vital epistemic unit in my formulations with these clients, and my recognition of it has become somewhat formulaic as I continue to gain experience in this area, however, there is still a clear need to situate this 'boom and bust' pattern within the wider formulation, and within what is known about the client's wider mode of functioning.

It is important to note that, much like the factor above, the degree of constraint/guidance upon this tailoring factor is relatively moderate. There may well be practical reasons to assume a more formulaic approach in a certain context even if committed to a 3e conception of mental disorder and dysfunction, such as their only being funding for a short

number of sessions available. In general though, a bespoke formulation remains the ideal. It should also be stressed that 'flexible and bespoke' is not synonymous with 'esoteric' or 'random'. The best bespoke suit-makers in the world still presumably follow a method which allows them to reflect on and improve their processes and to flexibly pivot their designs to fit diverse bodies. The same seems true for clinicians.

Degree of Collaboration. Under 3e Psychopathology, mental disorder is understood to be a significant disruption to one's agency and to one's ability to respond adaptively to the world as it changes around us. The ultimate intention of any therapeutic encounter or intervention then, should not simply be to alleviate 'symptoms' but to restore and develop agency. Patients are understood to be stuck in a pattern of making sense of and responding to the world that is working against them, not broken and in need of 'fixing'. In this sense, the therapeutic encounter is not usually understood as one whereby the clinician 'acts' on them, but one whereby patient and clinician come together to make sense of the problems the client is facing, and the clinician attempts to scaffold the client in developing and exploring some new ways forward. This is true both for clinicians utilizing psychotherapeutic intervention, and for those using psychopharmacological intervention—3e Psychopathology would encourage a view whereby the ameliorating effects of medication, coupled with support and behavioral change, are used to scaffold the development of agency over the challenges faced, rather than to simply 'alleviate symptoms'. Whatever the intervention tools available, active involvement by the patient/client is seen as essential to both therapeutic change and to the development of good enough formulations under a 3e view. In essence, it is the clients themselves that are figuratively stuck walking the same circular path, therefore they are the ones who have to lay a new path by walking it. Simply telling someone about an alternative path to take is unlikely to be sufficient, for even if they manage to find it, they have not learnt how to find it again in the future. Good formulation from a 3e view therefore, defaults towards being deeply collaborative and experiential in that the ultimate intention is to build agency and the patient's ability to find their own ways forward.

Much like with the earlier factors however, this leaning towards collaboration is only a general guideline, and practical considerations have to be made. To briefly draw on Vygotskian thought, agency is clearly not best developed simply through being told what is happening and what to do about it, but nor is it best developed by being left to one's own devices. Rather, agency is developed through being supported to come to understand, explore, and figure out ways forward, in a scaffolded way whereby the degree of support and guidance is matched to optimize skill development and is slowly reduced as skills and independence develop. While there remains a general default to, and spirit of, collaboration, insofar as a formulation is intended to be therapeutic a 3e orientation supports an approach where a clinician can flexibly shift between more collaborative and more didactic approaches in order to meet the needs and preferences of the patient in front of them.

As an example of this, I often see clients under our national accident compensation scheme and many of these clients are young men injured playing sport or while working in a physically demanding trade. For those unfamiliar with Aotearoa-New Zealand men, a common stereotype is that they are somewhat stoic and tend to deal with difficult emotions in non-verbal ways. While this remains a stereotyped view, it does exist for a reason. When working with these men on difficulties following injury such as mood, anxiety, acute trauma reactions, or suicidality, I often have a very limited number of funded sessions, hard won by a valiant occupational therapist who has argued for their access to sessions (and who has often put in a lot of work to convince the client to actually turn up). Given the time limitations, and the possibility that jumping into discussion of their thoughts and feelings may scare them out the door, I will often pivot to a much more didactic approach to formulation with these men. After a brief assessment and identification of some common mechanisms maintaining their difficulties (e.g., lack of positive activity, sleep disruption, fear-based avoidance) I will often feedback a simple formulation (often couched in an analogy to their favorite food, or something else they have a personal interest in), prescribe some clear steps that should help (e.g., positive activities, clear routine, reducing

alcohol), and motivate the client in reference to something that matters to them (e.g., responsibility to their partners or families, getting back to valued activities). This is a much more didactic approach than I would usually take, to the point that I often feel somewhat uncomfortable simply telling them what to do. However, once they see improvement in the problems they face, they will typically recognize they have more power to influence how they feel than they realized. At this point they will often return, hungry to discuss their insights as to how *they* managed to influence how they think and feel. It is always amazing to see how quickly the stereotype falls down when these men are given scaffolded opportunity. At this point I will shift back to a much more collaborative approach and the formulation continues to develop. In the language of 3e Psychopathology, this and other such therapeutic moves made the world over, are an example of flexibly matching the process of formulation to the client's mode of functioning. It is also perhaps a good example of how the collaborative-didactic spectrum is not as simple as it first seems, and that it is perhaps better understood as a dialectic. Sometimes, due to practical constraints or the particularities of the client, being more didactic is the most genuinely collaborative and agency-building option.

Selection of the Explanans. The conceptual model of mental disorder developed across previous chapters strongly accords with a system-wide approach to formulation, whereby causal factors are considered from across brain, body, and environment, all across time. This however, is far from endorsement of an unconstrained explanatory holism. As we will see when discussing the explanatory target within 3e approaches to formulation below, the target of explanation remains 'psychological' in nature—i.e., patterns/tendencies in sense-making that appear to contribute to the presenting problem. Clinicians oriented to a 3e conception should seek to consider: factors relating to the state of the environment (e.g., family and interpersonal dynamics, access to basic needs such as housing and food, current cultural milieu, exposure to social stress, inequity, or threat), factors relating to the state and experience of the body (e.g., nutrition, physical health/ill-health, pain, sensory differences, medications, drugs and alcohol), and historical/developmental factors including their wider mode of functioning (e.g., past learning, historical

environments, historical health, genetic history, disruptions to typical developmental trajectory). A 3e Psychopathology approach therefore is clearly a systemic one. The aim however is targeted and psychological in nature—i.e., to infer an explanation for the observed tendencies in sense-making that appear to contribute to the problem at hand.

The Explanatory Target. The explanatory target of formulations under a 3e Psychopathology approach are reliable tendencies or patterns within a patient's sense-making that appear to contribute to or maintain the difficulties they are facing. This approach aligns with ATOM and its respect of the data-phenomena distinction as described earlier, and indeed these reliable tendencies/patterns are reasonably analogous to the concept of clinical phenomena within ATOM (Bogen & Woodward, 1988; T. Ward et al., 1999). The importance of this data-phenomena distinction is hard to overstate. Clinicians are very frequently in the position of synthesizing large amounts of information in order to make recommendations for care, or simply to guide their clinical decision-making in the room. Distinguishing between data (e.g., individual items on test scores, or single comments made by patients) and phenomena (e.g., recurring processes of thought observed across sessions by both clinician and patient, consistent patterns emerging across tests scores and self-report, etc.) allows clinicians to not get caught on the tides of unreliable variation in the information they are presented with. In short, before spending effort trying to explain some apparent tendency or pattern, it is vital to pause and consider what evidence there is that it is occurring reliably, and important enough to justify the time and effort it will take to explain it.

This of course raises the question of what counts as reliable tendencies, patterns, or phenomena. Given the complexity recognized within the 3e conception of mental disorder/dysfunction, 3e Psychopathology does not conflict with established answers to this question concerning principles of good psychological or psychiatric assessment. In order to establish reliable patterns so that they may then seek to develop explanations for them, clinicians should utilize multiple sources of information wherever possible. If patterns continue to be present (or to vary meaningfully) across time, contexts, and multiple data sources (e.g., client reports, developmental histories, behavioral observations,

informant reports, psychometric assessments, neuropsychological testing, and historical data such as school reports and previous assessments) then this provides confidence that the patterns inferred are indeed real and worth trying to explain. Beyond this received answer however, 3e thinking, with its roots in phenomenology and its breaking down of the subjective-objective divide, also encourages recognition that the most important source of information is the patient themselves. Through the deep continuity thesis and recognition of the strive to survive inherent within all organisms, there comes a certain trust in the patient. Even in the rare case that patients are actively lying or malingering, it can be assumed that patients are trying to do their best to survive and adapt. Further, as constant witnesses to their own experience, patients themselves are in the best position to report on and gather information about their own sense-making. One must of course maintain an awareness of the imperfection of human memory and issues with biased reporting, but 3e Psychopathology also accords with a general spirit of trust in the client. Through this trust one listens to a patient's experience first, and doubts after the fact in light of conflicting evidence of incongruity, rather than a position whereby one doubts the validity of subjective report from the very beginning. Doing so also has great therapeutic potential through encouraging engaged collaboration and meta-cognitive reflection. Much like quantum phenomena, one cannot actively observe one's own thoughts without changing their behavior. As alluded to when discussing the collaborative-didactic factor, engaging the client collaboratively in the assessment process is often a great way to facilitate engagement and the development of agency. Further, as will be discussed in the following section on the 'nature of truth' within 3e informed formulations, it is not assumed *a priori* under the enactive worldview that subjective experiences are not objective, or are otherwise somehow less valid than 3rd person information.

The Explanatory Style. In accordance with the methodological pluralism prescribed, 3e Psychopathology does not align strongly with any one particular explanatory style. Rather a 3e Psychopathology perspective should encourage clinicians to reflect on how the explanatory style of their formulations can be adapted for different contexts and purposes. There is thereby a freedom to choose an explanatory style or to hybridize

as needed. For example, a mechanistic style of explanation is very useful for explaining the reafferent causal structures that support the maintenance of mental health problems under the 3e view. They are thus well suited for reasoning about the selection of treatment targets and interventions. Mechanistic explanations are also fairly easily summarized into a list format which can be helpful for keeping report writing concise but informative. Narrative approaches to explanation meanwhile, accord well with the historically informed nature of experience and mental disorder under the 3e view. By 'telling the story' of how a client and the challenges they are facing, narrative style explanations often highlight how the problems faced have emerged through the client striving to survive and adapt to difficult circumstance, thus evoking compassion. This can be of clear utility when explaining one's formulation to self-critical clients or support people, and when trying to argue with the powers-at-be for access to resources. Through 3e cognition's roots in phenomenology and complex systems theory there is also clear place for dynamic style explanations. For example, when the precise mechanism is unknown or still coming to be understood, one can simply comment on an observed reliable tendency as one developing part of a wider case formulation, analogous to a differential equation tracking system dynamics in complex systems. For example, one could share with a client, 'it seems like every time you are presented with X you feel Y and respond with Z, does that sound accurate?'. Such a simple dynamic explanation/description, while of minimal explanatory value for the mechanistically minded, can have clear utility in facilitating client engagement and curiosity, and may simply reflect an honest understanding of the problem when clinician and client are not yet able to confidently infer reasonable mechanisms for the observed dynamic. Such descriptions of the relationships between phenomena still have a degree of explanatory value in that they continue to refine the understanding of the problem at hand. In sum, 3e Psychopathology endorses a flexible relationship with explanatory style in order to best meet the contextual demands on one's formulation.

The Nature of 'Truth'. As discussed earlier, the complex and multiscale conceptualization of mental disorder within 3e Psychopathology encourages methodological and explanatory pluralism in the development of formulations. In other words, under 3e Psychopathology it is

assumed that there are multiple valid ways to go about developing a formulation and, even for a single client, there are likely to be multiple valid ways to explain their presenting problem—i.e., different angles from which to understand the client thus capturing different aspects of the challenges they face. This explanatory pluralism thereby brings with it a deep valuing of *epistemic humility*—i.e., holding one's own understanding with appropriate confidence, yet doing so lightly and always being open to alternative viewpoints and ideas. When faced with an alternative perspective on a patient's difficulties, whether from the client themselves or from a collaborating professional, it is natural to want to argue and defend ones understanding. Explanatory pluralism as endorsed here however, encourages us to ask: how is their perspective different to mine, what does it capture that my formulation does not, and what is potentially useful about this perspective?

Clearly the underlying notion of 'truth' here is not one where truth or accuracy can be simply inferred due to predictive value or alignment with current clinical evidence. As mentioned earlier, 3e Psychopathology rather aligns with a pragmatic and epistemically informed approach to 'truth' whereby formulations are subject to a diverse array of both epistemic and pragmatic values/principles that often conflict with each other. Thus, alternative perspectives can be more or less accurate and/or more or less useful in a myriad of different ways. Rather than discuss these various values and principles again in this section, a core selection are summarized in Table 6.1 below. Such principles and values can and should be used in the evaluation of formulations. A related point here is that, under 3e Psychopathology, formulations should be understood to be continually developing things. The same epistemic humility that holds for the comparison of one's own formulation with the perspective of another, also holds for comparison with one own perspective at different time points. Just as sense-making is a continually unfolding process, so too is the process of sense-making about sense-making. One never arrives at 'the truth', rather one should seek to develop a *good enough sketch* during initial assessment and then continually strive to improve it and adapt it to new challenges as they arise.

Table 6.1 Principles/values for the development and evaluation of formulations

Principle/value	Source/reason	Key evaluative questions
Methodological pluralism	Complexity, Humility	How has my process of assessment and formulation constrained my understanding? Given my context, are there other approaches to formulation that may provide greater epistemic or pragmatic value?
Local explanatory pluralism and epistemic humility	Complexity, Humility	Who else has a formulation/perspective on the problem at hand? Does considering these different perspectives highlight areas where my formulation could be improved?
External coherence	Epistemic	Does my formulation fail to cohere with established knowledge in any way I cannot justify?
Internal coherence	Epistemic	Does my formulation make sense? Does my formulation conflict with itself in any way?
Breadth	Epistemic	Does my formulation sufficiently account for all of the phenomena/reliable tendencies inferred during assessment?
Parsimony	Epistemic, Pragmatic	Is my formulation overly complicated? Can I simplify my formulation at all while maintaining its usefulness and accuracy?
Proximity to loci of control	Pragmatic	Does my formulation sufficiently capture parts of the brain-body-enviroment system that I and the client actually have influence over?
Pragmatism/a good enough sketch	Pragmatic	In critiquing and evaluating my formulation, have I considered the practical limitations placed upon me?
Communicability	Pragmatic	Can I communicate my formulation, or at least the necessary parts of it, clearly to the appropriate parties?
Ecological coherence	Embedment	Do I have sufficient information about the client's current contexts? Does my formulation cohere with this information?
Developmental coherence	Embedment	Do I have sufficient information about the client's history and development? Does my formulation cohere with this information? Do I have a sense of the client's mode of functioning, and does my formulation accord with this?

(*continued*)

Table 6.1 (continued)

Principle/value	Source/reason	Key evaluative questions
Physiological coherence	Embodiment	Do I have sufficient information about the client's medical and neurodevelopmental history, physical capacity, and experience of their body? Does my formulation cohere with this information?
Experiential coherence	Enactivism	Do I have sufficient information about the client's experiences/difficulties to be trying to explain them? Does my formulation cohere with what the client's description of their experiences? What do the patient's experiences mean to them, and have I considered this in my formulation?
Cultural coherence	Enactivism, Embedment	Do I have sufficient information about the client's cultural identity as they understand it? What cultural spheres does the client navigate and what is their experience of this? Does my formulation cohere with this information?
Reflexivity	Enactivism, Humility	Are there any assumptions I may be making based on my own culture, politics, or experience?
Awareness of process	Enactivism	What has my client's experience of assessment and formulation development been? Has this process supported their sense of agency/ know-how regarding the challenges they are facing?
Situated normativity	Enactivism	Is this presenting problem a problem for the client or for someone else?' Am I imparting any wider normative judgments based on statistical normality or my own cultural biases?

The Sense-Making Spiral, a Tool for Clinicians

Synthesizing many of the themes discussed across this section on formulation, Fig. 6.2 presents the sense-making spiral, a clinical tool to provide structure for the exploration of moment-to-moment sense-making in a

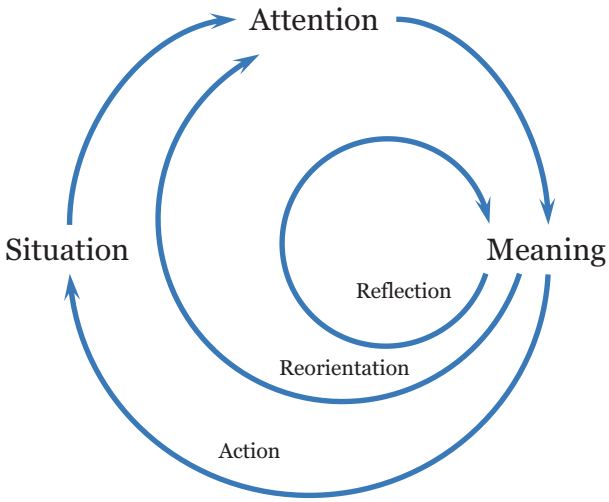

Fig. 6.2 The Sense-Making Spiral. This clinical tool is designed to assist with the collaborative exploration of a client's sense-making. When used across multiple situations, analyzing a patient's experience using this tool may facilitate the identification of reliable tendencies in their engagement with the world that may contribute to the challenges they are facing (copyright retained by Kristopher Nielsen, ©CC-BY-NC-SA)

collaborative manner. The sense-making spiral breaks down moment-to-moment sense-making, allowing for analysis of meaning-making and engagement with the world at any particular point—currently or in recent history—in a way that highlights the circular, continuous, and embedded nature of this engagement. I have developed this tool to be helpful in the identification of reliable tendencies within client's sense-making during formulation, and—more therapeutically—for encouraging client meta-cognitive awareness and emotional agency. This tool presents sense-making as an ongoing and embedded process. I have so far received positive feedback from clients that I have used this tool with (after having attained their informed consent to utilize a developing tool for which there is no current evidence-base). To be clear, the sense-making spiral simply represents one 3e congruent way to identify tendencies in sense-making, and is intended as a supplement to, not a replacement for, thorough assessment. It may be used explicitly in collaboration with a patient, or as an implicit guide when exploring a patient's experience.

When using the sense-making spiral explicitly, as presented in Fig. 6.2, I first introduce it as a tool for breaking down and thinking about one's experiences and how we responded in a given situation. I will then go through each part explaining the meaning of each and listing out some relevant questions, as I will do now. *Situation* refers to the context—to what was going on. Relevant questions include, where were you, who was there, what was happening, how were you feeling? *Attention* simply refers to what you are paying attention to paying attention to at the time, and how. Relevant questions include: what were you aware of, what were you focused on, was it something in the world, in your body, or in your thoughts? The nature or quality of their attention should also be discussed, i.e., whether this attention is focused, diffuse, or perhaps dissociative. Questions focused on this might include: was your attention tightly focused like spotlight, or were you aware of other things too? *Meaning* then refers to what emotions, bodily reactions, and salient affordances (i.e., pulls toward potential actions) someone experiences in response to what they are attending to. Relevant questions here include: how did that make you feel, were you aware of any responses or sensations in your body, were you aware of anything you wanted to do but didn't? At this point I highlight that we can respond to the situation and how we feel about it in different ways, either through reflection, reorientation, or action. I will then explain that *reflection*, or 'the reflective loop', refers to thinking, considering, remembering, imagining, or other cognitive activities that may feedback and alter how we feel in a situation. *Reorientation* refers to shifting or changing the nature of one's attention. *Action* refers to behaviors that change the situation and context such as speaking, moving, or otherwise doing something. Once I have explained the tool, I will then example its use with a recent situation that the client is interested in exploring. Then, once they are familiar with how it works, I ask them to take some notes for 'homework' on situations where the problem we are exploring occurs or is exacerbated, or simply situations where they felt they did not understand their own reactions.

Using this tool collaboratively in the room to discuss how a patient was making sense of and engaging with various situations will often reveal common patterns. For example, it may be noticed that the client often gets stuck in ruminative or existential thoughts, circling around the

reflective loop (see Fig. 6.1), encouraging a passivity of action and suppressing mood over time. Alternatively, a tendency to focus in on stimuli that exacerbate feelings of threat, anxiety, disgust or judgment may be observed—e.g., ambiguous social cues in social anxiety, a racing heartbeat during panic, potential criticism in depression, or triggers for particular obsessions in OCD. As other examples, one can also notice tendencies to reorient away from and thus avoid certain stimuli, to be pulled into intrusive memories, to belittle oneself or self-pathologize in the face of difficulty, or to fail to reflect at all and to jump in and act when angry. It can be advantageous if the client themselves notices the patterns in their sense-making, and drawing out the progress through the spiral as the situation unfolds appears to facilitate this.

Once the client is familiar with using the sense-making spiral as a data collection tool it can also serve therapeutic functions beyond simply passively developing meta-cognitive or mindful awareness. For example, the visual structure of the sense-making spiral also highlights many of the choices one has available when managing difficult experiences. We can reflect on and challenge our own thoughts, reorient our attention towards distracting or calming stimuli such as our breath or features of the environment, or we can act on, change, or leave the situation. These various options are all reflected in different pathways through the spiral. In this way the sense-making spiral already introduces scaffolding for the continued development of emotional agency right from early in the assessment process. Another advantageous function of the sense-making spiral is that it can be used to explore or acknowledge or explore hypothesized relationships between the recurring tendencies uncovered and the client's history or context. Visually I may encircle the entire spiral, labeling this 'history', re-present some of the identified tendencies in sense-making, and then attempt to facilitate discussion about where these recurring tendencies in engagement with the world may have been learnt. An example question might be: can you think of a time in your life, or perhaps a current context, where responding in this way may have been helpful, kept you safe, or met some other need? In effect such recurring tendencies represent embodied predictions as to the state of the world and how best to respond to it, shaped by the client's embodiment and history.

In sum, the sense-making spiral provides one way to explore client's moment-to-moment sense-making, and repeated use may reveal recurring tendencies/patterns in the way that the client engages with the world. It is such patterns/tendencies, revealed and validated through thorough assessment, that are seen as the appropriate targets of explanation under 3e Psychopathology during the formulation process. Using this tool in a treatment-oriented way also represents a plausible and 3e congruent way to facilitate meta-cognitive reflection and emotional agency.

6.4 Explanation in Short

In this chapter I have begun exploring some implications of 3e Psychopathology for the task of explanation. After briefly reviewing some resent insights into the nature of explanation itself I then considered explanation as it is undertaken by researchers—i.e., the development of nomothetic explanatory theory. I argued in broad terms, and in accordance with Fuchs (2017), that 3e ideas demand a methodologically and explanatorily pluralistic approach to psychopathology research. Given the scale of this topic I then focused in on two common ways to organize and target research and explanatory theory, namely: DSM-ICD syndromes and using the RDoC framework. I briefly highlighted some common concerns that DSM-ICD syndromes lack specificity and stability in that they are overly heterogeneous. I then argued that the RDoC approach, when viewed through the lens of 3e Psychopathology, is really aimed at producing neurologically focused component ingredients for our explanations of psychopathology, rather than at developing comprehensive explanations in their own right. I then briefly re-presented the RAP as one plausible way to systematically develop 3e congruent explanatory theories.

In the latter half of this chapter I then shifted to considering explanation as it is undertaken by clinicians—i.e., the task of developing formulations. I presented a rough methodological taxonomy capturing some of the core ways that approaches to formulation differ. I argued that 3e Psychopathology does not point clinicians to any one particular way to formulate, arguing instead for a flexible and pluralistic approach,

gently constrained and guided by 3e principles alongside wider epistemic and pragmatic values. I also argued that 3e Psychopathology encourages clinicians not to understand their clients' difficulties as 'locked away' in their mind, but as complex processes of making sense of and engaging with the world that continue to unfold and develop. I claimed that 3e Psychopathology places strong emphasis on genuine collaboration, epistemic humility, and the fostering of patient agency above and beyond alleviation of 'symptoms'. Towards the end of this discussion I presented two clinically focused resources that summarize core themes from this half of the chapter: Table 6.1 presented a list of principles/values and related questions that can be used in the evaluation of formulations and the processes used to develop them, while Fig. 6.2 presented the sense-making spiral as one way to collaboratively break down and explore moment-to-moment sense-making.

Considering the task of explanation as a whole, a central theme is that 3e Psychopathology as developed within these pages does not point us to any one approach. The complexity highlighted by the 3e view pushes us towards a critical and gently constrained pluralism in our efforts to explain and understand mental disorders.

References

Aftab, A., & Nielsen, K. (2020). *3E approach to psychopathology: Kristopher Nielsen, PhD* [Interview in Psychiatric Times]. https://www.psychiatrictimes.com/view/three-approach-psychopathology-kristopher-nielsen-phd

Bechtel, W. (2009a). Explanation: Mechanism, modularity, and situated cognition. *The Cambridge Handbook of Situated Cognition*, 155–170.

Bechtel, W. (2009b). Looking down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology, 22*(5), 543–564.

Beck, J. S. (2020). *Cognitive behavior therapy: Basics and beyond*. Guilford Publications.

Berenbaum, H. (2013). Classification and psychopathology research. *Journal of Abnormal Psychology, 122*(3), 894.

Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review, 97*(3), 303–352.

Bokulich, A. (2018). Representing and explaining: The eikonic conception of scientific explanation. *Philosophy of Science, 85*(5), 793–805.

Borsboom, D., Cramer, A., & Kalis, A. (2018). Brain disorders? Not really… Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences, 42*, 1–54.

Brigandt, I. (2013). Explanation in biology: Reduction, pluralism, and explanatory aims. *Science & Education, 22*(1), 69–91.

Bringmann, L. F., Elmer, T., & Eronen, M. I. (2022). Back to basics: The importance of conceptual clarification in psychological science. *Current Directions in Psychological Science, 31*(4), 340–346.

Bruch, M. (2015). *Beyond diagnosis: Case formulation in cognitive behavioural therapy*. John Wiley & Sons.

Casey, B., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S., & Ressler, K. J. (2013). DSM-5 and RDoC: Progress in psychiatry research? *Nature Reviews Neuroscience, 14*(11), 810.

Chang, H. (2017). Is pluralism compatible with scientific realism? In *The Routledge handbook of scientific realism* (pp. 176–186). Routledge.

Clack, S., & Ward, T. (2020). Modeling the symptoms of psychopathology: A pluralistic approach. *New Ideas in Psychology, 59*, 100799.

Contractor, A. A., Roley-Roberts, M. E., Lagdon, S., & Armour, C. (2017). Heterogeneity in patterns of DSM-5 posttraumatic stress disorder and depression symptoms: Latent profile analyses. *Journal of Affective Disorders, 212*, 17–24.

Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry, 13*(1), 28–35.

Cuthbert, B. N., & Insel, T. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine, 11*(1), 126.

Cuthbert, B. N., & Kozak, M. J. (2013). Constructing constructs for psychopathology: The NIMH research domain criteria. *Journal of Abnormal Psychology, 122*, 928–937.

Dallos, R. (2006). Integrative formulation: CAT and ANT. In *Formulation in psychology and psychotherapy* (pp. 199–224). Routledge.

Dallos, R., & Stedmon, J. (2013). Systemic formulation: Mapping the family dance. In *Formulation in psychology and psychotherapy* (pp. 87–115). Routledge.

de Haan, S. (2020b). *Enactive psychiatry*. Cambridge University Press.

Dickinson, D., Pratt, D. N., Giangrande, E. J., Grunnagle, M., Orel, J., Weinberger, D. R., Callicott, J. H., & Berman, K. F. (2017). Attacking heterogeneity in schizophrenia by deriving clinical subgroups from widely available symptom data. *Schizophrenia Bulletin, 44*(1), 101–113.

Donovan, C., & Murphy, D. (2020). De Haan on sense-making and psychopathology. *Philosophy, Psychiatry, & Psychology, 27*(1), 29–30.

Eells, T. D. (2015). *Psychotherapy case formulation*. American Psychological Association.

Fuchs, T. (2017). *Ecology of the Brain: The phenomenology and biology of the embodied mind*. Oxford University Press.

Fulford, K., & Jackson, M. (1997). Spiritual experience and psychopathology. *Philosophy, Psychiatry, & Psychology, 4*(1), 41–65.

Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 ways to have posttraumatic stress disorder. *Perspectives on Psychological Science, 8*(6), 651–662.

García Otero, E. (2022). *Participatory sense-making in psychotherapy*, PhD, University of the Basque Country/Universidad del País Vasco. http://hdl.handle.net/10810/56213

Gauld, C., Nielsen, K., Manon, J., Bottemanne, H., & Dumas, G. (2022). From analytic to synthetic-organizational pluralisms: A pluralistic enactive psychiatry. *Frontiers in Psychiatry, 13*. https://doi.org/10.3389/fpsyt.2022.981787

Glennan, S., & Illari, P. (2017). *The Routledge handbook of mechanisms and mechanical philosophy*. Taylor & Francis.

Greenberg, L. S. (2004). Emotion–focused therapy. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 11*(1), 3–16.

Haig, B. D. (2014). *Investigating the psychological world; scientific method in the behavioural sciences*. Massachusetts Institute of Technology.

Harper, D., & Spellman, D. (2006). Social constructionist formulation: Telling a different story. In *Formulation in psychology and psychotherapy* (pp. 115–142). Routledge.

Harris, R. (2009). *ACT made simple: A quick-start guide to ACT basics and beyond*. New Harbinger.

Hawkins-Elder, H., & Ward, T. (2020). The explanation of eating disorders: A critical analysis. *Behaviour Change, 37*(2), 93–110.

Hawkins-Elder, H., & Ward, T. (2021). From competition to co-operation: Shifting the "one best model" perspective. *Theory & Psychology, 31*, 821–841. https://doi.org/10.1177/0959354321995900

Hershenberg, R., & Goldfried, M. R. (2015). Implications of RDoC for the research and practice of psychotherapy. *Behavior Therapy, 46*(2), 156–165.

Hoffman, G. A., & Zachar, P. (2017). RDoC's metaphysical assumptions: Problems and promises. In *Extraordinary science: Responding to the crisis in psychiatric research* (pp. 59–86). MIT Press.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry, 167*, 748–751.

Insel, T., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science, 348*(6234), 499–500.

Jerotic, S., & Aftab, A. (2021). Scientific pluralism is the only way forward for psychiatry. *Acta Psychiatrica Scandinavica, 143*(6), 537–538.

Johnstone, L. (2018). Psychological formulation as an alternative to psychiatric diagnosis. *Journal of Humanistic Psychology, 58*(1), 30–46.

Johnstone, L., Boyle, M., Cromby, J., Dillon, J., Harper, D., & Kinderman, P. (2018). *The power threat meaning framework*. British Psychological Society.

Johnstone, L., & Dallos, R. (2013). Introduction to formulation. In *Formulation in psychology and psychotherapy* (pp. 21–37). Routledge.

Jones, E. G., & Mendell, L. M. (1999). Assessing the decade of the brain. *Science, 284*(5415), 739–739.

Karter, J. M., & Kamens, S. R. (2019). Toward conceptual competence in psychiatric diagnosis: An ecological model for critiques of the DSM. In *Critical Psychiatry* (pp. 17–69). Springer.

Kendler, K. (2008). Explanatory models for psychiatric illness. *American Journal of Psychiatry, 165*(6), 695–702.

Kendler, K. (2012b). The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. *Molecular Psychiatry, 17*(4), 377.

Kendler, K. (2019). From many to one to many – The search for causes of psychiatric illness. *JAMA Psychiatry, 76*(10), 1085–1091.

Kirmayer, L. J., & Crafa, D. (2014). What kind of science for psychiatry? *Frontiers in Human Neuroscience, 8*, 435.

Krueger, J., & Colombetti, G. (2018). Affective affordances and psychopathology. In *Philosophical perspectives on affective experience and psychopathology: Vol. XXVIII–2* (pp. 221–247). Quodlibet.

Larøi, F., Luhrmann, T. M., Bell, V., Christian, W. A., Jr., Deshpande, S., Fernyhough, C., Jenkins, J., & Woods, A. (2014). Culture and hallucinations: Overview and future directions. *Schizophrenia Bulletin, 40*(Suppl_4), S213–S220.

Leo, J., & Lacasse, J. R. (2008). The media and the chemical imbalance theory of depression. *Society, 45*(1), 35–45.

Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy, 62*, 129–139.

Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology, 12*, 435–463.

MacNeil, C. A., Hasty, M. K., Conus, P., & Berk, M. (2012). Is diagnosis enough to guide interventions in mental health? Using case formulation in clinical practice. *BMC Medicine, 10*(1), 1–3.

Malan, D., & Parker, L. (1995). *Individual psychotherapy and the science of psychodynamics*. CRC Press.

Miller, G. A. (2010). Mistreating psychology in the decades of the brain. *Perspectives on Psychological Science, 5*(6), 716–743.

Mitchell, S. D. (2002). Integrative pluralism. *Biology and Philosophy, 17*(1), 55–70.

Monroe, S. M., & Anderson, S. F. (2015). Depression: The shroud of heterogeneity. *Current Directions in Psychological Science, 24*(3), 227–231.

Morris, S. E., & Cuthbert, B. N. (2012). Research Domain Criteria: Cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience, 14*(1), 29.

NiaNia, W., Bush, A., & Epston, D. (2016). *Collaborative and indigenous mental health therapy: Tātaihono–stories of Māori healing and psychiatry*. Taylor & Francis.

Nielsen, K. (2020b). *What is mental disorder? Developing an embodied, embedded, and enactive psychopathology*. PhD thesis, Victoria University of Wellington. http://hdl.handle.net/10063/8957

Nielsen, K. (2022b). Same diagnosis, different problem: The challenge of heterogeneity in mental disorder. *MIND Foundation*. https://mind-foundation.org/same-diagnosis-different-problem-the-challenge-of-heterogeneity-in-mental-disorder/

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 8*(6), 800–822. https://doi.org/10.1177/0959354318808394

Nielsen, K., & Ward, T. (2020b). Phenomena complexes as targets of explanation in psychopathology: The Relational Analysis of Phenomena (RAP) approach. *Theory & Psychology, 30*(2), 164–185.

NIMH. (2022). *Developmental and environmental aspects*. https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/developmental-and-environmental-aspects

Papaspirou, P., & Moussas, X. (2013). A brief tour into the history of gravity: From Emocritus to Einstein. *American Journal of Space Science, 1*(1), 33–45.

Potochnik, A. (2016). Scientific explanation: Putting communication first. *Philosophy of Science, 83*(5), 721–732.

Potochnik, A. (2017). *Idealization and the aims of science*. University of Chicago Press.

Radomsky, A. S., Alcolado, G. M., Abramowitz, J. S., Alonso, P., Belloch, A., Bouvard, M., Clark, D. A., Coles, M. E., Doron, G., & Fernández-Álvarez, H. (2014). Part 1 – You can run but you can't hide: Intrusive thoughts on six continents. *Journal of Obsessive-Compulsive and Related Disorders, 3*(3), 269–279.

Seli, P., Risko, E. F., Purdon, C., & Smilek, D. (2017). Intrusive thoughts: Linking spontaneous mind wandering and OCD symptomatology. *Psychological Research, 81*(2), 392–398.

Sullivan, J. A. (2017). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations, 20*(2), 129–145.

Thagard, P. (2017). *Natural philosophy: From social brains to knowledge, reality, morality, and beauty* (draft 3).

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press. https://books.google.co.nz/books?id=OVGna4ZEpWwC

Vanderbeeken, R., & Weber, E. (2002). Dispositional explanations of behavior. *Behavior and Philosophy, 30*, 43–59.

Veit, W. (2020). Model pluralism. *Philosophy of the Social Sciences, 50*(2), 91–114.

Wakefield, J. C. (2014b). Wittgenstein's nightmare: Why the RDoC grid needs a conceptual dimension. *World Psychiatry, 13*(1), 38–40.

Ward, T., & Clack, S. (2019a). From symptom to clinical phenomena. *New Ideas in Psychology, 54*, 40–49.

Ward, T., & Clack, S. (2019b). From symptoms of psychopathology to the explanation of clinical phenomena. *New Ideas in Psychology, 54*, 40–49. https://doi.org/10.1016/j.newideapsych.2019.01.004

Ward, T., Clack, S., & Haig, B. D. (2016). The abductive theory of method: Scientific inquiry and clinical practice. *Behaviour Change, 33*(4), 212–231.

Ward, T., Vertue, F. M., & Haig, B. D. (1999). Abductive method and clinical assessment in practice. *Behaviour Change, 16*(1), 49–63.

Wegerhoff, D. (2022). *Understanding gangs: Developing an epistemically pluralist framework for gang research*.

Wegerhoff, D., Ward, T., & Dixon, L. (2020). A pluralistic approach to the definition, classification, and explanation of gangs. *Aggression and Violent Behavior, 58*, 101546.

Wegerhoff, D., Ward, T., & Dixon, L. (2022). Epistemic pluralism and the justification of conceptual strategies in science. *Theory & Psychology, 32*(3), 443–466.

Wilshire, C. E., Ward, T., & Clack, S. (2021). Symptom descriptions in psychopathology: How well are they working for us? *Clinical Psychological Science, 9*(3), 323–339.

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164*(4), 557–565.

Zachar, P., & Kendler, K. S. (2017). The philosophy of nosology. *Annual Review of Clinical Psychology, 13*, 49–71.

7

Summing Up and Moving Forward

In this concluding chapter I will briefly summarize the development of ideas across this book. This will firstly be achieved via a brief review of previous chapters, and secondly by way of a summary example comparing conceptualizations of disordered eating and the explanatory approaches they encourage or endorse. Finally, I highlight some limitations to 3e Psychopathology and draw the book to a close.

7.1 Looking Back

In Chap. 1 I introduced the idea that the study of psychopathology can be broken down into component tasks and situated the current project predominantly within the conceptual phase—right on the artificial border between naturally minded philosophy and conceptually minded science. I stated a wider commitment to pluralism and attempted to motivate the need for ongoing conceptual work within the study and treatment of psychopathology. I also highlighted a key observation that helped inspire this project—that peoples understanding of what mental

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7_7

disorder is appears to be deeply related to their wider understanding of human functioning and 'how the mind works'.

In Chap. 2 I overviewed a selection of extant conceptual models, breaking these down into those that make structural claims and those that propose a particular normative basis for using the label of disorder/dysfunction. We saw that all conceptual models had particular strengths, but so too did they fall down under closer scrutiny. We saw that across the models considered, understanding of dysfunction/disorder did indeed seem to be conceptually related to the underlying understanding of human functioning. This raised the question of whether assuming a well-suited framework of human functioning might facilitate the development of a novel and fruitful conceptual model of mental disorder. I also argued, following Thornton (2000), that assuming a value-inclusive but naturalistic framework of human functioning may allow us to move beyond the evaluative-objectivist divide.

Following on from these observations, Chap. 3 introduced 3e cognition or enactivism, a naturalistic yet non-reductionistic understanding of human functioning that includes an account of values, meaning, and normativity. I then reviewed three other frameworks that have recently attempted to consider the wider concept of mental disorder from the enactive viewpoint. I argued that Fuchs' (2017) notions of horizontal and vertical causality, dual aspectivity, and unfavorable patterns of behavior and experience offer a plausible sketch of mental disorders through a thoroughly embodied lens. Particular strengths were seen in his phenomenological observations, but key aspects of the conceptual model being proposed remain underspecified. I then argued that de Haan (2020) offers a more a more developed perspective. I suggested that de Haan's framework presents a perfectly good description of what mental disorders are, but provides little sense of how they might work (i.e., emerge and maintain) and how we should seek to investigate this. I also suggested that the normative basis for defining disorder de Haan provides is practical, but under justified, and that the inclusion of her 'common-sense' characteristic is problematic. Finally I reviewed Maiese's (2021) perspective. I argued that her perspective has similar strengths to de Haan's

structurally, but endorsed her more mechanistic focus on component habits and actions rather than on sense-making as a whole. I also supported Maiese's use of maladaption and 'dis-attunement' within the normative fabric of her model, highlighting its congruency with my own framework. However, I also shared concerns about the use of language such as 'identities' and 'selves' within her model and queried the reliance on social norms that do no functional work for the individual.

Starting in Chap. 4 I began to sketch out my own framework of 3e Psychopathology. I first outlined a set of theoretical/conceptual tools found within the enactive viewpoint that bring clear advantage for the consideration of mental disorder. These included notions of organizational causality, constitution, and dual aspectivity, the naturalized understanding of normativity, the thoroughgoing role of affectivity and culture in shaping human functioning, the developmental perspective it brings, and the way that it demands a plurality of perspectives. I then outlined a view of the structure of mental disorders as *recurrent patterns in sense-making supported by reafferent stable dynamic patterns of causal relations within the brain-body-environment system*. I also introduced the notion of a 'mode of functioning' in summary reference to the way that an individual makes sense of and survives in the world as a complex functioning of their embodiment and development. I then shifted to considering the normative basis of mental disorder under 3e Psychopathology. I argued that *mental disorder is a recurring pattern in sense-making that runs counter to the individual's functional norms to a significant or atypical degree, disrupting their wider mode of functioning in the world*. I tried to emphasize that this description holds functionality to be deeply contextual, and that the central question when evaluating if a pattern in sense-making counts as a disorder or not is: 'is this working for the person, or significantly working against their ability to survive, flexibly adapt to changing circumstance, and fare-well by their own standards?'

In Chap. 5 I took the conceptual bones of 3e Psychopathology presented in Chap. 4 and tried to stitch them together into a fuller conceptual model. I suggested that mental disorders under 3e Psychopathology can be seen as dysfunctional process-structures within the adaptive

striving of agents, distinguishable by their negative functional effects, inflexibility, and often complex and circular causal structures. I then attempted to riff on this idea by describing it though different perspectives, metaphors and examples. I explored how on this view we can sometimes view mental disorders as 'parasitic partial autonomies' interwoven with human functioning, how we can view them as analogous to significant backflows and stagnancies in the fluid dynamics of a river, and I applied the conceptual model to problems with anxiety. I then tried to take a more precise and contained approach, rating the 3e Psychopathology model on Zachar and Kendler's (2007) conceptual taxonomy. Some important questions related to the task of classification were then considered. I argued that 3e Psychopathology supports principles of classificatory pluralism and humility, a moderate anti-universalist stance, and the idea that alternative modes of functioning are not legitimate disorders.

Chapter 6 changed tack somewhat and focused in on implications of 3e Psychopathology for the task of explanation. The central question here was 'if this is what mental disorders are then how should we seek to explain them?' This question was explored at two different levels; explanation as it is engaged with by researchers, and explanation as it is engaged with by clinicians. I argued that 3e Psychopathology supports methodological and explanatory pluralism in both cases, but also offers gentle guidance and constraint. At a research level I briefly considered the DSM-ICD, the RDoC, and symptom-focused approaches, as ways to organize research findings before overviewing the RAP (Nielsen & Ward, 2020). I suggested that the RAP represents one way to go about developing 3e Psychopathology congruent explanatory theories regarding the maintenance of mental disorder. At a clinical level I reviewed some of the way that approaches to the task of formulation differ in the form of a methodological taxonomy. I suggested that 3e Psychopathology does not point directly to 'one best way' to formulate but explored some of the guidance that it offers. I argued that 3e Psychopathology supports a collaborative, pragmatic, and systemically minded approach to formulation that is mindful of the ways that assessment processes can enhance or stifle client agency. In an effort to demonstrate the real-world utility of conceptual approaches I also presented two 3e Psychopathology informed clinical resources toward the end of this chapter.

7.2 Disordered Eating as a Summary Example

Actual application of the ideas in this book to the development of an explanatory model is well beyond the constraints of the current project. However, as a summary exemplar only, it is useful to consider how the 3e approach here developed might reconfigure our conceptualization of disordered eating in comparison to other positions explored. Of particular interest is the relation between these conceptualizations and the kind of explanatory attempts we might take when holding these positions. A selection of conceptual positions is explored in this way within Table 7.1.

7.3 Limitations of the Current Project

There are limitations to the work done that bear considering as I draw this book to a close. These limitations concern appropriate use of 3e Psychopathology, how we will be able to falsify/evaluate it, and how applicable 3e Psychopathology is to the range of currently recognized mental disorders.

Appropriate Use

The work of this book is conceptual, and not explanatory. This was clearly stated at the outset but bears remembering. The concern here is that the framework developed may foster a sense of understanding when considering a particular mental disorder, even though we do not yet have quality explanations for it. Consider the example of anxiety given in Chap. 5. By offering a description of pathological anxiety that incorporates many of the known causal factors into a model of the agent-in-the-world, the description may spark a feeling of understanding, in that the pattern of anxiety no longer seems surprising when all of these potential causal factors are highlighted. While there is an interesting question to be asked about whether causally heterogeneous descriptions of mental disorders can do limited explanatory work—by highlighting *potential* causes and *ruling out* other causes, for discussion see Maung (2016)—we must

Table 7.1 How various conceptual positions relate to conceptual and explanatory approaches in the study of Anorexia Nervosa

Conceptual position	Conceptualization of anorexia	Congruent explanatory strategy
Biological Essentialism	Anorexia is a lesion/ difference in someone's brain/biology that produces a pattern of self-starvation. No necessary commitment to why this is a disorder (but often coupled with statistical or evolutionary functionalism).	Study the brain. Compare the brains of those who do and do not self-starve in an attempt to locate the lesion. Seek to understand how the lesion produces self-starvation. RDoC seems a viable approach.
Psychological Essentialism	Anorexia is a lesion/ difference in cognition that produces a pattern of self-starvation. No necessary commitment to why this is a disorder.	Study people's thinking. Compare the cognitive processes of those who do and do not self-starve in an attempt to locate the lesion. Seek to understand how the lesion produces self-starvation.
Socio-cultural Realism (Natural or Discrete Kind)	Anorexia is a distinct pattern of self-starvation caused by the pressures of society (e.g., media representations, the thin-ideal). No necessary commitment to why this is a disorder.	Study the social locations of those who self-starve. Compare to the social locations of those who don't. Infer the social pressures of relevance and seek to understand how they produce self-starvation.
Social Constructionism (Deflationary)	Anorexia is an unduly pathologizing label given by society to those that self-starve. Those captured by the label may not represent a meaningful group, may be expressing normal distress, may be responding to problems in society, or may not have a problem (e.g. pro-ana).	There is potentially nothing to explain. Instead we need to question the institutions that are pathologizing people, and offer non-medical support for those in distress.

(*continued*)

Table 7.1 (continued)

Conceptual position	Conceptualization of anorexia	Congruent explanatory strategy
Fuzzy MPC Kind	Anorexia is a pattern of self-starvation behavior brought about and maintained by a fuzzy network of causal mechanisms, potentially spanning the brain, body, and environment. No necessary commitment to why this is a disorder.	Multiple approaches needed (i.e., methodological and explanatory pluralism). Intention is to identify and understand the causal mechanisms supporting the pattern of behavior. The RAP seems a viable approach.
Circular Causes and Dual Aspectivity (Fuchs, 2017)	Anorexia is a disruption to normal vertical circular causality (i.e., the integrated functioning of the body and experience) and horizontal circular causality (i.e., the functioning and adaptive agency of the person in the world), supporting self-starvation behavior. This is a disorder because it negatively affects flexible attunement and agency.	A 'poly-perspectival' approach, with an emphasis on experientially focused methodologies. Biological features should not be considered causes because they are only one part of the wider pattern.
Enactive Psychiatry (de Haan, 2020)	Anorexia is a bias in sense-making resulting in self-starvation. It has physiological, experiential, socio-cultural, and existential aspects. Anorexia is a disorder because, variously: others see it as inappropriate or against common-sense, it results in inflexible thought and action, it results in suffering.	Study the physiological, experiential, socio-cultural, and existential aspects of the pattern. Mechanisms/causes identified do not *underlie* the bias but are part of what constitute it. Personalized network models and the modelling of developmental trajectories are endorsed.

(*continued*)

Table 7.1 (continued)

Conceptual position	Conceptualization of anorexia	Congruent explanatory strategy
Enactive Medical Model (Maiese, 2021)	Anorexia is a disruption to sense-making resulting in self-starvation. It has physiological, experiential, socio-cultural, and existential aspects. Anorexia is a disorder because it is maladaptive and gets in the way of the enaction of a person's various regional identities.	Explanatory approach is currently unspecified but appears to endorce a simular pluralistic approach to other enactive authors.
3e Psychopathology	Anorexia is a complex pattern in the way that a person makes sense of and engages with the world, maintaining self-starvation behavior. This pattern is constituted by and can be modeled as a network of causal mechanisms spanning brain, body, experience, and environment. This network/pattern is understood to be fuzzy, i.e., differing across people. Anorexia is a disorder because the person is acting against their own ability to survive, flexibly adapt to changing contexts, and to flourish by their own standards.	Multiple approaches and perspectives needed (i.e., methodological and explanatory pluralism). Study the pattern across biology, experience, and context. Intention is to identify and understand the network of causal mechanisms supporting dysfunctional sense-making—i.e., its dynamical constitution. Mechanisms do not *underlie* the pattern but are part of what constitute it. The RAP is proposed as one viable approach to developing theories of maintenance. A pluralistic, collaborative, and process-conscious approach to formulation is endorsed.

remember that the 'causes' mentioned are only potential, insufficiently evidenced, and will certainly not apply in every token case of anxiety. Explicit explanatory models which postulate, on the basis of good evidence, how exactly constituent phenomena within particular disorders

are related is a key *next* step in the development of 3e Psychopathology, as outlined in the RAP. The claim made in this book is that the conceptual framework here developed will be helpful for the development of explanatory models and clinical methodologies, and for guiding the general stance that people—both lay and professional—take towards those with mental disorders. The claim is not that the conceptual framework itself or examples given perform any meaningful explanatory work.

Falsifiability/Explanatory Value

Given the conceptual nature of the work, this book essentially makes no empirically testable claims. This is potentially problematic because good science is falsifiable science. Thought must therefore be given to what would it mean for the framework developed in this book to be incorrect or useless. The answer to this problem is that, while the framework is not *directly* testable, it is *indirectly* falsifiable and open to evaluation through its products, i.e., the explanatory models and clinical approaches it fosters. If the framework helps produce valuable explanatory models of disorder and/or components thereof (i.e., explanatory models that stand up to empirical testing, that are parsimonious, that point to successful treatments), or fosters the development of useful and effective clinical approaches, then this will constitute reason to believe that the wider framework is mapping on to reality in a useful way. If the framework instead facilitates the production of explanatory models that consistently fail to generate accurate predictions, or encourages clinical approaches that don't seem useful or effective, then this will constitute reason to revise or disregard 3e Psychopathology.

Applicability

The framework here developed presents a particular understanding of what mental disorder is. While it is argued that this perspective can provide an interesting and useful way to think about all mental disorders, I do not want to make the dogmatic claim that this will be the *best* perspective to take for *all* mental disorders currently recognized. Rather, the way

of thinking about mental disorder developed in this book may turn out to be more useful for some conditions (and purposes) than for others. For example, consider neuropsychological conditions such as Attention Deficit Hyper-Activity Disorder (ADHD). Perhaps ADHD is best seen as a difference in people's behavioral and attentional processes supported by a network of causal factors that span brain, body, and environment, all across development. On the other hand, if I am a neuroscientist trying to understand how the brains of those that experience ADHD differ from others, then it may actually be more useful to me if I approach ADHD as simply a brain disorder. Utilizing a simpler conceptual model may allow me to ignore environmental and developmental factors as extraneous and doing so may well facilitate my discovery of reliable differences (i.e., lesions) in the brains of those with ADHD or subtypes thereof. This would represent very useful information in our understanding of the condition. A 3e Psychopathology view meanwhile, does not seem as helpful in the same way. It would likely to encourage us to view ADHD not as a mental disorder in its sense of the world, but as a neurodevelopmental difference and alternative mode of functioning which brings with it certain functional challenges and vulnerabilities to other difficulties. I would argue that this is a useful and compassionate perspective to hold, but it is not as clearly useful to those trying to understand the neurobiology of the condition. While the framework developed in this book seems likely to be useful across the study of most mental disorders, this does not mean that the framework will be the best option for all disorders and investigatory tasks. As stated in Chap. 1 I am supportive of a conceptual pluralism beyond the bounds of 3e Psychopathology.

7.4 Returning to Our Starting Questions

At the closing it is interesting to revisit the three general questions I listed in Chap. 1.

Are mental disorders something you get or something you do? In highlighting the agential nature of disorder, the 3e view developed here sides with the idea that mental disorders are something people do. This of course not to suggest that mental disorders are entirely volitional, only

that they concern a person acting in the world, the functionality of this action, and the flexibility with which action is altered when it is not serving the self-maintenance, adaption, and faring well of the organism within its environment.

Are mental disorders defined by brute facts or by social norms and values? On the developed view mental disorders are factual, yet still normative in a functional sense. Deciding whether someone's behavior constitutes a mental disorder concerns a thoroughgoing evaluation of the behavior—of whether it is working for them.

Finally, does a mental disorder exist inside someone's brain or is it dispersed across their brain, body and environment? On the 3e view the answer to this question is complex. Structurally mental disorders are complex networks of causal factors interacting within the brain-body-environment system, holding a person stuck in unhelpful ways of making sense-of and engaging with the world. These structures are then defined as mental disorders on the basis of the dysfunctional nature of the relation between the organism's environment and the organism's pattern of behavior over time.

7.5 Conclusion

This book took as its central question 'what exactly is mental disorder?'. It was proposed that answers to this question depend on our fundamental assumptions about human functioning. As a set of assumptions seemingly fit for purpose, the position of enactivism was explored and the nature of mental disorder from this perspective considered. The 3e Psychopathology approach here developed allows for the convergence of psychological, neuroscientific, and phenomenological perspectives around a central conception of mental disorder without prejudice, and encourages deep normative consideration whenever a label of disorder or dysfunction. The view presented: moves beyond the internalist bias of many current conceptual models, defines an ethically and scientifically justifiable role for normativity within the nature of disorder, encourages consideration of cultural and individual variance, endorses a gently

constrained pluralism in our attempts at explanatory theory and clinical formulation, and can sit comfortably within a wholly natural worldview.

Under 3e Psychopathology, mental disorders are understood as recurring patterns in the way that someone makes sense of and engages with the world that work against that person's ability to survive, flexibly adapt to changing circumstance, and to fare-well by their own embodied standards. These patterns are understood to be supported by a complex network of causal factors across the person's brain-body-environment system, holding them stuck in behaviors and ways of thinking and feeling that do not work for them. This represents one plausible alternative answer to the question 'what is mental disorder?' If the sciences of psychopathology and its associated therapeutic approaches are to progress, we need to keep asking this question and refining our answers.

References

de Haan, S. (2020). *Enactive psychiatry*. Cambridge University Press.

Fuchs, T. (2017). *Ecology of the Brain: The phenomenology and biology of the embodied mind*. Oxford University Press.

Maiese, M. (2021). An enactivist reconceptualization of the medical model. *Philosophical Psychology, 34*, 1–27.

Maung, H. H. (2016). Diagnosis and causal explanation in psychiatry. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 60*, 15–24.

Nielsen, K., & Ward, T. (2020). Phenomena complexes as targets of explanation in psychopathology: The Relational Analysis of Phenomena (RAP) approach. *Theory & Psychology, 30*(2), 164–185.

Thornton, T. (2000). Mental illness and reductionism: Can functions be naturalized? *Philosophy, Psychiatry, & Psychology, 7*(1), 67–76.

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164*(4), 557–565.

References

Aftab, A., & Nielsen, K. (2020). *3E approach to psychopathology: Kristopher Nielsen, PhD* [Interview in Psychiatric Times]. https://www.psychiatrictimes.com/view/three-approach-psychopathology-kristopher-nielsen-phd

Aftab, A., & Nielsen, K. (2021). From Engel to Enactivism: Contextualizing the biopsychosocial model. *European Journal of Analytic Philosophy, 17*(2), M2–M22.

Albert, P. R., Benkelfat, C., & Descarries, L. (2012). *The neurobiology of depression – Revisiting the serotonin hypothesis. I. Cellular and molecular mechanisms*.

American Psychiatric Association. (2013a). Anxiety disorders. In *Diagnostic and statistical manual of mental disorders*. https://doi.org/10.1176/appi.books.9780890425596.dsm05

American Psychiatric Association. (2013b). *Diagnostic and statistical manual of mental disorders, 5th edn (DSM-5)* (5th ed.).

Andersen, H. (2014a). A field guide to mechanisms: Part I. *Philosophy Compass, 9*(4), 274–283.

Andersen, H. (2014b). A field guide to mechanisms: Part II. *Philosophy Compass, 9*(4), 284–293.

Andersen, H. (2016). Biomedical sciences. *The Routledge Companion to Philosophy of Medicine, 81*.

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7

Andrews, G., Slade, T., & Issakidis, C. (2002). Deconstructing current comorbidity: Data from the Australian National Survey of Mental Health and Wellbeing. *The British Journal of Psychiatry, 181*(4), 306–314.

Baier, A. L., Kline, A. C., & Feeny, N. C. (2020). Therapeutic alliance as a mediator of change: A systematic review and evaluation of research. *Clinical Psychology Review, 82*, 101921.

Banner, N. F. (2013). Mental disorders are not brain disorders. *Journal of Evaluation in Clinical Practice, 19*(3), 509–513.

Barandiaran, X. E. (2017). Autonomy and enactivism: Towards a theory of sensorimotor autonomous agency. *Topoi, 36*(3), 409–430.

Bechtel, W. (2009a). Explanation: Mechanism, modularity, and situated cognition. *The Cambridge Handbook of Situated Cognition*, 155–170.

Bechtel, W. (2009b). Looking down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology, 22*(5), 543–564.

Beck, A. T., & Bredemeier, K. (2016). A unified model of depression: Integrating clinical, cognitive, biological, and evolutionary perspectives. *Clinical Psychological Science, 4*(4), 596–619.

Beck, J. S. (2020). *Cognitive behavior therapy: Basics and beyond*. Guilford Publications.

Berenbaum, H. (2013). Classification and psychopathology research. *Journal of Abnormal Psychology, 122*(3), 894.

Bergner, R. M. (1997). What is psychopathology? And so what? *Clinical Psychology: Science and Practice, 4*(3), 235–248.

Bergner, R. M. (2004). An integrative framework for psychopathology and psychotherapy. *New Ideas in Psychology, 22*(2), 127–141.

Bergner, R. M., & Bunford, N. (2017). Mental disorder is a disability concept, not a behavioral one. *Philosophy, Psychiatry, & Psychology, 24*(1), 25–40.

Bingham, R., & Banner, N. (2014). The definition of mental disorder: Evolving but dysfunctional? *Journal of Medical Ethics, 40*(8), 537–542.

Bird, A., & Tobin, E. (2018). *Natural kinds*. The Stanford Encyclopedia of Philosophy.

Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review, 97*(3), 303–352.

Bokulich, A. (2018). Representing and explaining: The eikonic conception of scientific explanation. *Philosophy of Science, 85*(5), 793–805.

Bolton, D., & Gillett, G. (2019). *The biopsychosocial model of health and disease: New philosophical and scientific developments*. Springer Nature.

Boorse, C. (1975). On the distinction between disease and illness. *Philosophy & Public Affairs, 5*, 49–68.

Boorse, C. (1977). Health as a theoretical concept. *Philosophy of Science, 44*(4), 542–573.

Boorse, C. (2014). A second rebuttal on health. *Journal of Medicine and Philosophy, 39*(6), 683–724.

Borrell-Carrió, F., Suchman, A. L., & Epstein, R. M. (2004). The biopsychosocial model 25 years later: Principles, practice, and scientific inquiry. *The Annals of Family Medicine, 2*(6), 576–582.

Borsboom, D., Cramer, A., & Kalis, A. (2018). Brain disorders? Not really… Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences, 42*, 1–54.

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies, 61*(1–2), 127–148.

Brigandt, I. (2013). Explanation in biology: Reduction, pluralism, and explanatory aims. *Science & Education, 22*(1), 69–91.

Bringmann, L. F., Elmer, T., & Eronen, M. I. (2022). Back to basics: The importance of conceptual clarification in psychological science. *Current Directions in Psychological Science, 31*(4), 340–346.

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review, 125*(4), 606.

Bruch, M. (2015). *Beyond diagnosis: Case formulation in cognitive behavioural therapy*. John Wiley & Sons.

Casey, B., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S., & Ressler, K. J. (2013). DSM-5 and RDoC: Progress in psychiatry research? *Nature Reviews Neuroscience, 14*(11), 810.

Chang, H. (2017). Is pluralism compatible with scientific realism? In *The Routledge handbook of scientific realism* (pp. 176–186). Routledge.

Chang, H. (2020). Pragmatism, perspectivism, and the historicity of science. In M. Massimi & C. D. McCoy (Eds.), *Understanding perspectivism* (pp. 10–27). Taylor & Francis.

Chapman, R. (2021). Neurodiversity and the social ecology of mental functions. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620959833

Christensen, W. D. (2012). Natural sources of normativity. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43*(1), 104–112.

Clack, S., & Ward, T. (2020). Modeling the symptoms of psychopathology: A pluralistic approach. *New Ideas in Psychology, 59*, 100799.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19.

Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest, 18*(2), 72–145.

Colombetti, G. (2014). *The feeling body*. The MIT Press.

Contractor, A. A., Roley-Roberts, M. E., Lagdon, S., & Armour, C. (2017). Heterogeneity in patterns of DSM-5 posttraumatic stress disorder and depression symptoms: Latent profile analyses. *Journal of Affective Disorders, 212*, 17–24.

Cooper, R. (2013). Avoiding false positives: Zones of rarity, the threshold problem, and the DSM clinical significance criterion. *The Canadian Journal of Psychiatry, 58*(11), 606–611.

Cramer, A. O., Waldorp, L. J., van der Maas, H. L., & Borsboom, D. (2010). Complex realities require complex theories: Refining and extending the network approach to mental disorders. *Behavioral and Brain Sciences, 33*(2–3), 178–193.

Craver, C., & Kaplan, D. M. (2018). Are more details better? On the norms of completeness for mechanistic explanations. *The British Journal for the Philosophy of Science, 1*, 287–319.

Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry, 13*(1), 28–35.

Cuthbert, B. N., & Insel, T. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine, 11*(1), 126.

Cuthbert, B. N., & Kozak, M. J. (2013). Constructing constructs for psychopathology: The NIMH research domain criteria. *Journal of Abnormal Psychology, 122*, 928–937.

Dallos, R. (2006). Integrative formulation: CAT and ANT. In *Formulation in psychology and psychotherapy* (pp. 199–224). Routledge.

Dallos, R., & Stedmon, J. (2013). Systemic formulation: Mapping the family dance. In *Formulation in psychology and psychotherapy* (pp. 87–115). Routledge.

de Haan, S. (2017). The existential dimension in psychiatry: An enactive framework. *Mental Health, Religion & Culture, 20*(6), 528–535.

de Haan, S. (2020a). An enactive approach to psychiatry. *Philosophy, Psychiatry and Psychology, 27*(1), 3–25.

de Haan, S. (2020b). *Enactive psychiatry*. Cambridge University Press.

de Haan, S. (2021). Two enactive approaches to psychiatry: Two contrasting views on what it means to be human. *Philosophy, Psychiatry, & Psychology, 28*(3), 191–196.

de Haan, S., & Fuchs, T. (2010). The ghost in the machine: Disembodiment in schizophrenia–two case studies. *Psychopathology, 43*(5), 327–333.

De Haan, S., Rietveld, E., Stokhof, M., & Denys, D. (2013). The phenomenology of deep brain stimulation-induced changes in OCD: An enactive affordance-based model. *Frontiers in Human Neuroscience, 7*, 653.

De Jaegher, H. (2013). Embodiment and sense-making in autism. *Frontiers in Integrative Neuroscience, 7*.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences, 4*(4), 429–452. https://doi.org/10.1007/s11097-005-9002-y

Di Paolo, E. (2010). Overcoming autopoiesis: An enactive detour on the way from life to society. In *Advanced series in management*. Emerald Group Publishing Limited.

Di Paolo, E., Cuffari, E. C., & De Jaegher, H. (2018). *Linguistic bodies: The continuity between life and language*. MIT Press.

Dickinson, D., Pratt, D. N., Giangrande, E. J., Grunnagle, M., Orel, J., Weinberger, D. R., Callicott, J. H., & Berman, K. F. (2017). Attacking heterogeneity in schizophrenia by deriving clinical subgroups from widely available symptom data. *Schizophrenia Bulletin, 44*(1), 101–113.

Donovan, C., & Murphy, D. (2020). De Haan on sense-making and psychopathology. *Philosophy, Psychiatry, & Psychology, 27*(1), 29–30.

Doust, J., Walker, M. J., & Rogers, W. A. (2017). Current dilemmas in defining the boundaries of disease. *Journal of Medicine and Philosophy, 42*(4), 350–366.

Durt, C., Fuchs, T., & Tewes, C. (2017). *Embodiment, enaction, and culture: Investigating the constitution of the shared world*. MIT Press. https://books.google.co.nz/books?id=OJakDgAAQBAJ

Eells, T. D. (2015). *Psychotherapy case formulation*. American Psychological Association.

Elbau, I. G., Binder, E. B., & Spoormaker, V. I. (2019). Symptoms are not the solution but the problem: Why psychiatric research should focus on processes rather than symptoms. *Behavioral and Brain Sciences, 42*, E7.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science, 196*(4286), 129–136.

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network pschometrics: Combining network and latent variable models. *Psychometrika, 82*(4), 904–927.

Foucault, M. (2003). *Madness and civilization*. Routledge.

Fried, E. I., & Cramer, A. O. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science, 12*(6), 999–1020.

Fried, E. I., van Borkulo, C. D., Cramer, A. O., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology, 52*(1), 1–10.

Fuchs, T. (2017). *Ecology of the Brain: The phenomenology and biology of the embodied mind*. Oxford University Press.

Fuchs, T. (2022). Understanding as explaining: How motives can become causes. *Phenomenology and the Cognitive Sciences*, 1–16.

Fuchs, T., & Röhricht, F. (2017). Schizophrenia and intersubjectivity: An embodied and enactive approach to psychopathology and psychotherapy. *Philosophy, Psychiatry, & Psychology, 24*(2), 127–142.

Fulford, K. (2001). 'What is (mental) disease?': An open letter to Christopher Boorse. *Journal of Medical Ethics, 27*(2), 80–85.

Fulford, K. (2002). Values in psychiatric diagnosis: Executive summary of a report to the chair of the ICD-12/DSM-VI Coordination Task Force (Dateline 2010). *Psychopathology, 35*(2–3), 132–138.

Fulford, K., & Colombo, A. (2004). Six models of mental disorder: A study combining linguistic-analytic and empirical methods. *Philosophy, Psychiatry, & Psychology, 11*(2), 129–144.

Fulford, K., Davies, M., Gipps, R., Graham, G., Sadler, J., Stanghellini, G., & Thornton, T. (2013). *The Oxford handbook of philosophy and psychiatry*. OUP Oxford.

Fulford, K., & Jackson, M. (1997). Spiritual experience and psychopathology. *Philosophy, Psychiatry, & Psychology, 4*(1), 41–65.

Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 ways to have posttraumatic stress disorder. *Perspectives on Psychological Science, 8*(6), 651–662.

Gallagher, S. (2006). *How the body shapes the mind*. Clarendon Press.

Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford University Press. https://books.google.co.nz/books?id=Z28sDwAAQBAJ

Gallagher, S., & Varga, S. (2015). Conceptual issues in autism spectrum disorders. *Current Opinion in Psychiatry, 28*(2), 127–132.

García Otero, E. (2022). *Participatory sense-making in psychotherapy*, PhD, University of the Basque Country/Universidad del País Vasco. http://hdl.handle.net/10810/56213

Gardner, A., & Boles, R. G. (2011). Beyond the serotonin hypothesis: Mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(3), 730–743.

Garson, J. (2017). Mechanisms, phenomena, and functions. In *The Routledge handbook of mechanisms and mechanical philosophy* (pp. 122–133). Routledge.

Gauld, C., Nielsen, K., Manon, J., Bottemanne, H., & Dumas, G. (2022). From analytic to synthetic-organizational pluralisms: A pluralistic enactive psychiatry. *Frontiers in Psychiatry, 13*. https://doi.org/10.3389/fpsyt.2022.981787

Ghaemi, S. N. (2009). The rise and fall of the biopsychosocial model. *The British Journal of Psychiatry, 195*(1), 3–4.

Gibson, J. (1977). The concept of affordances. *Perceiving, Acting, and Knowing*, *1*.

Gleaves, D. H. (1996). The sociocognitive model of dissociative identity disorder: A reexamination of the evidence. *Psychological Bulletin, 120*(1), 42.

Glennan, S., & Illari, P. (2017). *The Routledge handbook of mechanisms and mechanical philosophy*. Taylor & Francis.

Graham, G. (2013). *The disordered mind: An introduction to philosophy of mind and mental illness*. Routledge.

Greenberg, L. S. (2004). Emotion–focused therapy. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 11*(1), 3–16.

Haig, B. D. (2014). *Investigating the psychological world; scientific method in the behavioural sciences*. Massachusetts Institute of Technology.

Haig, B. D., & Vertue, F. M. (2010). Extending the network perspective on comorbidity. *Behavioral and Brain Sciences, 33*(2–3), 158–158.

Harper, D., & Spellman, D. (2006). Social constructionist formulation: Telling a different story. In *Formulation in psychology and psychotherapy* (pp. 115–142). Routledge.

Harris, R. (2009). *ACT made simple: A quick-start guide to ACT basics and beyond*. New Harbinger.

Hartner, D. F., & Theurer, K. L. (2018). Psychiatry should not seek mechanisms of disorder. *Journal of Theoretical and Philosophical Psychology, 38*, 189–204.

Harvey, M. I. (2015). Content in languaging: Why radical enactivism is incompatible with representational theories of language. *Language Sciences, 48*, 90–129. https://doi.org/10.1016/j.langsci.2014.12.004

Haslam, N. (2002). Kinds of kinds: A conceptual taxonomy of psychiatric categories. *Philosophy, Psychiatry, & Psychology, 9*(3), 203–217.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry, 27*(1), 1–17.

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine, 42*(5), 903–920.

Hawkins-Elder, H., & Ward, T. (2019). Theory construction in the psychopathology domain: A multi-phase approach. *Theory & Psychology*.

Hawkins-Elder, H., & Ward, T. (2020). The explanation of eating disorders: A critical analysis. *Behaviour Change, 37*(2), 93–110.

Hawkins-Elder, H., & Ward, T. (2021). From competition to co-operation: Shifting the "one best model" perspective. *Theory & Psychology, 31*, 821–841. https://doi.org/10.1177/0959354321995900

Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Hershenberg, R., & Goldfried, M. R. (2015). Implications of RDoC for the research and practice of psychotherapy. *Behavior Therapy, 46*(2), 156–165.

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.

Hochstein, E. (2019). How metaphysical commitments shape the study of psychological mechanisms. *Theory & Psychology, 29*(5), 579–600. https://doi.org/10.1177/0959354319860591

Hoffman, G. A., & Zachar, P. (2017). RDoC's metaphysical assumptions: Problems and promises. In *Extraordinary science: Responding to the crisis in psychiatric research* (pp. 59–86). MIT Press.

Hucklenbroich, P. (2014). Medical criteria of pathologicity and their role in scientific psychiatry – Comments on the articles of Henrik Walter and Marco Stier. *Frontiers in Psychology, 5*, 128.

Hume, D. (1978). A treatise of human nature [1739]. *British Moralists*, 1650–1800.

Humphry, S. M., & McGrane, J. A. (2010). Is there a contradiction between the network and latent variable perspectives? *Behavioral and Brain Sciences, 33*(2–3), 160–161.

Hutto, D. D., & Myin, E. (2012). *Radicalizing enactivism: Basic minds without content*. MIT Press.

Hutto, D. D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. MIT Press.

Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology, 6*, 155–179.

Illari, P., & Glennan, S. (2017). Varieties of mechanisms. In *The Routledge handbook of mechanisms and mechanical philosophy* (pp. 109–121). Routledge.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry, 167*, 748–751.

Insel, T., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science, 348*(6234), 499–500.

Jefferson, A. (2014). Mental disorders, brain disorders and values. *Frontiers in Psychology, 5*, 130.

Jerotic, S., & Aftab, A. (2021). Scientific pluralism is the only way forward for psychiatry. *Acta Psychiatrica Scandinavica, 143*(6), 537–538.

Jilek, W., & Mattelaer, J. (2006). Koro: The psychological disappearance of the penis. *De Historia Urologiae Europaeae, 13*, 53–73.

Johnstone, L. (2018). Psychological formulation as an alternative to psychiatric diagnosis. *Journal of Humanistic Psychology, 58*(1), 30–46.

Johnstone, L., Boyle, M., Cromby, J., Dillon, J., Harper, D., & Kinderman, P. (2018). *The power threat meaning framework*. British Psychological Society.

Johnstone, L., & Dallos, R. (2013). Introduction to formulation. In *Formulation in psychology and psychotherapy* (pp. 21–37). Routledge.

Jones, E. G., & Mendell, L. M. (1999). Assessing the decade of the brain. *Science, 284*(5415), 739–739.

Kaplan, D. M. (2015). Moving parts: The natural alliance between dynamical and mechanistic modeling approaches. *Biology & Philosophy, 30*(6), 757–786.

Karter, J. M., & Kamens, S. R. (2019). Toward conceptual competence in psychiatric diagnosis: An ecological model for critiques of the DSM. In *Critical Psychiatry* (pp. 17–69). Springer.

Kendall, P. C., Cummings, C. M., Villabø, M. A., Narayanan, M. K., Treadwell, K., Birmaher, B., Compton, S., Piacentini, J., Sherrill, J., & Walkup, J. (2016). Mediators of change in the Child/adolescent anxiety multimodal treatment study. *Journal of Consulting and Clinical Psychology, 84*(1), 1.

Kendler, K. (2008). Explanatory models for psychiatric illness. *American Journal of Psychiatry, 165*(6), 695–702.

Kendler, K. (2012a). Levels of explanation in psychiatric and substance use disorders: Implications for the development of an etiologically based nosology. *Molecular Psychiatry, 17*(1), 11.

Kendler, K. (2012b). The dappled nature of causes of psychiatric illness: Replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. *Molecular Psychiatry, 17*(4), 377.

Kendler, K. (2016). The nature of psychiatric disorders. *World Psychiatry, 15*(1), 5–12.

Kendler, K. (2019). From many to one to many – The search for causes of psychiatric illness. *JAMA Psychiatry, 76*(10), 1085–1091.

Kendler, K., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*(6), 1143–1150.

Khalidi, M. A. (2013). *Natural categories and human kinds: Classification in the natural and social sciences*. Cambridge University Press.

Kingma, E. (2007). What is it to be healthy? *Analysis, 67*(2), 128–133.

Kirmayer, L. J., & Crafa, D. (2014). What kind of science for psychiatry? *Frontiers in Human Neuroscience, 8*, 435.

Kirmayer, L. J., & Jarvis, G. E. (2019). Culturally responsive services as a path to equity in mental healthcare. *HealthcarePapers, 18*(2), 11–23.

Kirmayer, L. J., & Ramstead, M. J. (2017). Embodiment and enactment in cultural psychiatry. In C. Durt, T. Fuchs, & C. Tewes (Eds.), *Embodiment, enaction, and culture: Investigating the constitution of the shared world* (p. 397). MIT Press.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., & Clark, L. A. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877–883.

Krueger, J., & Colombetti, G. (2018). Affective affordances and psychopathology. In *Philosophical perspectives on affective experience and psychopathology: Vol. XXVIII–2* (pp. 221–247). Quodlibet.

Kvaale, E. P., Haslam, N., & Gottdiener, W. H. (2013). The 'side effects' of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review, 33*(6), 782–794.

Larøi, F., Luhrmann, T. M., Bell, V., Christian, W. A., Jr., Deshpande, S., Fernyhough, C., Jenkins, J., & Woods, A. (2014). Culture and hallucinations: Overview and future directions. *Schizophrenia Bulletin, 40*(Suppl_4), S213–S220.

Lebowitz, M. S., & Appelbaum, P. S. (2019). Biomedical explanations of psychopathology and their implications for attitudes and beliefs about mental disorders. *Annual Review of Clinical Psychology, 15*, 555–577.

Leo, J., & Lacasse, J. R. (2008). The media and the chemical imbalance theory of depression. *Society, 45*(1), 35–45.

Lewis-Fernández, R., Guarnaccia, P. J., Martínez, I. E., Salmán, E., Schmidt, A., & Liebowitz, M. (2002). Comparative phenomenology of ataques de nervios, panic attacks, and panic disorder. *Culture, Medicine and Psychiatry, 26*(2), 199–223.

Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy, 62*, 129–139.

Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology, 104*(3), 411–420.

Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology, 12*, 435–463.

MacNeil, C. A., Hasty, M. K., Conus, P., & Berk, M. (2012). Is diagnosis enough to guide interventions in mental health? Using case formulation in clinical practice. *BMC Medicine, 10*(1), 1–3.

Magnus, P. (2012). *Scientific enquiry and natural kinds: From planets to mallards*. Springer.

Magnus, P. (2014a). Epistemic categories and causal kinds. *Philosophy Faculty Scholarship*. https://doi.org/10.1016/j.shpsc.2014.10.001

Magnus, P. (2014b). NK≠ HPC. *The Philosophical Quarterly, 64*(256), 471–477.

Maiese, M. (2016). *Embodied selves and divided minds*. Oxford University Press. https://books.google.co.nz/books?id=w_quCgAAQBAJ

Maiese, M. (2017). Can the mind be embodied, enactive, affective, and extended? *Phenomenology and the Cognitive Sciences*, 1–19.

Maiese, M. (2021). An enactivist reconceptualization of the medical model. *Philosophical Psychology, 34*, 1–27.

Malan, D., & Parker, L. (1995). *Individual psychotherapy and the science of psychodynamics*. CRC Press.

Mallon, R. (2016). *The construction of human kinds*. Oxford University Press.

Markon, K. E. (2013). Epistemological pluralism and scientific development: An argument against authoritative nosologies. *Journal of Personality Disorders, 27*(5), 554–579.

Massimi, M. (2021). *Perspectival realism*. Oxford University Press.

Massimi, M., & McCoy, C. D. (2020). *Understanding perspectivism: Scientific challenges and methodological prospects*. Taylor & Francis.

Maung, H. H. (2016). Diagnosis and causal explanation in psychiatry. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 60*, 15–24.

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy, 86*, 95–104.

Megone, C. (1998). Aristotle's function argument and the concept of mental illness. *Philosophy, Psychiatry, & Psychology, 5*(3), 187–201.

Miller, G. A. (2010). Mistreating psychology in the decades of the brain. *Perspectives on Psychological Science, 5*(6), 716–743.

Mitchell, S. D. (2002). Integrative pluralism. *Biology and Philosophy, 17*(1), 55–70.

Molenaar, P. C. (2010). Latent variable models are network models. *Behavioral and Brain Sciences, 33*(2–3), 166–166.

Monroe, S. M., & Anderson, S. F. (2015). Depression: The shroud of heterogeneity. *Current Directions in Psychological Science, 24*(3), 227–231.

Morris, S. E., & Cuthbert, B. N. (2012). Research Domain Criteria: Cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience, 14*(1), 29.

Muders, S. (2014). On the concept of the normative in the assessment of mental disorder. *Frontiers in Psychology, 5*, 129.

Murphy, D. (2017). Can psychiatry refurbish the mind? *Philosophical Explorations, 20*(2), 160–174.

Murphy, D., & Woolfolk, R. L. (2000). The harmful dysfunction analysis of mental disorder. *Philosophy, Psychiatry, & Psychology, 7*(4), 241–252.

Nesse, R. M. (2001). On the difficulty of defining disease: A Darwinian perspective. *Medicine, Health Care and Philosophy, 4*(1), 37–46.

NiaNia, W., Bush, A., & Epston, D. (2016). *Collaborative and indigenous mental health therapy: Tātaihono–stories of Māori healing and psychiatry*. Taylor & Francis.

Nielsen, K. (2020a). Think of mental disorders as the mind's 'sticky tendencies.' *Aeon*. https://aeon.co/ideas/think-of-mental-disorders-as-the-minds-sticky-tendencies

Nielsen, K. (2020b). *What is mental disorder? Developing an embodied, embedded, and enactive psychopathology*. PhD thesis, Victoria University of Wellington. http://hdl.handle.net/10063/8957

Nielsen, K. (2021a). Comparing two enactive perspectives. *Philosophy, Psychiatry, & Psychology, 28*(3), 197–200.

Nielsen, K. (2021b). Comparing two enactive perspectives on mental disorder. *Philosophy, Psychiatry, & Psychology, 28*(3), 175–185.

Nielsen, K. (2022a). Affordances and 3E psychopathology. In *Affordances in Everyday Life* (pp. 149–156). Springer.

Nielsen, K. (2022b). Same diagnosis, different problem: The challenge of heterogeneity in mental disorder. *MIND Foundation*. https://mind-foundation.org/same-diagnosis-different-problem-the-challenge-of-heterogeneity-in-mental-disorder/

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 8*(6), 800–822. https://doi.org/10.1177/0959354318808394

Nielsen, K., & Ward, T. (2020a). Mental disorder as both natural and normative: Developing the normative dimension of the 3e conceptual framework for psychopathology. *Journal of Theoretical and Philosophical Psychology, 40*(2), 107–123. https://doi.org/10.1037/teo0000118

Nielsen, K., & Ward, T. (2020b). Phenomena complexes as targets of explanation in psychopathology: The Relational Analysis of Phenomena (RAP) approach. *Theory & Psychology, 30*(2), 164–185.

NIMH. (2022). *Developmental and environmental aspects*. https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/developmental-and-environmental-aspects

Nordenfelt, L. (2007). The concepts of health and illness revisited. *Medicine, Health Care and Philosophy, 10*(1), 5.

O'Connor, B. (2017). Mental disorder as a practical psychiatric kind. *Philosophy, Psychiatry, & Psychology, 24*(4), E-1-E-13.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–973.

Okrent, M. (2017). *Nature and normativity: Biology, teleology, and meaning*. Routledge.

Olbert, C. M., Gala, G. J., & Tupler, L. A. (2014). Quantifying heterogeneity attributable to polythetic diagnostic criteria: Theoretical framework and empirical application. *Journal of Abnormal Psychology, 123*(2), 452.

Ossorio, P. (1985). Pathology. In *Advances in descriptive psychology* (Vol. 4, pp. 151–202). JAI Press.

Papaspirou, P., & Moussas, X. (2013). A brief tour into the history of gravity: From Emocritus to Einstein. *American Journal of Space Science, 1*(1), 33–45.

Potochnik, A. (2010). Levels of explanation reconceived. *Philosophy of Science, 77*(1), 59–72.

Potochnik, A. (2016). Scientific explanation: Putting communication first. *Philosophy of Science, 83*(5), 721–732.

Potochnik, A. (2017). *Idealization and the aims of science*. University of Chicago Press.

Potochnik, A., & McGill, B. (2012). The limitations of hierarchical organization. *Philosophy of Science, 79*(1), 120–140.

Radden, J. (2006). *The philosophy of psychiatry: A companion*. Oxford University Press.

Radomsky, A. S., Alcolado, G. M., Abramowitz, J. S., Alonso, P., Belloch, A., Bouvard, M., Clark, D. A., Coles, M. E., Doron, G., & Fernández-Álvarez, H. (2014). Part 1 – You can run but you can't hide: Intrusive thoughts on six continents. *Journal of Obsessive-Compulsive and Related Disorders, 3*(3), 269–279.

Ramírez-Vizcaya, S., & Froese, T. (2019). The enactive approach to habits: New concepts for the cognitive science of bad habits and addiction. *Frontiers in Psychology, 10*, 301.

Rietschel, M. (2014). Mental disorders are somatic disorders, a comment on M. Stier and T. Schramme. *Frontiers in Psychology, 5*, 53.

Roberts, T., Krueger, J., & Glackin, S. (2019). Psychiatry beyond the brain: Externalism, mental health, and autistic spectrum disorder. *Philosophy, Psychiatry, & Psychology, 26*(3), E-51.

Sadler, J. Z. (1999). Horsefeathers: A commentary on "Evolutionary versus prototype analyses of the concept of disorder". *Journal of Abnormal Psychology, 108*(3), 433–437.

Sadler, J. Z. (2005). *Values and psychiatric diagnosis* (Vol. 2). Oxford University Press.

Sadler, J. Z., & Agich, G. J. (1995). Diseases, functions, values, and psychiatric classification. *Philosophy, Psychiatry, & Psychology, 2*(3), 219–231.

Seli, P., Risko, E. F., Purdon, C., & Smilek, D. (2017). Intrusive thoughts: Linking spontaneous mind wandering and OCD symptomatology. *Psychological Research, 81*(2), 392–398.

Stier, M. (2013). Normative preconditions for the assessment of mental disorder. *Frontiers in Psychology, 4*, 611.
Sullivan, J. A. (2017). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations, 20*(2), 129–145.
Szasz, T. S. (1960). The myth of mental illness. *American Psychologist, 15*(2), 113.
Szasz, T. S. (1963). *Law, liberty, and psychiatry: An inquiry into the social uses of mental health practices*. Syracuse University Press.
Szasz, T. S. (1974). *The myth of mental illness: Foundations of a theory of personal conduct, Rev*. Harper & Row.
Tabb, K. (2016). Philosophy of psychiatry after diagnostic kinds. *Synthese*, 1–19.
Tekin, S., & Bluhm, R. (2019). *the bloomsbury companion to philosophy of psychiatry*. Bloomsbury Publishing.
Telles-Correia, D., Saraiva, S., & Gonçalves, J. (2018). Mental disorder – The need for an accurate definition. *Frontiers in Psychiatry, 9*, 64.
Thagard, P. (2017). *Natural philosophy: From social brains to knowledge, reality, morality, and beauty* (draft 3).
Thomas, J. G., & Sharp, P. B. (2019). Mechanistic science: A new approach to comprehensive psychopathology research that relates psychological and biological phenomena. *Clinical Psychological Science, 7*(2), 196–215. https://doi.org/10.1177/2167702618810223
Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press. https://books.google.co.nz/books?id=OVGna4ZEpWwC
Thompson, E., & Cosmelli, D. (2011). Brain in a vat or body in a world?: Brainbound versus enactive views of experience. *Philosophical Topics, 39*(1), 163–180.
Thompson, E., & Stapleton, M. (2009). Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi, 28*(1), 23–30.
Thornton, T. (2000). Mental illness and reductionism: Can functions be naturalized? *Philosophy, Psychiatry, & Psychology, 7*(1), 67–76.
Vanderbeeken, R., & Weber, E. (2002). Dispositional explanations of behavior. *Behavior and Philosophy, 30*, 43–59.
Varela, F. J. (1992). *Autopoiesis and a biology of intentionality*, pp. 4–14.
Varela, F. J., Thompson, E., & Rosch, E. (2017). *The embodied mind: Cognitive science and human experience*. MIT Press.
Varga, S. (2011). Defining mental disorder. Exploring the 'natural function' approach. *Philosophy, Ethics, and Humanities in Medicine, 6*(1), 1.
Veit, W. (2020). Model pluralism. *Philosophy of the Social Sciences, 50*(2), 91–114.

Wadhams, G. H., & Armitage, J. P. (2004). Making sense of it all: Bacterial chemotaxis. *Nature Reviews Molecular Cell Biology, 5*(12), 1024–1037.

Wakefield, J. C. (1992a). Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder. *Psychological Review, 99*(2), 232.

Wakefield, J. C. (1992b). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47*(3), 373.

Wakefield, J. C. (1997a). Diagnosing DSM-IV – Part I: DSM-IV and the concept of disorder. *Behaviour Research and Therapy, 35*(7), 633–649.

Wakefield, J. C. (1997b). Normal inability versus pathological disability: Why Ossorio's definition of mental disorder is not sufficient. *Clinical Psychology: Science and Practice, 4*(3), 249–258.

Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry, 6*(3), 149.

Wakefield, J. C. (2013). The DSM-5 debate over the bereavement exclusion: Psychiatric diagnosis and the future of empirically supported treatment. *Clinical Psychology Review, 33*(7), 825–845.

Wakefield, J. C. (2014a). *The biostatistical theory versus the harmful dysfunction analysis, part 1: Is part-dysfunction a sufficient condition for medical disorder?, 39*, 648–682.

Wakefield, J. C. (2014b). Wittgenstein's nightmare: Why the RDoC grid needs a conceptual dimension. *World Psychiatry, 13*(1), 38–40.

Wakefield, J. C. (2015). DSM-5, psychiatric epidemiology and the false positives problem. *Epidemiology and Psychiatric Sciences, 24*(3), 188–196.

Walker, M. J., & Rogers, W. A. (2018). A new approach to defining disease. *The Journal of Medicine and Philosophy, 43*, 402–420.

Ward, D., Silverman, D., & Villalobos, M. (2017). Introduction: The varieties of enactivism. *Topoi, 36*(3), 365–375.

Ward, T., & Clack, S. (2019a). From symptom to clinical phenomena. *New Ideas in Psychology, 54*, 40–49.

Ward, T., & Clack, S. (2019b). From symptoms of psychopathology to the explanation of clinical phenomena. *New Ideas in Psychology, 54*, 40–49. https://doi.org/10.1016/j.newideapsych.2019.01.004

Ward, T., Clack, S., & Haig, B. D. (2016). The abductive theory of method: Scientific inquiry and clinical practice. *Behaviour Change, 33*(4), 212–231.

Ward, T., & Fischer, R. (2019). Behavioral and brain sciences commentary on Borsboom, Cramer and Kalis: Families of network structures – We need both phenomenal and explanatory models. *Behavioral and Brain Sciences, 42*(E31) https://doi.org/10.1017/S0140525X1800122X

Ward, T., Vertue, F. M., & Haig, B. D. (1999). Abductive method and clinical assessment in practice. *Behaviour Change, 16*(1), 49–63.

Wegerhoff, D. (2022). *Understanding gangs: Developing an epistemically pluralist framework for gang research*.

Wegerhoff, D., Ward, T., & Dixon, L. (2020). A pluralistic approach to the definition, classification, and explanation of gangs. *Aggression and Violent Behavior, 58*, 101546.

Wegerhoff, D., Ward, T., & Dixon, L. (2022). Epistemic pluralism and the justification of conceptual strategies in science. *Theory & Psychology, 32*(3), 443–466.

Wilshire, C. E., Ward, T., & Clack, S. (2021). Symptom descriptions in psychopathology: How well are they working for us? *Clinical Psychological Science, 9*(3), 323–339.

Zachar, P. (2014). *A metaphysics of psychopathology*. MIT Press.

Zachar, P. (2015). Psychiatric disorders: Natural kinds made by the world or practical kinds made by us? *World Psychiatry, 14*(3), 288.

Zachar, P. (2018). Diagnostic nomenclatures in the mental health professions as public policy. *Journal of Humanistic Psychology*. https://doi.org/10.1177/0022167818793002

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164*(4), 557–565.

Zachar, P., & Kendler, K. S. (2017). The philosophy of nosology. *Annual Review of Clinical Psychology, 13*, 49–71.

Zautra, N. (2015). Embodiment, interaction, and experience: Toward a comprehensive model in addiction science. *Philosophy of Science, 82*(5), 1023–1034.

Index[1]

[1] Note: Page numbers followed by 'n' refer to notes.

K. Nielsen, *Embodied, Embedded, and Enactive Psychopathology*, Palgrave Studies in the Theory and History of Psychology, https://doi.org/10.1007/978-3-031-29164-7