The Concept of Emotional Disorder

INTERNATIONAL PERSPECTIVES IN PHILOSOPHY AND PSYCHIATRY

Series editors: Bill (K. W. M.) Fulford, Lisa Bortolotti, Matthew Broome, Katherine Morris, John Z. Sadler, and Giovanni Stanghellini

VOLUMES IN THE SERIES:

Anatomy of an Avatar: The Virtual Self in Pathologies of Self-awareness
Gerrans

Philosophical Perspectives on Psychedelic Psychiatry
Letheby and Gerrans

Vice and Psychiatric Diagnosis
Sadler

Intruders in the Mind: Interdisciplinary Perspectives on Thought Insertion
Lopez-Silva and McClelland

Philosophy of Psychedelics
Letheby

Psychiatry Reborn: Biopsychosocial Psychiatry in Modern Medicine
Savulescu, Roache, Davies, and Loebel

Madness and the Demand for Recognition: A Philosophical Inquiry into Identity and Mental Health Activism
Rashed

Thomas Szasz: An Appraisal of HisL
Haldipur, Knoll, and Luft (eds.)

Psychiatric Neuroethics: Studies in Research and Practice
Glannon

Madness and Modernism: Insanity in the Light of Modern Art, Literature, and Thought
Sass

Philosophical Issues in Psychiatry IV: Psychiatric Nosology
Kendler and Parnas

Lost in Dialogue: Anthropology, Psychopathology, and Care
Stanghellini

Vagueness in Psychiatry
Keil, Keuck, and Hauswald

The Virtue of Defiance and Psychiatric Engagement
Potter

The Healing Virtues: Character Ethics in Psychotherapy
Waring

Embodied Selves and Divided Minds
Maiese

Naturalism, Hermeneutics, and Mental Disorder
Varga

The Abraham Dilemma: A Divine Delusion
Graham

Experiences of Depression: A Study in Phenomenology
Ratcliffe

Alternative Perspectives on Psychiatric Validation: DSM, ICD, RDoC, and Beyond
Zachar, St. Stoyanov, Aragona, and Jablensky (eds.)

Philosophical Issues in Psychiatry III: The Nature and Sources of Historical Change
Kendler and Parnas (eds.)

Is Evidence-based Psychiatry Ethical?
Gupta

Diagnostic Dilemmas in Child and Adolescent Psychiatry: Philosophical Perspectives
Perring and Wells (eds.)

Addiction and Weakness of Will
Radoilska

One Century of Karl Jaspers Psychopathology
Stanghellini and Fuchs

Emotions and Personhood: Exploring Fragility—Making Sense of Vulnerability
Stanghellini and Rosfort

Recovery of People with Mental Illness: Philosophical and Related Perspectives
Rudnick (ed.)

Autonomy and Mental Disorder
Radoilska (ed.)

Philosophical Issues in Psychiatry II: Nosology
Kendler and Parnas (eds.)

Discursive Perspectives in Therapeutic Practice
Lock and Strong (eds.)

The Sublime Object of Psychiatry: Schizophrenia in Clinical and Cultural Theory
Woods

Maladapting Minds: Philosophy, Psychiatry, and Evolutionary Theory
Adriaens and De Block

Thinking Through Dementia
Hughes

The Concept of Emotional Disorder

Gloria Sibson Ayob

Senior Lecturer in Philosophy and Mental Health
University of Central Lancashire

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number is on file at the Library of Congress

ISBN 9780198909606

DOI: 10.1093/9780198909620.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

For my family

Contents

Acknowledgements

I would like to thank the reviewers and colleagues, anonymous and non-anonymous, who read through the manuscript and provided incisive and constructive suggestions that have made it a significantly better piece of work than it would otherwise have been. Particular thanks go to Richard Gipps, Mark Pestana, Rowland Stout, and an anonymous reviewer for their detailed comments. This work is much improved thanks to their help. All mistakes are mine alone. My thanks also go to my editors at OUP—Martin Baum, Kayley Gilbert, Jade Dixon, Vinothini Thiruvannamalai, and Leo Mosquline—who have supported this manuscript in the journey it has taken from submission to production. For his patience and diligence in combing through the manuscript with such care and thoroughness, therein saving me from innumerable errors, I would like to express a special thanks to Tim Beck, my copy editor.

This book would never have been written without the philosophical guidance, encouragement, and generosity of my colleague, tutor, mentor, and dear friend Tim Thornton. I have grown and learned so much from our years of philosophical conversation—a conversation that has been, above all, so very enjoyable. Thank you, Tim.

I owe an intellectual debt to Bill Fulford, whose revolutionary work on the concepts of health and pathology constitutes the broad conceptual foundation upon which rests the analysis of the concept of emotional disorder I defend here. Without Bill's incisive analysis of the debate about the concept of mental disorder, and his own values-in analysis of this concept, it would have been significantly harder to articulate, in a familiar and easily recognizable way, the analytical framework presented in this book.

I have been deeply fortunate to work with some fine and wonderful colleagues at the University of Central Lancashire, whose presence has made for a friendly and intellectually stimulating milieu. Sharing together in the life of the mind with this very special community has been invaluable. My particular thanks go to Simon Baverstock, Chris Connell, David Cornell, Peter Herissone-Kelly, Emma Jones, Peter Lucas, Mick McKeown, Brian Rosebury and, of course, Tim Thornton.

I am very grateful for the conversations I have had with my students over the years on the MA in Philosophy and Mental Health programme. It was in

our collective grappling with the perplexing phenomena of psychopathology that the seeds of the argument in this book came to be sowed. I am especially thankful to Shirley Ward, whose boundless intellectual curiosity, and warmth and friendship, have been an important source of inspiration to me in writing this book.

My own curiosity about the assumptions that shape our lives—which is the soil in which this book is rooted—goes back some way. I am deeply grateful to David and Carmen Guard, M. P. Prabhakar, and the whole team at MABECS, who afforded me the privilege of studying philosophy and learning to engage in philosophical inquiry in a disciplined way. My tutors at Manchester and Warwick Universities reinforced my early enthusiasm: Tim Thornton, Rowland Stout, Helen Beebee, Julian Dodd, Harry Lesser, Jay Kennedy, Anthony Hatzimoysis, Thomas Uebel, and Michael Luntley all gave me a sound introduction to the subject and, in so doing, set me on the path that would eventually lead to the writing of this book.

In the many twists and turns my life has taken over the past decade, two notable events have directly contributed to the conception and progress of this book: first, meeting my husband-to-be Paul Sibson in 2014; and second, the birth of our son Leo seven years later. Leo's arrival in the autumn of 2021 signalled the start of a marvellous adventure—an adventure that continues right into the present moment. Through our evolving relationship, I learn anew, each day, what it feels like to love unconditionally. I have also learnt, in an immediate way, the power of the proverbial village in bringing up a child. But for the sustained support and wisdom of my parents, and of the magnificent Kate Reed, Ellie Gregory, Catherine Doyle Lauren Seddon-Roberts, Anita and Mike Spillane, Caroline Blunt, Olivia Fern, and of course Joan Dawson (Nana Joan), Leo wouldn't be growing into the curious, gleeful, and socially engaged little person that he is. My debt of gratitude to you all is beyond measure. My deep gratitude too to my dear friends Mila Petrova, Nimmi Prabhakar, Kate Reed, and Jo-Ann Lee for being the wonderfully brilliant and inspiring companions that you are. My life has been enriched in countless ways over the years through the gift of your friendship. I'd also like to express my special thanks to my sisters, Shanti and Rina Ayob, and to Paul's sister Joanne Sibson, for being lovely aunts to Leo. It is a marvel to behold: this wonderful circle of human love and warmth that Leo is growing up in.

Life with Paul was too immense, too full, too vivid to be described adequately in words (if I could only *paint* my gratitude). Our shared fascination about the workings of the human mind gave Paul and me countless hours of probing, rich, and often humorous conversation whilst walking together in the Lakeland fells. These have been amongst the most precious experiences

of my life, and it was through Paul's penetrating insights as a gifted psychotherapist that I began to truly wonder about our emotional lives. The wonder Paul sparked led me, in turn, to begin to appreciate the multifaceted nature of our emotional experiences and the complexity of the myriad ways by which we interpret their significance. These ideas form the kernel of the analytical framework that I have sought to articulate in this book. Paul's enthusiasm for this project is what sustained me through the ups and downs of the writing process, and it is for this reason now a completed piece of writing. Thank you, my beloved.

Above all, I have my parents to thank for so generously nurturing my nascent interest in reflecting upon the guiding assumptions that forge our lives. Without your enduring support, I simply wouldn't have been able to get started, let alone to continue to pursue this interest for as long and as wholeheartedly as I have done. Thank you, Papa and Mummy.

Gloria Sibson Ayob
June 2024

The far-off land may have bays, forelands, angles in and out to any number and extent; yet at a distance you see none of these . . . nothing but a grey unbroken line upon the water.

Edwin Abbott, *Flatland: A Romance of Many Dimensions* (2010: 18)

Introduction

I.1 Emotional Disorder: The Very Idea

The notion that it is possible, literally, to fall prey to emotional illness is as familiar as it is intriguing. For although we have grown used to casting states of extreme anxiety or despondency as pathological, the implication that these states are subsumable under the same broad category as bronchitis, pneumonia, or measles (say), prompts some basic questions. In virtue of what pathogen (or pathogen-like entity) is a state of dejection rendered pathological? Or if this question incorrectly presumes that the notion of emotional illness is sufficiently closely related to an ailment like pneumonia, then where in the space of pathology should emotional illness be situated? What sort of non-emotional ailment does disordered despondency most closely resemble, a resemblance in terms of which we might make sense of the claim that despondency can take on a pathological character? Perhaps emotional illness is *sui generis*, unlike any other sort of pathological kind. In that case, these questions are misleading ones to ask, but now the idea that we can fall ill emotionally becomes all the more intriguing. What exactly is meant by the claim that extreme anxiety can constitute a *sui generis* pathological state? Why do we find it useful to speak in this way about our emotional distress? And even if it has its uses, can we make good on this claim of pathology? How might we go about establishing the legitimacy of the concept of emotional illness?

These questions are the subject of inquiry in what follows. For all that we have become accustomed to diagnosing and treating disorders of emotion, the concept of emotional disorder itself has received little scrutiny in its own right.[1] Complaints that psychiatry over-pathologizes our emotional lives tend to focus on specific kinds of emotional experience (sadness or anxiety), without these complaints being anchored in a worked out general notion of emotional disorder. Even Jerome Wakefield's prominent criticism of psychiatry on this point assumes, rather than articulates, a concept of emotional disorder. As we'll see, Wakefield's criticism of psychiatry is anchored in an

[1] I will be using the terms 'disorder', 'pathology', and 'illness' interchangeably in what follows because nothing of significance in the inquiry that follows turns upon suggested differences between the three terms.

The Concept of Emotional Disorder. Gloria Sibson Ayob, Oxford University Press. © Oxford University Press 2024.
DOI: 10.1093/9780198909620.003.0001

analysis of the general concept of disorder, which he then presumes is applicable in a straightforward way to the emotions. But the concept of emotional disorder is—as I hope to show—multi-faceted and far more complex than we might initially suppose it to be, which makes this presumption a problematic one.

The task of clarifying the charge of over-pathologization alone merits close investigation of the concept of emotional disorder. But given the central position occupied by emotions in human life, this task has important consequences that stretch far beyond this specific debate within the philosophy of psychiatry. Appreciating the complex structure of judgements about emotional disorder will shape how we think about disorder and illness generally, as well as bearing upon our view of ourselves as self-governing agents capable of believing, deciding, and acting in morally and personally significant ways. The concept of emotional disorder also informs our understanding of what it means to flourish or to languish, so grasp of its complexity is essential for properly appraising our lives at the broadest of levels. All this testifies to the central position that our emotions occupy in our lives. It also underscores the need to carefully examine the notion of emotional pathology in its own right, an inquiry that is long overdue.

I.2 The Charge of Over-Pathologization: Horwitz and Wakefield vs Kendler

I want to begin this investigation into the concept of emotional disorder by taking a closer look at the charge that psychiatry over-pathologizes or over-medicalizes our emotions. Critics of psychiatry who raise this charge worry that the concept of emotional disorder is being applied too liberally and that we are wrongly consigning far too much of our emotional lives to the container of pathology. The charge of over-pathologization is familiar in psychiatry generally, but it is worth remarking that, when raised in the context of emotional distress, this concern is put forward even by authors who are otherwise broadly sympathetic to psychiatry and accept that it is a legitimate branch of medicine. In order to explore the charge from this moderate position, I'll turn to a recent exchange between the authors Jerome Wakefield and Kenneth Kendler.

Wakefield and his co-author Allan Horwitz have co-authored two monographs (2007, 2012) in which they argue that contemporary psychiatry over-pathologizes our emotional lives—most notably, with respect to the experience of anxiety and of sadness. Referring to the revisions made to the

diagnostic criteria for major depressive disorder (MDD) and the eight types of anxiety disorder in the latest edition of the Diagnostic and Statistical Manual of Mental Disorders—the DSM 5— Horwitz and Wakefield contend that these diagnostic criteria are problematic because they wrongly subsume normal states of human sadness and anxiety under these categories of disorder, thereby incorrectly pathologizing them. The psychiatrist Kenneth Kendler seeks to defend psychiatry against this charge (Kendler 2008).[2]

Horwitz and Wakefield begin their 2007 book, *The Loss of Sadness*, by noting the manifold increase of diagnoses of disordered sadness over the past thirty years in the US, reporting that the treatment of depression in outpatient settings (where most depressed people are treated) increased by 300 per cent between 1987 and 1997, and that the overall percentage of the US population in treatment for depression in a particular year grew from 2.1 per cent in the early 1980s to 3.7 per cent in the early 2000s, an increase of 76 per cent in just twenty years (Horwitz and Wakefield 2007: 4).[3] Horwitz and Wakefield believe that this rapid increase in the rate of depression is best explained by the erroneous pathologization in the DSM of ordinary human sadness. They state:

> The basic flaw . . . of the DSM definition of MDD [Major Depressive Disorder] . . . is simply that it fails to take into account the *context* of the symptoms and thus fails to exclude from the disorder category intense sadness . . . that arises from the way human beings *naturally* respond to major losses . . . . Moreover, as we show, the problem has been getting much worse in recent years, with growing pressure to use a lower number of symptoms, sometimes as few as two, as sufficient criteria for diagnosing a disorder. The potential for false-positive diagnoses—that is, people who meet the DSM's diagnostic criteria but do not in fact have a mental disorder—increases exponentially as the number of symptoms required for a diagnosis decreases. (Horwitz and Wakefield 2007: 14, italics added)

According to Horwitz and Wakefield, the context in which sadness is experienced ought to play a determining role in our assessment of that sadness as pathological, and the source of the DSM's error lies in its failure to recognize this point. They point out that when sadness is experienced as a response to personal loss, it is a normal (non-disordered) state; sadness only becomes pathological when it lacks this vital context, that is, when it is experienced

[2] Kendler served as a member of the DSM-5 Work Group for Mood Disorders from 2007 to 2010.

[3] In the UK, almost as dramatic an increase as been noted in the same period, at least with respect to the use of antidepressant medication. The National Institute for Clinical Excellence (2004) . . . has shown that antidepressant usage has increased by 234 per cent between 1992 and 2002 (Davies 2012: 61).

in spite of its not being a discernible response to any personal loss. The DSM category, Major Depressive Disorder (henceforth 'MDD_{DSM}'), fails to register this intuitively compelling distinction because its diagnostic criteria focus exclusively on the 'symptoms' of sadness, without any reference to the contexts in which people experience these symptoms. This blocks from view the possibility that these symptoms might be responses to specific circumstances—in this case, to circumstances of personal loss, the natural human response to which is feeling sad. In effect, Horwitz and Wakefield's criticism of psychiatry is that the DSM fails to recognize normal human sadness for what it is and, in so doing, it places the boundary between normality and pathology in the wrong place.

Horwitz and Wakefield are pointing at something that we ordinarily take for granted and which seems so obvious as to be a truism, namely that sadness can be (and often is) a natural response to personal loss. However distressing the sadness felt may be, feeling sad seems to us to be an integral part of what it means to lose someone or something we value. Indeed, it is arguably through the experience of sadness that the personal loss comes to have the significance that it does for us: not merely an absence of what was once present, but as *loss*. Far from being necessarily disordered in and of itself then, sadness is what enables us to experience loss and, in this, to recognize all that we valued in the relationship, the situation, or the object that is now gone.

In the face of this point, the occlusion in category of MDD_{DSM} of the possibility that sadness could be a response to loss seems so counterintuitive as to be obviously dubious. So it comes as a surprise to read that Kenneth Kendler's defence of the category against Horwitz and Wakefield's critique *also* appeals to our intuitions. Kendler focuses on the distressing quality of sadness, describing a patient Ms. S who is in her late 20s but has a life-threatening condition which could prove fatal at any moment and who struggles to function in her daily life as a result. She seeks help from her psychiatrist, Kendler, hoping to find something that will help her feel better, and Kendler prescribes an antidepressant drug in response. Arguing against Horwitz and Wakefield, Kendler states:

> She had an excellent response over the subsequent 6 weeks and achieved a full remission of her symptoms. Although I found her depression *to be entirely appropriate to* her tragic situation, the argument that I should not have treated her so that she could experience an 'inherent and valuable part of the human condition' obviously rings rather hollow. (Kendler 2008: 149, italics added)

Kendler identifies a situation in which sadness is experienced as a normal response to loss (in this case, not just the loss of good health, but the imminent threat of losing one's own life), but which he thinks merits medical treatment. Here at the outset, it is worth taking note of the following point: while Kendler's response to Ms. S seems intuitively compelling, it also seems to trade on a crucial ambiguity, which it will be useful to disentangle. Although pharmacological interventions in healthcare are usually understood to be aimed at treating and removing an underlying pathology, pharmacological substances can be used for enhancing a person's quality of life even in the absence of any such pathology (e.g. in recreational drug use). So in prescribing the antidepressant, it might be argued that all Kendler seems to have done is to invoke the latter type of use—i.e. medication for enhancement. So long as we accept the plausible claim that the two uses of pharmacology are distinct, then it seems possible to medicate for enhancement without presupposing that, in so doing, one has therein treated a pathological condition. On this reading, Kendler's response would seem to be consistent with viewing the sadness that Ms. S feels as a natural (non-pathological) response to her life-threatening condition. There is thus no dispute between Horwitz and Wakefield, on the one hand, and on the other, Kendler.

But Kendler himself contends that his response does bear upon the question of where we draw the boundary between normal and pathological sadness. Drawing upon his discussion of the case of Ms. S, Kendler responds to Horwitz and Wakefield's charge that MDD_{DSM} over-pathologizes sadness by arguing that it is in fact Horwitz and Wakefield who fail to recognize genuine instances of pathological sadness, and that they do this because they have 'over-romanticiz(ed) the suffering associated with the condition of MDD' (Kendler 2008: 149) The fact that Kendler takes his response to Ms. S to have this demarcational significance indicates that Kendler understands himself to be invoking the first use of pharmacology just noted (medication for treating pathology), rather than its second use (medication for enhancement).

Having clarified the import of Kendler's response, we can now say something about its structure. Alongside its intuitive appeal, Kendler's view becomes yet more plausible when we take into account the fact that the alleviation of suffering has always been one of the basic goals of healthcare. Admittedly, since suffering comes in a multitude of forms, not all of which consist in (or are caused by) pathology, it seems reasonable to seek to qualify this point in one important way: this basic goal of healthcare refers specifically to the alleviation of *pathological* suffering, where this refers (roughly) to the

idea of suffering that constitutes a pathological condition.[4] Kendler's point, it seems to me, is that a key variable that shapes our intuitions about whether suffering is pathological—at least in the domain of human emotion—is the degree of its severity. If emotional distress is sufficiently severe to immobilize its subject (as is the case for Ms. S, who reports being unable to carry out her day-to-day functioning), then this inclines us towards the view that it is the kind of suffering that medicine *ought to* alleviate if it has the means to. Importantly, Kendler contends, the alleviation of suffering here does not seem to be aimed merely at enhancement. The suffering seems pathological. If Kendler is right about this, then the degree of severity of sadness experienced would appear to be an independent variable in determining whether the suffering wrought is pathological, i.e. separate from the fact that Ms. S's severe distress is a reasonable response to excruciating life circumstances. Placing greatest emphasis on this variable, Kendler concludes that the boundary between normal and pathological sadness is located in the right place after all vis-à-vis the category of MDD_{DSM}, *pace* Horwitz and Wakefield's charge of over-pathologization.

I.3 Who Is Right?

Both Horwitz and Wakefield's and Kendler's respective positions seem to have at least *some* intuitive appeal in their own right. This points at a striking and important feature of our concept of emotional pathology: our intuitions about this form of pathology appear to shift depending on which aspects of an emotional experience we foreground. This kaleidoscopic character of emotional experience compounds the challenge of making sense of the charge of over-pathologization and that of over-romanticization and deciding between them, in the following way. Horwitz and Wakefield and Kendler all assume that there is a fact of the matter as to whether or not a given emotional response is disordered. This assumption is what makes intelligible the charge that a swathe of normal human suffering is being wrongly pathologized (Horwitz and Wakefield's charge), or that genuine illness is being incorrectly under-diagnosed (Kendler's counter-charge). What both parties further assume, though, is that it is possible to determine the pathology status of an emotional experience directly by reference to a particular aspect of this experience—its reasonableness as a response to specific life circumstances, say, or the severity of the suffering that is wrought by it. But, as I shall seek to show in the next

[4] Kendler writes: 'Our central goal as a medical discipline is the alleviation of the human suffering that results from dysfunctional alterations in certain domains of first-person, subjective experience, such as mood, perception, and cognition' (Kendler 2005: 433).

chapter, once we begin to recognize the kaleidoscopic character of our emotional experiences, this recognition puts pressure on the simple appeal to individual features of an emotional experience to settle its pathology status.

This point is worth underscoring for its importance. If a single aspect of an emotional experience were sufficient for determining its pathology status, it would be possible to 'read off' disorderedness from the experience directly (so long as the relevant aspect of the emotional experience is a sensible property). Disorderedness would be transparent. But the exchange between Horwitz and Wakefield and Kendler suggests that judgements of emotional disorder do not work this way. In his response to Horwitz and Wakefield's critique, Kendler has to appeal to some particular aspect of sadness of the kind that the category of MDD_{DSM} aims at capturing, in order to defend his counter-claim that sadness of this sort is genuinely pathological. And that is precisely what Kendler does: he anchors his defence of psychiatry in a particular aspect of MDD_{DSM}-sadness, namely its degree of severity (MDD_{DSM}-sadness is so severe as to be immobilizing). Once this move is registered, we see that Kendler's response involves offering us a reason for thinking that Horwitz and Wakefield are wrong in their charge that psychiatry over-pathologizes human sadness, but that this reason doesn't deliver a knock-down blow against Horwitz and Wakefield's critique. By foregrounding different aspects of MDD_{DSM}-sadness, Horwitz and Wakefield and Kendler respectively arrive at contrary conclusions concerning its pathology status. Absent any further reason for favouring the aspect Kendler emphasizes (the severity of the sadness suffered) over the one emphasized by Horwitz and Wakefield (sadness that is experienced in spite of there being no personal loss suffered), what we have is a stand-off between two rival responses to the boundary problem, each appealing to a different aspect of sadness to justify this response.[5]

[5] Although the dispute between Wakefield and Kendler has unfolded further since (see Zachar et al. 2017 for a review of this debate), the essential fault lines of this dispute are set out with particular clarity in this early exchange, which makes it a useful point of departure into the present inquiry. As the debate unfolds, we see Wakefield and Kendler disagreeing about which comparisons between groups are the relevant ones to make in deciding whether or not to retain the DSM-IV's bereavement exclusion clause. Claiming that the DSM-IV's bereavement exclusion clause results in 'logically inconsistent' stances adopted towards (A) those individuals with bereavement-related depression (deemed by the fourth edition of the DSM—DSM-IV—not to be a medical condition) and (B) those individuals with other stress (i.e. non-bereavement)-related depression (deemed by DSM-IV to be a medical condition), Kendler argues that both classes of depression merit a medical diagnosis of depression. He therein concludes that the bereavement exclusion clause should be eliminated from the DSM. Note that in drawing this comparison, Kendler makes no reference to (C) those individuals who manifest depressive symptoms but who do so without any obvious precipitating stress-inducing loss. Yet, it is exactly this latter category that Wakefield thinks is the *most* salient group to compare with those individuals who suffer from bereavement (and other loss)-related depression; and it is on this basis that Wakefield argues for the retention of the bereavement exclusion clause. In other words, Wakefield grants Kendler his point about consistency between groups (A) and (B), but he contends that Kendler is failing to draw the contrast that is of greatest importance to the delineation of the boundary of depression as a medical disorder, namely the contrast between groups (A) and (B) on the one hand, and on the other, group (C).

That each party has to appeal to a particular aspect of sadness in order to motivate their respective responses, and that we are confronted with the question of how to adjudicate between the responses, points at one crucial thing: the concept of disordered sadness appears not to be transparent. In contrast with delusions or verbal hallucinations, say, the disorderedness of an emotional experience doesn't wear itself on the sleeve of the experience itself: neither of the features of an emotional experience emphasized by Horwitz and Wakefield and by Kendler respectively seems decisive in determining its pathology status. In order to arrive at a principled response to the boundary problem, the fact is that we need to adjudicate between the respective considerations that both parties appeal to. It is this need for adjudication that renders the concept of emotional disorder opaque.

How then are we to determine where the boundary lies between normal and pathological sadness? More generally, what goes into the evaluation of our emotional experiences as pathological or otherwise? For reasons that will become clear in the following chapters, an adequate response to the boundary problem with respect to emotional pathology must be informed by an appreciation of the opacity of the concept of disordered sadness. I will begin to explore the boundary problem in the first chapter with the aim of elucidating this claim by undertaking an exercise in conceptual anthropology, observing a range of practices of evaluating our emotional experiences that both directly and indirectly shape judgements of emotional disorder. While it is now widely accepted that the concept of mental disorder is at least partly normative, one conclusion that this exercise in conceptual anthropology points towards is that, because of the opacity of the concept of emotional disorder, judgements about emotional disorder are value-laden in a distinctive way. An appreciation of the unique way in which such judgements are normative is essential for understanding and navigating debates about the legitimate scope of the concept of emotional pathology. As the ensuing discussion will show, resolving the dispute between Horwitz and Wakefield and Kendler may be less a matter of getting right the facts about emotional pathology, and more a matter of

As will become clear, the Basic Evaluative Dimension (BED) framework that I offer in this book sheds light on this disagreement about which comparison matters most to the delineation of the boundary of emotional disorder with regard to our states of sadness, anhedonia, and suchlike. To anticipate the discussion in Section 1.8, I will suggest that it is because Kendler prioritizes the basic evaluative dimensions of emotional valence and (high-level) agential efficacy that he pays little attention to the presence or absence of a precipitating personal loss. By contrast, Wakefield prioritizes the basic evaluative dimensions of representational accuracy and (lower-level) volitional efficacy, and this explains why he invests in the presence/absence of a precipitating personal loss a boundary-delineating significance.

understanding what values are at stake in deciding whether or not an emotional experience of a specific kind is pathological.

One important qualification is worth making at the outset: in examining responses to the boundary problem with respect to emotional pathology, I shall not be concerned with the appeal to, and use of, pharmacological substances for the purpose of functional enhancement. My concern will be exclusively with the justification of medical intervention on the grounds of treating pathology. Thus discussions of the use of mood-altering substances (say) for improving a person's mood purely for the sake of it—or to enable him to function more efficiently in daily life (where the latter concern is not used to anchor the *further* claim that the person suffers from a mood disorder on account of his not functioning as effectively as he'd like to)—lie beyond the scope of the present book.

It is important also to note a vital terminological point before I press on: although I speak of 'disordered emotion', and philosophers have rightly noted that there are significant and interesting differences between emotions and mood states, my discussion in this book will encompass emotions, mood, and other affective states, reflecting the broad use to which the term is put in the psychiatric and psychotherapeutic literature. As well, unless otherwise indicated, I will use the term 'emotional states' to cover a range of mental state kinds, including emotional responses (token and type) and emotional dispositions (again, token and type). This is because, unless specifically indicated, these distinctions do not make a difference to the arguments put forward in this book.

I.4 Looking Ahead

Philosophical discussions about the concept of emotional disorder tend to be informed primarily by the broader debate about the concept of mental disorder (or that of disorder more generally). Little if any serious consideration is given to the separate philosophical debate about the nature of emotions. But it is a mistake to overlook the latter debate. Our emotions are remarkably complex mental states—as even a brief survey of the latter debate will reveal—and it would be premature to assume in light of this complexity that we will be able to arrive at an adequate understanding of the concept of emotional disorder simply by fitting it into existing analyses of the concept of disorder. These analyses simply may not be up to the job of handling the unique complexity of

emotional states. It may be that we will need to fundamentally revise our sense of what the adequacy constraints are on analyses of the disorder concept once we begin to appreciate the demands placed by a fuller, more accurate understanding of our emotional lives. I will now turn to the task of filling out the concept of emotional disorder in light of our understanding of the nature of the emotions generally.

1
Setting out the Basic Evaluative Dimensions (BED) Framework

What is an emotion?

James (1884: 188–205)

1.1 Introduction

In discussions about the concept of emotional pathology, emphasis is usually placed on the *pathology* constituent of this concept. In this way, our attention tends to be constrained by the framing concept of mental disorder. Horwitz and Wakefield's conceptualization of depression and anxiety in terms of Wakefield's harmful dysfunction (HD) analysis of the disorder concept exemplifies this tendency clearly. The notion of HD, which serves as the analysans of the concept of disorder, forms the analytical framework within which the concept of emotional disorder is characterized. Within this framework, emotions are conceived of as psychological mechanisms that have been naturally selected for and it is assumed that we have a clear understanding of what constitutes a failure of functioning (a dysfunction) in these emotion mechanisms. It is taken as self-evident, for instance, that sadness is a 'loss response mechanism' and that anxiety is a 'threat response mechanism', and thus that disordered sadness and disordered anxiety consist in a dysfunction in these two mechanisms respectively.

By turning our gaze away from the framing concept of disorder and towards that of emotion, however, and emphasizing instead the *emotion* constituent of the concept of emotional pathology, a whole new horizon comes into view. The philosophical and psychological study of the emotions reveals just how complex and multifaceted our emotional experiences are. Once we bring this complexity into view, a natural question arises: granting Horwitz and Wakefield's claim that sadness is a response to loss, why do they

The Concept of Emotional Disorder. Gloria Sibson Ayob, Oxford University Press. © Oxford University Press 2024.
DOI: 10.1093/9780198909620.003.0002

emphasize just this particular facet of sadness in their conceptualization of pathological sadness? True, our emotional experiences (in general) purport to provide us with information about our immediate circumstances and they are a means by which we come to have concerns and to experience things as valuable. The emphasis on sadness as a response to loss makes sense in light of these aspects of our emotional lives. But our emotional experiences offer us a whole lot more besides this. They motivate us to act; they are integral to our experience of life as pleasant and enjoyable (or otherwise); they are defining elements in the constitution of human character; they provide a kind of social glue without which human life would look very different. How, if at all, are these latter aspects of our emotional experiences captured by a concept of disordered sadness that refers just to its loss response aspect? If they aren't captured by this concept, what justification do we have for prioritizing just this aspect of the experience of sadness in our concept of emotional pathology?

One might wonder why we should want to capture these latter aspects of our emotional experience in our concept of emotional disorder in the first place. We can only really answer this question properly by initially adopting an open-ended stance in our inquiry into the concept of emotional disorder. Foregrounding the concept of emotion in our thinking about the nature of emotional pathology helps us do this. Specifically, it helps us to recognize that we in fact appraise our emotional experiences in a variety of ways because of the many roles that our emotions play in our lives, all of which matter to us. We appraise emotions with regard to the information they purport to provide us with; we appraise our emotions for their effectiveness in motivating action; we appraise our emotions as a source of our concerns and values.[1] With this horizon now in view, we can avoid immediately pegging the notion of emotional pathology to just one aspect of our emotional experiences, and we can do this by asking two more basic questions still: how do judgements of emotional pathology relate to the various appraisals we make of our emotional

[1] Note that the term 'appraisal', as I use it here, is not derived from or connected to its use in the 'appraisal theory' of emotion (put forward by Frijda 1986 and 1993, and subsequently developed amongst others by Lazarus 1991 and Scherer 2005). In the latter context, an 'emotional appraisal' refers to the appraisal of a situation that a subject makes *within* an emotional experience (thus: in feeling fear, the subject appraises her present circumstances as threatening or dangerous). By contrast, I am speaking here of the appraisals we make *of* our emotional experiences—as rational or otherwise, as pleasant or otherwise, as agentially efficacious or otherwise, etc. Alisdair Macintyre highlights this 'double connection between emotion and evaluation', remarking that 'In judging that . . . some particular expression of emotion is appropriate or inappropriate, we evaluate it. And in feeling . . . our own emotions, . . . we characteristically treat their objects as deserving and inviting our responses to them' (Macintyre 2016: 153).

experiences? And does the notion of emotional pathology represent one further, independent dimension of appraisal? Or is it instead derived from the foregoing appraisals?

1.2 The Concept of Emotional Disorder: A Meta-evaluative Concept

It is a striking characteristic of our talk of emotional disorder that we seem always to arrive at the idea of disorderedness only through an appraisal of our emotional experiences that does not itself overtly invoke the notion of pathology. We doubt the reasonableness of our emotional tendencies, or we feel oppressed by the deeply unpleasant quality of a distressing emotional experience, or we are frustrated by the lethargy induced by this experience, and it is through these basic appraisals that we begin to wonder if such emotions are pathological. This suggests that the concept of emotional disorder depends upon these more basic appraisals.

To make vivid this idea that judgements of emotional disorder are dependent upon appraisals that do not themselves appear to refer to the notion of pathology, imagine that we are presented with a questionnaire in which we are asked to make judgements about disordered sadness. The questionnaire comprises questions such as the following:

- On a sadness scale of 1 to 10, where 1 is least intense and 10 is most intense, would:
  a. A score of 3 in response to a favourite garment that has disappeared be normal or disordered?
  b. A score of 5 in response to a poor performance in one's annual appraisal at work be normal or disordered?
  c. A score of 8 in response to the demise of an intimate friendship be normal or disordered?

How are we to respond to these questions? Any answer we give will feel arbitrary because, as it stands, the questionnaire doesn't seem to contain sufficient information about the meaning of this sadness scale. To be in a position to offer sensible answers, we would at the very least need more information about the notion of intensity being invoked. The sort of information it is necessary to have here includes answers to the following questions: in Question c., is a score of 8 *proportionate* (in the intensity of the sadness felt) to the loss

suffered?[2] Does this score signify *a crippling of our agency*? Does this score stand for an experience that is *deeply unpleasant*? It is only by filling out the notion of emotional intensity in these more basic evaluative terms that we will be in a position to respond meaningfully to this questionnaire.

The radical incompleteness of this hypothetical questionnaire—and the kind of information that would have to be added to make it meaningful—points at the opacity of judgements of emotional disorder, in the sense discussed in the Introduction. The meaning of the notion of 'emotional intensity', in terms of which the concept of pathology is framed in this questionnaire, is not self-evident. Rather, it is given meaning by reference to further aspects of the experience of sadness, e.g. its affective quality (as pleasant/unpleasant), its impact on our agency, its proportionality to the loss that is suffered. It is only by reference to the appraisals made with respect to these various aspects of sadness that we can begin to form judgements about whether the sadness in each case is normal or disordered.

In fact, there seem to be at least five basic dimensions along which we evaluate our emotional experiences. The judgement that an emotional experience is disordered is a meta-evaluative judgement that is anchored in our appraisal of emotional experiences along these five more basic evaluative dimensions. This is the central claim of this book. Support for this claim derives from observations of the actual practices we engage in when judging the pathology status of our emotional experiences—so the discussion that follows may be understood as an exercise in conceptual anthropology. These five basic evaluative dimensions are as follows:

1. Emotional rationality.
2. Existential significance.
3. Affective valence.
4. Volitional and agential efficacy.
5. Interpersonal functioning.[3]

The judgement of emotional disorder depends on the appraisals we make of our emotional experiences in these five basic evaluative dimensions. Slightly more specifically, it results from our negotiation, either explicit or implicit, of

[2] The notion of proportionateness is itself opaque and requires further analysis, which I will offer shortly.

[3] While the five basic evaluative dimensions aim at comprehensiveness, it is possible that the list presented here is not exhaustive: we may with further investigation find it useful to recognize other dimensions of basic evaluation, or to individuate the existing ones in a finer-grained way. Still, the five dimensions to be presented in this chapter help us make a good start in understanding the structure of the concept of emotional disorder.

the appraisals made in these dimensions. While each of these basic dimensions appears to be distinct and irreducible, they can and do mutually influence and constrain each other in judgements of emotional disorder in ways that we shall see. For brevity, I shall refer to this framework for understanding the concept of emotional pathology as the basic evaluative dimensions (BED) framework.

Once we recognize that the concept of emotional disorder is a meta-evaluative concept that derives from the appraisals made along these five more basic dimensions of evaluation, we are in a position to make the following diagnostic point. Disputes about the boundary problem—and about the pathology status of particular emotional experience—may be due, at least in part, to the different relative weightings that each party to the dispute assigns to these basic evaluative dimensions. I will return to this point in Section 1.8, once we have explored each of the five basic evaluative dimensions. It is to this latter task that I shall now turn.

1.3 The First Basic Evaluative Dimension: Emotional Rationality

1.3.1 Cognitivism about Emotions

The idea that our circumstances can warrant or fail to warrant emotional responses of particular kinds has an ancient provenance, finding expression in Aristotle's famous dictum that a person merits praise when he is angry at the right things 'as he ought, when he ought, and so long as he ought' (*EN* IV.5). This notion implies that there is a mind-to-world fit between our circumstances and the emotions we feel in them, and it has been filled out in numerous ways. I will focus on a particular contemporary account for illustrative purposes. On this account, the idea that our emotional responses can fit (or fail to fit) the circumstances that elicit them is expressed by saying that our emotional responses can be representationally accurate.[4] An emotional response that is representationally accurate is judged to be reasonable or warranted, and one that is inaccurate may be judged to be unreasonable and unwarranted. To take a simple example of this idea: a fearful response is taken to

[4] Greenspan 1988, de Sousa 1988, Solomon 1993, 2008.

depict an object, or a situation, as threatening. Insofar as the object or situation is actually threatening (or rather insofar as the subject has good reason to suppose it is), the fearful response is said to accurately represent the threatening situation to the subject, i.e. as threatening. On this basis, the experience of fear is said to be *appropriate to* the situation—that is, it is *rational* to feel fear in this situation. This is one important sense in which the term 'emotional rationality' is understood.[5]

How an emotional experience can accurately (or inaccurately) represent an object or a situation to its subject has been, and continues to be, a question that is extensively debated. Unlike perceptual states and belief states, whose intentionality strikes everyone as obvious, emotion states have been viewed in diverse ways. Proponents of feeling-centred conceptions of the emotions tend to assimilate emotional experiences to bodily sensations like itches, tickles, or pain: while they may be pleasant or unpleasant, they lack the property of being directed at something.[6] Viewing emotional states in this way rules out the possibility that our emotions are apt for rational assessment. Against the feeling-centred conception of emotion, it has been pointed out that, when we feel angry, we feel angry about something; when we feel love, we feel love towards someone; again, it doesn't make sense to speak of feeling remorseful without feeling remorseful about something in particular. On the basis of these truisms, it is argued that emotions are intentional states, and this view of emotions is now a more widely accepted starting point for thinking about the nature of emotion (cf. Solomon 1993). Taking emotion states to be intentional enables us to see them as capable of representing our circumstances in some way or other, and on this basis, as apt for rational assessment in the manner just sketched out.[7] This view of emotions has come to be referred to as the cognitivist view.[8]

Once cognitivism is accepted, questions arise concerning the intentional objects of our emotional experiences and the manner by which these experiences represent their objects as being the way they are (or failing to). One thing that makes the intentionality of emotion states intriguing is that what

[5] There are two further common notions of emotional rationality: strategic rationality (which I will look at below) and rationality assessments of our emotional states deriving from their position in the inferential and causal network of beliefs, desires, intentions, etc. in which our emotional states are embedded.

[6] Deigh (2008: 40) attributes this conception of emotion to early modern philosophers—including Descartes, Locke, and Hume.

[7] In order to take into account generalized moods (such as anxiety and joy) that seem objectless, Deigh (ibid.: 41) remarks that intentionality is typical, rather than essential, for emotions.

[8] I use this term in its broadest sense, i.e. to contrast it with the feeling-centred conception of emotion. Thus, as the term 'cognitivism' is used here, the question remains open as to whether *qua* intentional states, emotions are more like judgements or beliefs, or whether they are better thought of as akin to perceptual states, or whether indeed they constitute a *sui generis* intentional state.

is represented by an emotional response is typically not a bare sensible property, but rather an evaluative one. Danger (the property represented by the emotional experience of fear), injustice (represented by the experience of anger), novelty (represented by the experience of surprise), personal loss (represented by the experience of sadness) are common examples of this. What shape do our emotional experiences take, such that they are capable of depicting or embodying these evaluations of danger, injustice, loss, etc.? In addition to have a feeling tone or affective quality (as emphasized by feeling-centred theorists), cognitivists hold that our emotional experiences have an evaluative content. This will constrain the response a cognitivist gives to this question, but there are still several ways of responding to it. Are the evaluative content of an emotional experience and its feeling tone distinct components within the experience, with just the evaluative content being intentionally structured? Or is there in fact just a single intentionally structured mental state, with the evaluative content of an emotional experience being carried in some way through its affective quality? And precisely what is the shape of this intentional structure through which an emotional experience evaluates: is it an evaluative judgement or belief, or is this evaluation more akin to perceptual content?

A wide range of responses have been offered to these questions, and they continue to be debated. Since our concern is with the concept of emotional pathology, it suffices simply to note two things: firstly, that the view that emotional experiences are intentional states (cognitivism) now receives broad acceptance, and secondly, that this is what makes intelligible the claim that our emotional experiences can provide us with evaluative information about our circumstances in such a way that makes them apt for rational assessment.

1.3.2 Emotional Rationality, Emotional Intensity, and Emotional Pathology

A quick survey of the fifth edition of the DSM shows that the notion of emotional intensity is used in judging emotional experiences to be disordered. As the list of DSM categories presented below shows, this appeal to emotional intensity is expressed through the language of proportionality, i.e. an emotional experience that is disproportionately intense is deemed to be pathological. This idea provides a vital link between the concepts of emotional pathology and emotional rationality: it is an important way by which the appraisal of an emotional experience as rational/irrational informs the judgement of emotional disorder. I will explore this point here, with the aim of showing that the

concept of emotional disorder is a meta-evaluative one that derives in part from the more basic concept of emotional rationality.

We can begin by noticing that the following categories of emotional disorder in the DSM explicitly employ the idea that an emotional experience that is disproportionately intense is pathological:

> Disruptive Mood Dysregulation Disorder. Criterion A: severe recurrent temper outbursts manifested verbally (e.g. verbal rages) and/or behaviourally (e.g. physical aggression toward people or property) that are grossly *out of proportion* in intensity or duration to the situation or provocation [emphasis added].
>
> Anxiety disorders. Since individuals with anxiety disorders typically overestimate the danger in situations they fear or avoid, the primary determination of whether the fear or anxiety is *excessive or out of proportion* is made by the clinician, taking cultural contextual factors into account [emphasis added].
>
> Specific phobias. Criterion D: the fear or anxiety is *out of proportion* to the actual danger posed by the specific object or situation and to the sociocultural context [emphasis added].
>
> Social Anxiety Disorder (Social Phobia). Criterion E. The fear or anxiety is *out of proportion* to the actual threat posed by the social situation and to the sociocultural context [emphasis added].
>
> Borderline (Emotionally Unstable) Personality Disorder. Criterion B1. Emotional lability (an aspect of Negative Affectivity): Unstable emotional experiences and frequent mood changes; emotions that are easily aroused, intense, and/or *out of proportion* to events and circumstances [emphasis added].
>
> Persistent Complex Bereavement Disorder. Criterion E. The bereavement reaction is *out of proportion* to or inconsistent with cultural, religious, or age-appropriate norms [emphasis added].

This association between the proportionality of emotional intensity experienced and emotional pathology, in the foregoing categories, reflects colloquial practice. Where categories of emotional disorder do not refer to the proportionality in emotional intensity to eliciting stimuli, they are criticized for not doing so: Horwitz and Wakefield's criticism of the category of Major Depressive Disorder furnishes us with a good example of this. But what are we getting at, or trying to get at, when we view as pathological an experience of sadness that we deem to be disproportionately intense relative to the eliciting circumstances?

Emotions are experienced in varying degrees of intensity and it is a remarkable feature of our emotions that their intensity is an integral means by which they convey evaluative information. Consider the following types of fear state: mild apprehension, creeping anxiety, intense fear, full-blown terror. All these states represent the subject's circumstances as dangerous or threatening, but they do so to varying degrees. Terror signifies that the subject's circumstances pose an imminent and extremely dangerous threat, whilst mild apprehension signifies that the circumstances are unpleasant but barely threatening, posing no real risk of serious harm to the subject. Emotional intensity is integral to the way in which our emotional experiences depict evaluative information because it matters not just that we are able to identify *that* a value is instantiated in the present circumstances (for instance, that the circumstances are threatening), but also that we are able to discern the degree to which that is the case (thus, determining *how* threatening the circumstances are). Emotions seem to be unique in this way: intensity is not a means by which belief states and perceptual experiences convey information about our circumstances.[9]

Since it is an emotional experience's accuracy in depicting our circumstances that we are assessing when we appraise it as rational, and since the intensity of the emotion is an integral means by which these circumstances are so depicted, the judgement that it is rational to experience an emotion with x degree of intensity is, in effect, the judgement that that degree of intensity accurately represents our circumstances. It would be rational for me to feel very intense fear (i.e. terror) when I am in fact in a situation that is extremely and imminently dangerous, and it is rational for me to feel this way in virtue of the fact that my terror is what depicts, to me, my situation as extremely and imminently dangerous. Conversely, it would be irrational for me to feel terror in circumstances that are not in fact threatening, or else in circumstances in which the degree of threat I am confronted with is neither grave nor imminent. In this last case, we speak of my terror as being disproportionate.

Thus when we judge an emotional experience to be disordered on account of its being disproportionate in intensity to the eliciting circumstances, in part what we are doing is appraising the emotional experience as irrational. The judgement of emotional disorder thus derives (in part) from this more basic appraisal. The qualifier 'in part' is important here because the notion of emotional intensity, like that of emotional pathology which it informs, is complex. The discussion so far has focused exclusively on the cognitive aspect of this

[9] Some beliefs are held more ardently than others, but this is usually taken to refer to the fact that some beliefs matter more to us than others, and that we hold them with greater conviction. These are not aspects of beliefs by means of which they convey information (about our circumstances) to us.

notion, but if intensity were a purely cognitive notion, it would be difficult to see how emotional experiences would be any different to evaluative beliefs. The notion of emotional intensity gets its substance, arguably, in virtue of the fact that emotional experiences generally also contain affective and agential elements. Sticking with the example of fear, we can note that the experience of fear is constituted by a cognitive-affective-agential configuration, such that in experiencing fear as severe, the experience instantiates the following specific configuration:

- A cognitive element: the threat is judged to be, or perceived as, imminent and dangerous.
- An affective element: the fear experienced is extremely unpleasant.
- An agential element: the fear experienced is strongly motivating.

We refer to all these things when we speak of intense fear, and the affective and agential elements of our emotional experiences also give rise to appraisals thereof, in distinct basic evaluative dimensions. It is for this reason that the qualifier 'in part' matters: our judgements of emotional disorder derive in part from the appraisals of emotional experiences as rational/irrational, but as we shall see, they can also derive from (amongst other things) appraisals of these experiences as unpleasant, or as motivating or as crippling.

1.3.3 Emotional Irrationality: Agential and Pathological

Given the thesis that the judgement of emotional pathology is a meta-evaluative one that is anchored in part in the basic evaluation of an emotional experience as rational or otherwise, an initially tempting move would be to completely align the concepts of irrationality and pathology. According to this move, providing that they aren't outweighed by positive appraisals in other basic evaluative dimensions (such as that of existential significance), those emotional responses that are deemed irrational are on this basis judged to be disordered. Conversely, an emotional response that is deemed to be rational or reasonable is understood to be a normal, ordinary (non-pathological) response to the subject's circumstances.

However, even within the basic evaluative dimension of emotional rationality, this blanket identification of irrationality with disorderedness overlooks the fact that we do make judgements about emotional irrationality that are altogether non-health related, and in particular, that it is possible to say of an irrational emotional response that it signifies a poor exercise of one's rational

agency (rather than being no exercise of one's rational agency at all, i.e. produced by or consisting in a pathological process). In issuing the appraisal that an emotional response is irrational in this agential sense, we typically mean to speak of an agent's poor or bad judgement. This appraisal of a person's exercise of her rational agency—through the emotional response she manifests in a given situation—is anchored in the view that her emotional response embodies a perception of value that can be assessed not only as accurate or inaccurate but, crucially, as warranted or unwarranted by the circumstances. Thus, taking our emotional responses to signify an exercise of our rational agency presupposes viewing human emotion in cognitivist terms ('cognitivist' in the broad sense articulated earlier in this section). In order to distinguish the two kinds of appraisal, I shall refer respectively to the 'agential' and 'pathological' forms of emotional irrationality. Given that the notion of emotional irrationality is ambiguous between these two kinds of appraisal, it would be more accurate to say that an appraisal of emotional irrationality can *tend us towards* issuing a judgement of emotional pathology, but that the relationship between the appraisal and the judgement isn't one of sufficiency.

Delineating the boundary between pathological and agential emotional irrationality is far from a straightforward undertaking. One important reason for this is that the way we draw the boundary will depend to a large extent on the background perspective we bring to bear on our judgements of emotional irrationality. Different perspectives will yield different answers, as we will see when we consider the following two perspectives: the evolution-theoretic approach adopted by Horwitz and Wakefield (examined in Section 5.2), and the Aristotelian approach to emotional irrationality (examined in Section 4.5.2). To anticipate, I suspect that the evolution-theoretic approach threatens to collapse all appraisals of emotional irrationality into the judgement of disorderedness. By contrast, the scope of agential emotional irrationality looks to be far wider according to the Aristotelian approach to emotional rationality, and the scope of pathological emotional irrationality correspondingly much narrower.

If it is possible that the notion of emotional irrationality may be construed either agentially or pathologically (as seems to be the case on the Aristotelian view), then it would be a mistake to blankly identify emotional irrationality with emotional disorderedness. As has just been noted, it would be more accurate instead to say that an appraisal of emotional irrationality can tend us towards issuing a judgement of emotional pathology, depending on the nature of the irrationality that is manifest and on the background perspective we bring to bear on our judgements of emotional irrationality. Just as importantly,

the appraisal of a particular emotional response or disposition as pathologically irrational likely, but not conclusively, settles its pathology status because it is possible that this appraisal has to be weighed up against appraisals of that very same response in other basic evaluative dimensions (if and when the latter appraisals are made). Again, this speaks to the point that an appraisal of emotional irrationality is not in itself sufficient for judging a particular emotional response to be disordered, although it can and often does provide *prima facie* grounds for issuing this judgement.

1.4 The Second Basic Evaluative Dimension: Existential Significance

There are things we value that find expression in our emotional lives but that aren't well captured by standard notions of emotional rationality. While it may be rational to carry on working in a job one finds unfulfilling because it pays the bills, the sense of *ennui* that results from these circumstances may eventually lead to chronic anhedonia. Or again, a life in which one works towards meeting one's perceived physical needs but seeks out no more (i.e. no leisure pursuits, no meaningful friendships) may seem impoverished or empty. The subject of this life may also likely suffer chronic anhedonia, but again, it is not straightforwardly clear that this is an irrational way to live. What chronic anhedonia in both these instances point towards is the fact that our emotional lives embody and manifest certain kinds of value, values that the standard notion of emotional rationality does not seem suited to capture. There are several reasons for this. Three (interrelated) reasons are identified here:

- *Background values*. Like judgements of practical rationality, judgements of emotional rationality (or irrationality) usually presume a backdrop of ultimate values against which a certain action, or a certain emotional response, is deemed to be rational (or otherwise). Thus, against the backdrop of loving and caring deeply for one's life partner, it is rational to feel bereaved when this personal relationship ends (through death or separation). That we care about the person with whom we are related to in this way, or that we care about working in a job that inspires our enthusiasm and excitement, enables us to respond emotionally to circumstances that concern our life partner, or our job (cf. Price 2015). Because these values enable and shape our emotional responses, I'll refer to them as 'background values'. The object of judgements of emotional rationality are our

particular emotional responses, but the appraisal of our relationship to these background values themselves lies outwith the judgements of emotional rationality made about our particular emotional responses.

- *Scale and scope*. As we've just noted, judgements of emotional rationality tend to focus on emotional experiences as responses to particular circumstances characterized in terms of their formal objects (loss, danger, etc.). The timescale of these responses tends to be relatively brief, or at least temporally circumscribed. But our relationship to background values tends to encompass a much longer time span—going back into our early childhood and/or stretching forward into our old age. While our background values can and do change over the lifespan, a distinctive feature of these values is that (at least some of) these values derive from the exercise of our reflective capacity to take stock of our lives as a whole, to ask ourselves e.g. what the point of our activities has been or is, or whether important decisions we have made have been the right ones. While philosophers have cautioned us against assuming that the question of 'the meaning of life' can coherently be asked of our lives as a whole—viewed from a detached perspective—it seems fairly coherent to wonder about the point of what we've been doing from an immersed perspective, i.e. in a comparative way. We ask questions such as these: 'Have I been engaging in activities that matter most to me, activities that connect me to what concerns me most?', 'Am I in an intimate relationship that helps me live in a way I would wish to?' Taking stock of our lives in these ways often involves appraising our emotional experiences because our background values find expression in our emotions (although not exclusively). I will say more about this shortly. For the moment, the main point to note is that the standard notion of emotional rationality seems to be ill-suited to capturing these appraisals, given their global character (i.e. they pertain to the entirety of our lives).
- *Transcendence*. This is a background value that merits special attention. Many people find it important to live their lives in a way that relates them to something that transcends it.[10] This something could be an omnipotent deity, but it needn't be. It could be a sense of being situated in a universe whose nature and origins invoke awe and inspire scientific curiosity or, more locally, it could be a sense of being situated in the intricate web of life that has evolved on our planet. More locally yet, it could be a sense of belonging to a human community with a shared purpose (e.g. pursuing

[10] For a recent philosophical exploration of this claim, see the fifth chapter of Cottingham (2022).

> environmental conservation). A life in which a person cares about being in this relationship may, on the surface, look much like a life in which a person is fairly indifferent to this possibility. Yet, for the person who cares in this way, the absence of a lived relationship with the transcendent can lead to significant emotional distress. Life can seem arid and barren, pointless, perhaps even anxiety-provoking. Again, because of the global scope of this background value or concern, the standard conception of emotional rationality is not suitably shaped to capture the appraisals we make of these distressing emotions.

So in addition to the dimension of emotional rationality (and the other three dimensions), it is helpful to recognize a further basic evaluative dimension along which we appraise our emotional experiences. I shall refer to this as the dimension of *existential significance*, since the values that are expressed through these emotional experiences are ultimate and/or have a global scope, i.e. they relate to our existence as a whole.

The relationship between appraisals of our emotions in the dimension of existential significance, and the meta-evaluative judgement of emotional disorder, can now be stated as follows. Where we are disconnected from what we value in this ultimate sense, this disconnection typically manifests itself in emotional distress—e.g. in anxious or depressed states. When a distressing emotional experience is taken to be a manifestation of this disconnection, we appraise it positively by judging the experience to have existential significance. A negative appraisal in this dimension consists in judging that the emotional distress bears no existential significance. We will explore this basic evaluative dimension at greater length in the third chapter.[11]

1.5 The Third Basic Evaluative Dimension: Affective Valence

It is a distinctive feature of our emotional experiences, one that sets them apart from other information-providing states such as beliefs and perceptual experiences, that many if not most such experiences have a feeling tone and

[11] The evaluative dimension of existential significance occupies a distinctive position in discourse about emotional pathology: it provides the broadest source of resistance to the meta-evaluative judgement that an emotional experience is pathological. To put the point in terms of an image, if the domain of emotional pathology is drawn most widely by Kendler, it is drawn somewhat more narrowly by Wakefield, and tends to be drawn even narrower still by those who incorporate evaluations of emotional experience as existentially significant.

that this tone is regarded broadly as pleasant or unpleasant. I shall refer to this feeling tone as the *affective quality* of an emotional experience, and to our appraisal of this tone as pleasant or unpleasant as the experience's *valence*.[12]

Whilst the fact that our emotions manifest a feeling tone is a familiar point in the empirical literature, the focus there has been largely if not squarely on the affective quality of our *occurrent* emotional episodes, typically spanning seconds or minutes. Cognitive psychologists James Russell and Lisa Feldman Barrett for instance identify the feeling tone of these episodes as being one of two dimensions of the 'core affect' of an emotional experience (the other dimension is that of arousal or activation—this latter dimension pertains to the volitional aspect of our emotional experiences, to be discussed in Section 1.6). The notion of 'core affect' in this context refers to 'the most elementary consciously accessible affective feelings' in our emotional experience, and Russell and Feldman Barrett suggest that both the dimensions that make up an emotion's core affect—valence and arousal—have a bipolar structure, comprising the poles of pleasantness/unpleasantness in the case of valence (Feldman Barrett and Russell 1998, Russell and Feldman Barrett 1999).[13]

This focus in the empirical literature on the affective quality of our occurrent emotional episodes stems from researchers'—including Russell and Feldman Barrett's—concern to formulate a valid and reliable terminology that will aid the empirical study of our emotions, an approach that tends to centre upon emotions experienced episodically (since these lend themselves most readily to observation under experimental conditions). As my concern in this book is with the concept of emotional pathology, rather than with the development of a metric that will facilitate the empirical study of human emotion, and because the concept of emotional pathology cuts across all levels of human emotional experience (as we shall see), it will be helpful in light of our present concern to note that affective quality appears to manifest not just locally in our occurrent emotional episodes, but also more globally in our dispositional affective traits.[14] Recognizing this enables us to appreciate that

[12] For 'valence' terminology, see e.g. Robert Solomon (2006), Lecture 15.

[13] The dimension of arousal instantiates a bipolar structure defined in terms of activation, on one end, and on the other, deactivation. It should be noted that since Russell and Feldman Barrett's work is aimed at uncovering the underlying structure of emotion in a manner that is both comprehensive yet economical, their proposal is embedded within a rich debate about whether the proposed two dimensions are indeed sufficiently comprehensive, and indeed whether bipolarity is the right way to think about either of these dimensions. However, whilst important, this debate is orthogonal to our current concern about the appraisals we make of our emotional experiences with regard to their valence quality—and the way these appraisals inform the judgement of emotional pathology.

[14] This claim calls for careful elaboration and study, which will take us beyond the parameters of the present inquiry. Still, it will be useful briefly to say the following: it seems to me that at least one way in which affective quality is made manifest in our dispositional affective traits is through the (regular) instantiation

affective quality is an aspect of emotion that can span rather longer time scales than the mere seconds or minutes of an occurrent emotional episode, which in turn will allow us to discern any broad practical significance that stems from the fact that we tend to appraise the affective quality of our emotional experiences in terms of pleasantness/unpleasantness.

Our conceiving of the affective tone of our emotional experiences in terms of this bipolarity does indeed seem to shape our lives in a wide-sweeping way. To fully grasp the practical significance of this claim of bipolarity, I want to begin by noting that it provides us with a helpful framework for understanding how our appraisals along the dimension of affective valence inform our judgements of emotional pathology. This is principally because the idea that the affective quality of our emotional experiences instantiates a bipolar structure constituted by the poles of pleasantness/unpleasantness appears to be exactly reflected in the appraisals we make of our emotions along the basic evaluative dimension of affective valence. The general tendency of sentient creatures to avoid painful stimuli and to seek out pleasurable ones informs the way we view the pleasantness or otherwise of our emotional experiences: experiences with a pleasant valence are spoken of as 'positive emotions' and those with an unpleasant valence are deemed to be 'negative emotions'.[15] This in turn feeds directly into the health/pathology distinction with respect to our emotional experiences: judgements of emotional pathology are generally underpinned by the appraisal of an emotional experience as negative (on account of its unpleasantness), whilst emotional experiences deemed positive (on account of their pleasantness) are generally associated with the idea of emotional health.

of these traits in a subject's occurrent emotional episodes (i.e. instantiations in virtue of which the ascription of a dispositional affective trait to the subject is warranted). The core affect that Russell and Feldman Barrett refer to are a component of these recurring episodes. Further, however, the affective quality of these dispositional traits—made manifest through the regular instantiations of these traits in occurrent emotional episodes—may over time 'colour in' the lifeworld of the subject whose traits they are (this point is especially visible in emotional experiences that tend to get diagnosed psychiatrically, where a mood—usually appraised as unpleasant—endures over time, becoming a baseline state for its subject). Just as occurrent emotional episodes have a core affective component that is structured around the pole of pleasantness/unpleasantness, so too it seems does the 'colouring in' of our lifeworld by the emotions we are prone to experience regularly and that constitute our dispositional affective traits. This too seems to have an affective component structured around the pole of pleasantness/unpleasantness. I recognize that this latter claim is speculative and calls for further investigation, and I have only tentatively included mention of it here because of its importance.

[15] That pleasant emotional experiences are appraised as positive, and unpleasant ones appraised as negative, is an assumption that constrains inquiries into the emotions. For instance, it constitutes a guiding assumption within the Positive Psychology paradigm. A critical commentator describes this normative schema as follows: 'For Seligman [a founding father of the Positive Psychology movement], positive emotions . . . are those that feel good. So love would be positive, anger and grief negative, and so forth' (Nussbaum 2008: S92).

I shall return to this point shortly. For the moment, I want to continue to trace out the practical significance of conceiving of the affective quality of our emotional experiences in these terms. It is difficult to overestimate the importance we place on the affective quality of our emotional experiences. This evaluative dimension has come to shape the way we think about what it means to live well. In modernity, the association between flourishing and experiencing a preponderance of pleasant subjective states is expressed most boldly in Benthamite utilitarianism,[16] and it continues to hold a firm grip on our understanding today of what it means to flourish. The prominence of positively valenced emotional experiences in our contemporary understanding of what it means to flourish is reflected in the importance placed—by empirical researchers studying the concept of flourishing—on the idea of subjective well-being. One author writes:

> Many philosophers and social scientists have concerned themselves with defining happiness or well-being . . . . A third meaning of happiness *comes closest to the way the term is used in everyday discourse*—as denoting a preponderance of positive affect over negative affect . . . . This definition of subjective well-being thus stresses pleasant emotional experience. This may mean either that the person is experiencing mostly pleasant emotions during this period of life or that the person is predisposed to such emotions, whether or not he or she is currently experiencing them. (Diener 1984: 543, italics added)[17]

In a similar vein, the World Database of Happiness advertises itself as 'an archive of research findings on subjective enjoyment of life', herein presupposing as self-evident that happiness consists in enjoying a life permeated with positively valenced emotional experiences.[18]

[16] Jeremy Bentham writes: 'Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne. They govern us in all we do, in all we say, in all we think: every effort we can make to throw off our subjection, will serve but to demonstrate and confirm it. In words a man may pretend to abjure their empire: but in reality he will remain subject to it all the while. The *principle of utility* recognizes this subjection, and assumes it for the foundation of that system, the object of which is to rear the fabric of felicity by the hands of reason and of law' (Bentham 1982, originally published 1781).

[17] Although researchers in this tradition acknowledge that flourishing also depends, to some degree, on what they call 'objective' conditions (the latter include conditions as disparate as health, comfort, virtue, and wealth), they tend to equate flourishing with subjective well-being, and this is justified by the idea that the objective conditions shape flourishing indirectly—that is, their contribution to flourishing is taken to be mediated by 'subjective processes' (Diener 1984: 552).

[18] https://worlddatabaseofhappiness.eur.nl/ (retrieved 17 July 2020). At a level of grain finer than that of flourishing, conceived of as a global assessment of individual well-being, it is revealing of the value we place on experiencing positively valenced emotions that psychologists speak of 'emotional regulation' mainly if not exclusively with reference to emotional experiences that are regarded as having a negative valence, e.g. anxiety, anger, shame (https://www.psychologytoday.com/gb/basics/emotion-regulation); Thompson 1994.

The weight we place on the valence quality of our emotional experiences (as reflected in our intuitions about human flourishing) gives rise to an important understanding of what it means to suffer or to languish. According to this notion of suffering, we tend to think of someone whose life is permeated with negatively valenced emotional experiences as a person who languishes or suffers. Again, the notion of emotional intensity appears here: the more intense these negatively valenced emotional experiences are, and the longer they are experienced for, the greater the suffering is deemed to be. This, combined with the foundational aim in healthcare of reducing or alleviating suffering, goes some way towards explaining why the valence quality of emotional experiences constitutes an important basic evaluative dimension upon which judgements of emotional disorder are based. In the fifth edition of the DSM, we see explicit reference being made to this basic evaluative dimension in the criteria for Major Depressive Disorder, i.e. markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (Criterion A.2). But this evaluative dimension arguably shapes and informs the conception of disordered affect in the DSM in a far more global way. With the exception of mania, all the affective conditions that appear in the DSM are usually deemed to have a negative valence quality—most notably the depressive and the anxiety disorders, but also elsewhere, e.g. irritability and anger in bipolar disorder (and its corollary in children and adolescents, 'disruptive mood dysregulation disorder', which is characterized by 'pervasive irritability and intolerance of frustration'), as well as intense anger, chronic feelings of emptiness, and irritability and anxiety in Borderline Personality Disorder.[19] It is by focusing principally on the evaluative dimension of affective valence (together with the dimension of personal-level agential efficacy) that Kendler arrives at the view of intense sadness and intense anxiety as pathological.

The polarities in this evaluative dimension (pleasantness as meriting a positive appraisal/unpleasantness as meriting a negative appraisal) seem to map neatly onto judgements of emotional health and disorderedness. Taken in isolation, i.e. without reference to appraisals made in the other basic evaluative

[19] A remark about mania: taken on its own, mania especially in its milder forms may appear to be the exception that challenges the claim being put forward here—namely, that only unpleasant affective experiences (as opposed to pleasant ones) are deemed pathological, and are judged so exactly on account of their unpleasantness. This appearance is arguably misleading: our very conceptualization of elevated or expansive states (as the DSM-V puts it) *as* episodes of mania arises in virtue of their close association with depressive states, i.e. standing in a cyclical relationship to the latter. When such elevated or expansive states occur without this association, then I suspect we generally don't recognize or conceive of these states as states of mania. Of course, the more severe a manic state becomes, the more distressing it is inherently experienced as being—due to the irritability, psychomotor agitation, and prolonged sleeplessness suffered—and thus the more likely the affective experience therein will on its own terms (i.e. without necessarily bearing any relationship to depressive states) be appraised negatively on account of its unpleasantness.

dimensions, emotional experiences with a negative valence are deemed pathological, whilst emotional experiences with a positive valence are deemed healthy. However, as is the case with emotional experiences negatively appraised in all the other basic evaluative dimensions, here again we ought to recognize that there exists a gap that prevents a straightforward alignment between negative valence and pathology, on the one hand, and on the other, between positive valence and health. The fact is that it is possible not to judge as pathological an emotional experience that is deemed to be unpleasant (even deeply unpleasant, to the point of being extremely distressing), because the negative appraisal of the experience in this evaluative dimension can be trumped by a positive appraisal of that very experience in some other evaluative dimension. I will return to this in Section 1.8. However, as I shall note in the fourth chapter, by conceiving of human flourishing exclusively in terms of subjective well-being, we are encouraged to place significant weight on negative appraisals of our emotional experiences in the dimension of valence. Indeed, Kendler's inclination to pathologize distressing emotional experiences on account of their negative valence quality looks to be bootstrapped by a particular conception of human flourishing that accords primacy to subjective well-being (I shall refer to this as the 'health-based conception of flourishing' and turn to look at it in greater detail in Section 4.2).

1.6 The Fourth Basic Evaluative Dimension: Volitional and Agential Efficacy

In addition to providing us with evaluatively significant information about our circumstances and being experienced as pleasant or as distressing, emotional experiences are also motivating. They do this in two ways. Firstly, in virtue of making evaluations about our social and physical world, our emotional experiences—and affective responses more broadly—serve as an attentional beam that highlights certain features of a given set of circumstances, making them salient to us.[20] Secondly, through their valence quality and/or through the evaluations they issue, our emotional experiences (and affective responses more broadly) move us to act and, equally, they inhibit us from acting. This striking feature of affective experience can be discerned across the entire hierarchy of agency, from the most basic level of instinctive responses, through to levels of intentional agency, and all the way up to full-blown autonomous

[20] Sherman 1997: 73.

agency. It is in order to capture the width of this spectrum that I use the term 'volitional efficacy' to refer to the appraisals we make of our affective experiences all along the spectrum in this evaluative dimension: 'volition', as the term is being used here, refers not just to the narrower class of voluntary action, but is intended to include the far wider class of 'endogenous action, particularly goal-directed [purposive] endogenous action, shared by humans and some other animals' (Haggard 2019). At the broadest level, affects motivate or inhibit their subjects brutely and immediately. Intense fear that is experienced in response to being pursued by a predator can directly motivate an animal to flee (the flight response) or to become immobilized (the freeze response). By contrast, the term 'agential efficacy', as I use it here, picks out a special class within the broader category of volitional efficacy, referring just to intentional and/or autonomous actions undertaken by agents who are endowed with the capacity for reason. Within this narrower domain of agential efficacy, emotions can motivate not merely brutely (although they can continue to do this, and this is most noticeable when they interfere with or thwart our pursuit of our intentional goals), but crucially, by providing an agent with reasons for acting in particular ways.

Because affects are a source of motivation across various levels of agency, we can understand them as enabling us to meet our basic needs, to satisfy our preferences, and to realize our intentional goals and purposes (when they motivate), and conversely, as hindering us from achieving all these things (when they inhibit). To unpack this claim somewhat: when an affective experience is motivating, and when this serves to satisfy its subject's needs, preferences, or goals, it is judged in positive terms. Conversely, when the emotional experience is inhibiting, and when this prevents needs being met (or goals being realized), it is judged as negative. An intensely negative emotional experience in this sense may cripple or immobilize its subject altogether. However, the converse is also possible: an inhibiting response can serve a protective function, and a motivating response can jeopardize its subject in some way or other. This complexity in the way we appraise the motivating/inhibiting aspect of our affective experiences underscores the point that volitional efficaciousness is only ever characterizable as such, and thus appraisable positively or negatively, from a determinate reference point. When it comes to non-rational animals, the perspective it makes most sense to adopt is an impersonal one that refers to the survival (and reproductive) advantage accrued by an organism through the volitional efficacy of its affective responses. Drawing on either a systems-theoretic or an evolution-theoretic approach, it seems plausible to say that affective capacities in a species (along with all other capacities manifest by members of that species) function to enable individuals to

respond to the environment in ways that enhance their survival and reproductive chances—and it is from this perspective that the motivating/inhibiting aspect of an organism's affective responses are appraised negatively or positively. Very generally, affective responses resulting in volitional changes that are biologically advantageous or beneficial are positively appraised, and those resulting in volitional changes that are biologically disadvantageous or harmful are negatively appraised. I shall refer to these as appraisals made in the basic evaluative dimension of *volitional efficacy*.[21]

Appraisals made in the more specific evaluative dimension of *agential efficacy* require us to refer to the personal perspective of the rational agent herself. This is because of the nature of intentional action itself: an action is only ever intentional under a particular description, and this description is the one the agent herself gives of her so acting. Any description of an action that the agent herself isn't aware of (and which she would deny doing, if made aware of it) is not a description of an action that is intentionally done—however accurate that description of that action, non-intentionally specified, might happen to be.[22] Appraisals of an emotional experience along the evaluative dimension of agential efficacy accordingly (either implicitly or explicitly) refer to the agent's own perspective. Emotional experiences that are motivating or inhibiting in a way that helps the agent to realize her intentions, or her goals, concerns, and projects more broadly, are appraised positively; conversely, we appraise in negative terms emotional experiences that motivate or inhibit in ways that thwart or undermine an agent's intentional pursuits.

Summarizing, at a lower or more basic level of agency, affective responses contribute to an organism's ability to avoid harms and to accrue benefits, where these harms and benefits are impersonally specified, i.e. without reference to an individual rational agent's point of view, but rather by reference to biological viability. At a higher level of agency, the basic reference point becomes the rational agent's own conscious perspective. It is worth noting here that appraisals along the dimension of lower-level volitional agency can pull against appraisals along the dimension of higher-level agential efficacy since rational agents are capable of engaging in projects, concerns, and activities that transcend biological imperatives. For instance, an environmentally conscious person may be passionately opposed to having children himself because he is deeply concerned about human overpopulation and its impact on

[21] On the role that affective responses have with regard to low-level volitional efficacy, see Silvan Tomkins' affect theory and, in particular, the concept of an 'affect program' (Tomkins 1981, 1984). Tomkins' affect theory has informed further developments along this theoretical trajectory, most notably Paul Ekman's theory of basic emotions (Ekman 1992a, 1992b).

[22] Cf. Anscombe 2000 (originally published 1957).

climate change and the erosion of natural habitats, say. Or to take an example of the kind we've already looked at, consider the intense, immobilizing sadness that a person is very likely to suffer as a result of becoming bereaved. From the evolution-theoretic perspective that grounds appraisals of our affective experiences in terms of their lower-level volitional efficacy, it might be argued that this immobilizing sadness serves a basic biological function, such as eliciting much-needed social support in the face of losing a primary attachment figure. On these grounds, the immobilizing sadness is appraised positively along the dimension of (low-level) volitional efficacy, and this provides *prima facie* grounds for resisting the pathologization of such sadness (as we'll see in Section 1.9.1, Horwitz and Wakefield argue in precisely this way). By contrast, this very same experience of immobilizing sadness may prevent its subject from pursuing her intentional projects, concerns, and goals—and on these grounds be appraised negatively when viewed from the angle of (higher-level) agential efficacy. This negative appraisal provides *prima facie* grounds for pathologizing the sadness.

The importance we place on this aspect of our affective experiences—on their volitional and agential efficacy—should be self-evident. Given that we care deeply about our capacity to exercise our rational agency effectively, it should come as no surprise to see that appraisals made of our affective experiences in this basic evaluative dimension feed into our judgements of emotional disorder. We see this clearly in the DSM-5's diagnostic criteria for both major depressive disorder and generalized anxiety disorder, one of which explicitly states that the affective condition in question must cause 'a clinically significant impairment in social, occupational, or other important areas of functioning'.[23] Situating the appraisals made along this basic evaluative dimension within the BED framework, we can observe the following: as a starting-point for thinking about the relationship between appraisals in this basic evaluative dimension and the meta-evaluative concept of emotional disorder, it is possible to say that a negative appraisal of an affective experience in this dimension gives us *prima facie* grounds for judging the experience to be disordered, and that a positive appraisal of the experience provides us with *prima facie* grounds for resisting this pathologization. But the emphasis must again be on the *prima facie* character of these grounds, for when it comes to forming judgements of disorder, these appraisals are seldom if ever taken in isolation. Instead, the judgement of disorder is formed as a result of negotiating appraisals made in several basic evaluative dimensions (as well as *within*

[23] This is Criterion B in major depressive disorder (APA 2013: 161), and Criterion D in generalized anxiety disorder (APA 2013: 222).

each of the dimensions themselves). As noted previously, I will explore this point at greater length in Section 1.8, once all five basic evaluative dimensions have been laid out, thereby enabling us to assemble them together.

1.7 The Fifth Basic Evaluative Dimension: Interpersonal Functioning

Human intersubjectivity is thoroughly permeated by emotion. Right from the very beginning of life (starting with the infant's bond with its primary caregiver), our interactions and relationships with others are characteristically shaped by our emotional dispositions. It is in virtue of possessing these dispositions that we are able to relate to each other in the rich and complex ways that we (as humans) do.

The idea that emotional dispositions are integral to human intersubjectivity is subject to interpretation in a number of ways. At the most general level, and similar to what we've seen when looking at appraisals of emotional experiences along the dimension of emotional rationality, the appraisals made along the dimension of interpersonal functioning are subject to being construed either agentially or pathologically. Whereas we speak of *moral* emotions when construing human emotional disposition(s) agentially, we speak broadly of prosocial and antisocial emotional traits when studying them empirically—and it is from this latter perspective that the claim that antisocial traits are pathological is generally made.

Construed agentially, our emotional dispositions are taken to figure prominently in the constitution of our character. Reference to the character-constituting nature of our emotional dispositions is especially visible in the context of the moral appraisals we make in everyday life and within certain ethical frameworks (most notably the Aristotelian one). We may speak for instance of a generous or a selfish person, a kind or a callous person, a person who is humble or haughty—and each of these adjectives assesses, in moral-evaluative terms, a person's character by reference to an interpersonally significant emotional disposition. This moral evaluation derives from the unifying role that character plays in our making sense of others as they exercise their moral agency, manifest in their feelings and their actions (including their utterances). The philosopher Bernard Williams gives us a succinct articulation of this point in the following passage:

> In some cases, the relevant unity in a man's behaviour, the pattern into which his judgements and actions together fit, must be understood in terms of **an emotional structure** underlying them, and that understanding of this kind may be essential. Thus we may understand a man's particular moral remark as being, if sincere, an expression of compassion. This may then be seen as part of a general current in his behaviour which, taken together, reveals his quality of being a compassionate man; and it may be that it is only in the light of seeing him as a compassionate man that *those* actions, judgements, even gestures, will be naturally taken together at all. It is understanding this set of things as expressions of **a certain emotional structure** of behaviour that constitutes our understanding them as a set. (Williams 1973: 222, italics in original, bold face added)

Williams is reminding us here of the fact that our making sense of another person's morally relevant actions is usually if implicitly informed by our view of the kind of person we take her to be—that is, on our view of her character as this is emotionally constituted. As we will see later in Section 4.5.3, this feature of our intersubjective lives—the fact that we exercise our moral agency in part through the emotional dispositions we manifest—also plays an important role in human moral education. Observing and learning from excellent others (others whose moral excellence is manifest in their actions and their emotional tendencies) is an important resource in learning how to become better moral agents ourselves.

The agential construal of interpersonally significant emotional dispositions is complemented and filled out by the empirical study of psychological capacities relevant to emotional functioning in the intersubjective domain. Most obviously, knowledge of these capacities sheds light on the causal enabling conditions that enable us to become moral agents and to exercise this agency effectively through our interpersonally significant emotional dispositions. For instance, empirical research into the capacity for (various forms of) empathy yields interesting and important insights into how it is that humans are capable of feeling and responding compassionately to each other (where having a compassionate disposition is understood as something that is apt for moral appraisal). Here, the empirical perspective underscores the capacities that are presupposed in the exercise of our moral agency, and the agential perspective foregrounds the exercise of that agency itself.

The empirical and the agential perspectives do however seem to stand in an uneasy tension with respect to the characterization of interpersonally significant emotional dispositions themselves. Take for instance the tendency to feel little or no remorse. Construed agentially, this tendency is viewed as a morally significant character failing for which a person can be held accountable.

Indeed, this tendency (so construed) occupies an especially prominent position in our practice of holding each other—and ourselves—morally accountable. This is evident in the variety of human institutions that are organized around remorse, for instance the confessional, the establishment of truth and reconciliation commissions, and the role that the expression of remorse (or lack thereof) plays in sentencing in the criminal justice system. In this last instance, the lack of remorse on the part of the defendant is often explicitly remarked upon by judges, and where the defendant is found guilty, remorselessness tends to be treated as an aggravating factor that leads to severer sentencing.[24] One reason why remorse occupies such a central position in human systems of morality and justice is that we take it to be restorative: inherent in a wrongdoer's genuine remorse is his or her recognition of the victim's dignity—dignity that was violated through the wrongdoing—and the wrongdoer's earnest desire to express this recognition to his or her victim.[25] Committing a grave wrongdoing against another, and then failing to restore the victim's dignity, is generally viewed as a serious moral failing on the part of the wrongdoer.

But from the empirical perspective, the tendency towards remorselessness is characterized as a personality trait that is viewed as criterial of antisocial personality disorder (ASPD). Whilst there continues to be much debate about the implications of being diagnosed with ASPD for its subject's agential status (it is far from clear for instance whether ASPD is an excusing condition—and indeed whether it exempts individuals from moral accountability altogether), this pathological construal of the tendency to feel no remorse does at the very least put pressure on a straightforward agential construal of this tendency as a morally significant character failing. This has given rise to the suspicion that the pathological construal of this tendency *displaces* the agential construal (Charland 2004).

In the remainder of this book, I will not have much to say about this tension between the agential and pathological construals of the negative appraisals made in the BED of interpersonal functioning, as the issues raised by this tension are wide-ranging and go beyond the remit of the present work. However, my hope is that a general elucidation of the structure of the concept of emotional disorder (as this book seeks to offer) will provide us with a further resource for understanding and negotiating this tension. Given our

[24] In the UK, for instance, the Sentencing Council prescribes that remorse should be treated as a mitigating factor in sentencing. (See Clause 1.27 here: https://www.sentencingcouncil.org.uk/wp-content/uploads/web_seriousness_guideline.pdf, page accessed 6 May 2024.)

[25] Feldman Bettencourt (2015). See also: https://theconversation.com/how-remorse-alone-can-sometimes-change-the-past-for-those-who-have-been-wronged-89257.

immediate concerns here, the main point to note is that negative appraisals of our emotional dispositions within the basic evaluative dimension of interpersonal functioning can *tend us towards* issuing a judgement of emotional pathology—even if this judgement turns out to be a contestable one.

1.8 Putting It All Together: The Meta-Evaluative Concept of Emotional Disorder and the Boundary Problem

Having set out the five basic dimensions of evaluation along which we appraise our emotional experiences, the central claim of this book can now be stated more completely. Judgements of emotional disorder are meta-evaluative judgements that are the upshot of negotiating the more basic appraisals we make of our emotional experiences along the five dimensions. It would be too strong to assert that judgements of emotional disorder are least contentious when negative appraisals of an emotional experience are made in *all* five basic evaluative dimensions, since judgements of pathology do not always draw on appraisals in all five dimensions. A more accurate understanding of the relationship between the basic evaluative dimensions and the meta-evaluative concept of emotional disorder would be this: the judgement of emotional disorder is always anchored in *at least one* negative appraisal in a basic evaluative dimension. This judgement is not contentious when the emotional experience is in no way positively appraised, i.e. along the other basic evaluative dimensions. A judgement of emotional disorder becomes more contentious when it is appraised negatively in one (or more) basic evaluative dimension but appraised positively in another basic evaluative dimension (or a set therein). For instance, we may be inclined to judge a distressing emotional experience as pathological on account of its negative valence, but this judgement may then be challenged if it turns out that the experience is also shown to be a rational response to a set of eliciting circumstances (thus being appraised positively in the dimension of emotional rationality). A substantial part of the task of arriving at final judgements of emotional disorder consists in negotiating the appraisals made along the five basic dimensions, weighing up conflicting appraisals against each other and deciding which ones to prioritize.

The fact that we have to negotiate the appraisals made in the five basic evaluative dimensions explains why the appraisal of an emotional experience as irrational can anchor the judgement of emotional disorder without entailing it. For example, while the negative appraisal of someone's depressed mood in the dimension of rationality (because there is no apparent notable personal loss to which the mood appears to be a response) may tend us towards judging

the mood as disordered, its positive appraisal in the dimension of existential significance (because the mood is expressive of an intrapsychic conflict, say)[26] pulls us away from the judgement of disorder. In view of these competing appraisals, the judgement of emotional disorder will turn upon how we weigh them up. Prioritizing the dimension of existential significance over that of emotional rationality will anchor the judgement that the depressed mood is normal, not disordered. More broadly, the necessity of negotiating appraisals made in the five basic dimensions of evaluation provides us with an explanation of the gap that exists between negative appraisals in the evaluative dimensions, and the judgement of emotional disorder. I will return to this explanatory schema in Section 1.9.3.

Summarizing, the argument for the claim that the concept of emotional disorder is a meta-evaluative concept comprises two steps: firstly, noticing that judgements of emotional disorder are opaque, meaning that we never arrive at such judgements directly, but only ever through more basic appraisals of our emotional experiences. The second step consists in identifying these more basic appraisals, and on this basis showing exactly how they contribute to the meta-evaluative judgement of emotional disorder. That has been the main undertaking in this chapter.

Now that we have the overall shape of this account before us, we can use it to make sense of the dispute between Horwitz and Wakefield and Kendler over the boundary problem (sketched out in the Introduction). Kendler defends the pathologization of intense sadness, and of intense anxiety, without necessarily making reference to the subject's circumstances, by emphasizing the negative appraisals made of these emotional experiences in the dimensions of affective valence and personal-level agential efficacy. Horwitz and Wakefield challenge this pathologization, insisting (by invoking the concept of 'emotional dysfunction') that reference be made to the subject's circumstances, by emphasizing the dimension of representational accuracy (a component of emotional rationality) and low-level volitional efficacy, i.e. at a lower level of specification than the level that interests Kendler. We are able to see now that both Kendler's and Horwitz and Wakefield's stances are well motivated: we do in fact appraise our emotions in all these ways, and the foregoing discussion shows that there is no obvious reason why, when it comes to the concept of emotional pathology specifically, we should prioritize the dimension of representational accuracy over that of affective valence (or vice versa). For as we have seen, in contrast with our conceptualization and assessment of doxastic states such as belief and judgement, the representational accuracy of an

[26] This example will be explored at greater length in the third chapter.

emotional experience is not the only thing we care about. The way our emotions make us feel, and their contributing role in volitional and/or agential efficacy, matter just as much to us—and all these ways of mattering inform the judgements we make about the pathology status of our emotional experiences.

Recognizing this, it becomes clear to see why no straightforward resolution seems forthcoming in the dispute between Horwitz and Wakefield and Kendler about the boundary problem, for we are able to see now that each party to the dispute prioritizes a different set of basic evaluative dimensions. This is what yields conflicting answers in response to the boundary problem. In the face of this, it is insufficient merely to insist that reference is made to a subject's circumstances in judging whether or not an emotional experience is pathological (emphasizing the dimension of emotional rationality); likewise, it is insufficient merely to emphasize the acuteness of the emotional distress or the degree to which an emotional experience immobilizes its subject (emphasizing the dimensions of affective valence and agential efficacy). For each move fails to recognize the grounds on which both positions rest, i.e. upon the privileging of different but equally valid dimensions of evaluation. The obvious way out of this impasse would be to appeal not just to the appraisals made in the basic dimensions themselves, but to provide a further argument for prioritizing one's favoured dimension(s) over the other dimensions, either in general terms or in a localized, more context-dependent way (or both). Recognizing this demand is essential to appreciating the nature of the epistemic predicament we are in when we try to work out where the boundary lies between normal emotional experiences and pathological ones, and to respond to this question in a non-arbitrary way.

1.9 Two Approaches: The Evolution Theoretic Framework and the Basic Evaluative Dimensions (BED) Framework

In this section, I want to compare the BED analytical framework—and the claim therein that judgements of emotional disorder are meta-evaluative—with the evolution-theoretic approach to understanding the concept of emotional pathology. Both approaches aim at analysing this concept in a general, systematic way, thereby promising to furnish us with the resources for responding to the boundary problem in a principled manner. But there is one crucial difference between the two approaches. Naturalistic accounts of pathology in general, and evolution-theoretic accounts as an instance thereof, aim at establishing that the concept is rooted in a value-free base. Human

evaluations may be superimposed upon this base—in order to enable us to draw distinctions we find important in conceptualizing pathology—but judgements of disorder should ultimately be rooted in this value-free base. Horwitz and Wakefield (Horwitz and Wakefield 2007, 2012) employ this evolution-theoretic framework, offering an analysis of the concept of emotional disorder that is shaped in this mold.[27]

The BED framework for understanding the concept of emotional disorder, by contrast, construes the concept as wholly evaluative. Human evaluations run all the way down—through to the appraisals made in the five basic evaluative dimensions—so that it does not seem to be possible to parse the meta-evaluative concept of emotional disorder into evaluative and value-free components. Indeed, viewed from the perspective of the BED framework, the idea that it is possible to ground judgements of emotional disorder in a value-free base seems to be defensible only in a problematically circular way, or so I shall argue in this section. And to anticipate: once we have before us a fuller view of the agential understanding of human emotions—I shall develop this

[27] Horwitz and Wakefield introduce the evaluative notion of harm in order to acknowledge that there are biological dysfunctions that we do not intuitively judge to be pathological; the condition of harm is aimed at capturing just those biological dysfunctions that are so judged. But this evaluative notion of harm is superimposed on a pre-existing value-neutral base that is set out by the concept of biological dysfunction. In addressing the evolutionary perspective, my concern in this monograph will be just with the *exhaustive* use of the evolutionary perspective to delineate the boundary of emotional disorder—that is, specifically with the move to ground the concept of emotional disorder in the notion of biological dysfunction. Thus, I will not consider the use of the evolutionary perspective to shed light more broadly on emotional disorders, i.e. the use of this perspective by theorists who are agnostic or even sceptical about whether emotional disorders—such as these are identified today—are necessarily biological dysfunctions (see Garson 2022, 2019, Nesse 2000, 1990), and who thus do not attach an *inherent* boundary-delineating significance to the notion of biological dysfunction. For authors in this latter group, the evolution-theoretic framework is held to be valuable in the first instance because it purports to offer us a schema for explaining why, for better or for worse, we manifest the emotional responses that we do. The question of how the boundary of emotional disorder ought to be drawn is secondary to this concern. Indeed, at least one author in the latter camp is altogether doubtful about the instructiveness of the evolution-theoretic perspective in delineating the current boundary of mental disorder. This is principally because even if an emotional response once served a biological function for our ancestors in the Environment of Evolutionary Adaptiveness (EEA) and is hence deemed functional by evolution-theoretic lights, that very same response may not be deemed adaptive in the present day. But in spite of its being deemed non-adaptive today (and this assessment of the response may incline some, including Kendler, towards the judgement of disorder), there is no dysfunction (I will return to this point in n. 38). On the basis of this possibility, Justin Garson concludes: 'I think it's important to recognize that mental disorders don't always involve dysfunctions. To take it a step further, it's good to see that we actually don't have a clue what mental disorders are. When we have a friend or a loved one with a severe mental disorder, it can be extremely scary, perplexing, or sad. We want to understand what's going on because that understanding gives us a sense of mastery, even if a false one. It gives us a chance to pretend that we know what's happening and that, in principle, we know how to fix it. It's better to stop pretending that we really understand what mental disorders are' (Garson 2019: 173–4). Since the overarching aim of the present inquiry is to make some headway in understanding what emotional disorders are by elucidating the structure of our concept of emotional disorder—and this includes examining how substantive understandings of this concept are put to use in the task of delineating the boundary of emotional disorder—I will set Garson's approach aside and address just those evolutionary theorists like Horwitz and Wakefield who are primarily concerned with this task of delineation and who, to this end, deploy the notion of biological dysfunction.

within the context of a broadly Aristotelian framework in Chapter 4—it will be possible to see that this circularity is indicative of a more general difficulty that the evolution-theoretic framework faces in recognizing and taking into account appraisals of our emotional experiences, agentially construed.[28] It is in virtue of being limited in precisely this way that the evolution-theoretic approach holds out the promise of anchoring the concept of emotional disorder in a value-free base. But it comes at a significant cost: this restricted view of human emotion simply fails to register, let alone to shed light on, the true complexity of appraisals that we make of our emotional experiences, appraisals that bear upon our use of the concept of emotional disorder. In recognizing this complexity, as the BED analytical framework enables us to do, we see quickly that the concept is thoroughly value laden. I will present a broad comparison between the two approaches in the fifth chapter (Section 5.2). In what immediately follows, I will articulate a more localized version of this comparison.

I'll begin the more localized comparison of the two frameworks in this section by providing a brief sketch of the evolution-theoretic understanding of the concept of emotional disorder, and on this basis, I will pinpoint the problematic circularity that besets this understanding. On this basis, I will then proceed to argue that the BED framework has several key advantages over the evolution-theoretic approach. To reiterate, the upshot of the discussion in this section will be that a comprehensive depiction of the concept of emotional disorder reveals it to be an inherently value-laden one—and once we recognize this, we will be in a position to see that disputes about the boundary of emotional disorder can only be satisfactorily addressed by understanding what values are at stake in deciding whether or not an emotional experience of a specific kind is pathological.

1.9.1 The Evolution-Theoretic Understanding of the Concept of Emotional Pathology

The evolution-theoretic approach aims at articulating a naturalistic understanding of emotional pathology, rooted in the notion of emotional dysfunction. Working within this framework, Horwitz and Wakefield conceive

[28] The appraisals we make of our emotional experiences are subject to being agentially construed in all the basic evaluative dimensions bar one, i.e. the dimension of affective valence. This point has been spelt out explicitly with reference to the dimensions of rationality, agential efficacy, and interpersonal functioning; it is implicit in the dimension of existential significance—it is only through the exercise of our rational agency that we are able to grasp matters of existential concern. This last point will be developed more fully in Chapter 4.

of emotions as psychological mechanisms that have been naturally selected for. Emotional pathology consists in a failure of functioning (a dysfunction) in these emotion mechanisms.[29] Thus, construing sadness a 'loss response mechanism', Horwitz and Wakefield contend that clinical depression is disordered sadness, and this means specifically that it consists in a dysfunction in a subject's loss response mechanism. Experiencing depression in the absence of a notable personal loss is judged as pathological since the subject's loss response mechanism has been activated without an appropriate or relevant cause, indicating a dysfunction in this mechanism.

This hypothesis raises the question: why do we have a dedicated psychological mechanism for responding to loss? What is its function? The notion of function here is interpreted in biological terms, given the background evolution-theoretic framework within which it is embedded. So the question can be further specified: what is the survival value of a loss response mechanism? How does it enhance our chances of survival (as well as our reproductive fitness)? Horwitz and Wakefield make the following suggestions. Depressive reactions (withdrawal and inhibition), in the face of the loss of a primary attachment relationship, elicit much-needed social support from one's community or tribe. Thus depression functions to attract support, and our survival may depend on this support. Another function of depression, suggest Horwitz and Wakefield, is to protect its subject from aggression, after a loss of status is suffered. The negative behaviour, mood, and thought that is characteristic of depression is understood here as arising as an adaptive response to circumstances of defeat and subordination. Horwitz and Wakefield suggest a third function of sadness: it helps people to 'disengage from their investments in unreachable goals, or in goals with a low probability of success . . . . The suspension of current activity, accompanied by the intense ruminative activity that is characteristic of depression, may facilitate the difficult shift of energy to new projects or attachments' (Horwitz and Wakefield 2007: 47–50). In this way, sadness *can be* adaptive. When it is experienced as a response to loss, the characteristic cognitive and affective features of depression serve a biological function that helps to enhance its subject's chances of survival.[30] It would therefore be a mistake to construe depression as pathological when it is experienced as a response to loss.[31] By contrast, if we experience depression in spite

[29] Wakefield adds the qualification that in order to count as pathological, the dysfunction has to be deemed harmful.

[30] These suggestions about the biological function of low mood as a response to loss have been discussed in the wider literature. See e.g. Nesse 2000: 15–16, and Beck and Alford 2009: 215.

[31] As noted in n. 27, not all evolutionary theorists follow Horwitz and Wakefield in drawing this conclusion. I will say more about this move in n. 36.

of not having suffered a notable personal loss in these ways, it would appear to be the case that our 'loss response mechanism' is not functioning as it has been designed to, i.e. to enhance our prospects of survival in the face of personal loss. It is, in short, dysfunctional. This is the basis for judging depression-without-personal-loss as disordered, according to Horwitz and Wakefield.

In effect, the evolution-theoretic approach and the BED framework give us two very different ways of articulating the worry that psychiatry over-pathologizes our emotional experiences (it is this concern that motivates the resistance to Kendler's stance). Spelling out the notion of over-pathologization in evolution-theoretic terms, Horwitz and Wakefield characterize ordinary (non-pathological) emotional experiences in terms of the notion of biological function—which in turn, as we shall see, leads them to focus exclusively on an organism's basic needs. 'Basic needs', as I am using the term here, refers to those conditions that have to be satisfied in order to for an organism to achieve the biological goals of survival and reproduction—thus, survival needs (e.g. the need for nutrition, for shelter and safety, the need to belong to a community), and reproductive needs (e.g. the need to mate successfully). Horwitz and Wakefield argue that, under specific conditions, an organism's basic needs are met in and through its depressive experience. On this basis, the depressive experience is said to be biologically adaptive, thus non-pathological. By contrast (and again, as we shall see soon), the BED framework spells out the concern that psychiatry over-pathologizes our emotional experiences in a way that incorporates both the organism's basic needs and, in the case of humans, their intentional-level concerns and goals. On this analysis of the notion of over-pathologization, psychiatry wrongly pathologizes depressive experiences if, in spite of being immobilizing, such experiences are (a) in fact rational and/or bear existential significance and (b) greater weight is placed on these latter dimensions than on the dimension of volitional efficacy.

A quick terminological note before we proceed further: the terms 'basic needs' and 'intentional-level concerns and goals' are placeholders. They aren't altogether ideal because our pursuit of intentional-level goals can be (and often are) informed by our basic needs.[32] A better way of marking the distinction that I believe is needed in order to clearly articulate the evolution-theoretic

[32] But to a first approximation, we may say the following. Basic needs are impersonally specified—that is, without reference to the agent's own conscious perspective on these needs. They are specified, and are wholly specifiable, by reference to the norms of survival and reproductive viability. Intentional level concerns and goals are, by contrast, specifiable only in personal terms—that is, an adequate specification makes essential reference to the agent's own conscious perspective on these concerns and goals.

concept of emotional disorder is to speak instead of human 'first nature' and 'second nature'. But this latter distinction becomes intelligible only once we have in view the Aristotelian account of human emotion, and we will arrive at this account only in the fourth chapter. So for now I will continue to speak (somewhat clumsily) of the distinction between 'basic needs' and 'intentional level concerns'.

1.9.2 Could There Be a Non-pathological Form of Intense, Immobilizing Sadness (IIS)? A Problematic Circularity in the Evolution-Theoretic Approach

To assess Horwitz and Wakefield's proposal, I want to consider a person who, in the face of a significant bereavement, finds herself in the pit of intense and immobilizing grief. Let's call her Mary. According to Horwitz and Wakefield's concept of emotional pathology as emotional dysfunction, Mary's grief is not disordered because her loss response mechanism has been appropriately activated, and the ensuing grief serves a biological function—e.g. it elicits social support. Further to this, let's imagine that Mary is determined to get back to work because she is in the middle of an important project that she finds valuable, both for its own sake and because she values doing her job well. Within the BED framework, this will generate mixed appraisals of Mary's grief *within* the dimension of volitional/agential efficacy. On the one hand, let's grant for the sake of argument that the grief is adaptive insofar as it meets Mary's basic need for social support (as per and Horwitz and Wakefield's suggestions about the biological functions of sadness), and this will yield a positive appraisal of the grief in this dimension—specifically, that of lower-level volitional efficacy. On the other hand, the crippling effect of the grief she suffers from prevents Mary from engaging in a project that she finds personally important and meaningful and which, for this reason, she would very much like to continue engaging in. This should yield a negative appraisal of the grief experienced, now in the dimension of personal-level agential efficacy. Kendler in fact appeals to this consideration in his example of Ms. S, whom as we saw in the Introduction struggles to be the mother and the wife she would like to be in the face of a life-threatening medical diagnosis. Kendler takes Ms. S's emotional distress to be pathological, warranting a medical response, precisely because of its devastating impact on her ability to pursue her intentional-level concerns and goals. Against Kendler, Horwitz and Wakefield would deny that this distress is pathological on the grounds that her loss response mechanism has been appropriately activated, i.e. by the loss

of her health and an imminent threat of the loss of her own life.[33] They would argue, further, that the immobilization of her agency that results directly from this emotional distress functions to attract social support and is in this way adaptive.[34] This is sufficient grounds for judging Ms. S's emotional distress to be non-disordered, according to Horwitz and Wakefield.

Here, we are faced with a dispute between Horwitz and Wakefield and Kendler about the pathology status of intense, immobilizing sadness ('IIS' for brevity). The dispute centres on how widely or narrowly we should take the *scope* of the pathology concept to be with respect to IIS. Given its adverse impact on our ability to exercise our agency with respect to intentional-level concerns, *all* instances of IIS are potentially pathological, or so Kendler argues. Locating this claim within the BED framework, we can register it by noting that IIS summons a negative appraisal in the dimension of agential efficacy due to this adverse impact. (The qualifier 'potentially' denotes the possibility that this appraisal may be outweighed by a competing appraisal in another basic evaluative dimension.) On Horwitz and Wakefield's evolution-theoretic account, by contrast, only a subset of instances of IIS are pathological. There are instances of IIS that are normal, i.e. not pathological, namely those instances that are positively appraised in the dimensions of lower-level volitional efficacy and of emotional rationality (when an experience of IIS is representationally accurate). So on Horwitz and Wakefield's account, addressing the boundary problem correctly involves undertaking the task of separating out IIS in its pathological form leading to the inability to execute intentional goals, from non-pathological forms thereof.

To better understand what this task involves, we can begin by taking note of a general idea that guides Horwitz and Wakefield's evolution-theoretic approach to understanding the concept of disorder—a guiding idea that is rooted in an analogy between physical pathology and psychopathology (including emotional pathology). According to this guiding idea, when we suffer from a condition of pathology that disables us from our daily functioning, this happens specifically on account of a dysfunctional biological mechanism. Applying the idea first to the case of physical pathology, a coronary or respiratory dysfunction might for instance stop us from participating in vigorous

[33] Casting this in terms of the BED analytical framework, reference to the idea of the emotional response being appropriate indicates that it is positively appraised along the dimension of representational accuracy (a component of emotional rationality).

[34] Reference to the idea that the emotional response helps the organism get its basic needs met indicates that it is positively appraised along the dimension of volitional efficacy.

physical activities, much as we'd like to participate in such activities. Now, we could of course be prevented from engaging in vigorous physical activities for other reasons: we do not live close enough to relevant opportunities; our daily schedules are too busy; we might not have the financial resources to purchase the equipment needed, etc. But our inability to participate in these latter cases does not stem from (biological) dysfunction. Analogously, in the case of psychopathology, an agent's inability to exercise their intentional-level agency, due to suffering from IIS, may take on either a pathological or a non-pathological form, according to Horwitz and Wakefield. If IIS is activated in spite of an absence of notable personal loss, then this instance of IIS is pathological, analogous to coronary dysfunction (in the previous example). By contrast, Mary and Ms. S's IIS-induced inability to exercise their intentional-level agency is not pathological since IIS has been activated in these cases in response to a notable personal loss. Being non-pathological, these instances of being unable to exercise intentional-level agency due to suffering IIS would be analogous to the condition of being too busy or being financially limited (in the previous example). This is the basic picture of the concept of disorder that Horwitz and Wakefield offer us.

However, there is a peculiarity about this analogy between physical pathology (the coronary/respiratory dysfunction example) and emotional pathology (IIS): whereas in the former case the causes of the inability to participate in desirable physical activities are cleanly disjunctive between pathological causes (i.e. coronary/respiratory dysfunction) and non-pathological causes (i.e. financial limitations), IIS as the cause of an agent's inability to engage intentional-level concerns straddles both sides of the disjunction between pathology and non-pathology (in the conceptual space carved out by Horwitz and Wakefield). That is, the term 'IIS' appears as a common factor here. By contrast, there is no analogous common factor in the physical pathology case: there is not a term, X (equivalent to IIS), that appears both in the characterization of the coronary dysfunction as the pathological cause of the inability to participate, and in one (or more) of the non-pathological causes of the inability to participate. This means that in the case of emotional pathology only, Horwitz and Wakefield are faced with the task of separating out IIS in its pathological form leading to the inability to execute intentional goals, from non-pathological forms thereof.[35] This task may be diagrammatically represented as follows:

[35] No analogous task arises in the case of physical pathology since there is no common factor that straddles both pathological causes and non-pathological causes, which then requires sorting out.

Table 1.1 Delineating the boundary of disorder: a comparison between IIS and a concept of physical disorder

	I. Impediments to engaging in vigorous physical activity		**II. Inability to engage intentional-level concerns**	
Common factor	[None]		**Intense immobilizing sadness (IIS)**	
Cause(s)	A1. Pathological causes(e.g. coronary/ respiratory dysfunction)	A2. Non-pathological causes(e.g. time scarcity, financial limitations)	B1. Pathological IIS(= IIS in the absence of any personal loss, i.e. caused by a dysfunction in the subject's loss response mechanism)	B2. Non-pathological IIS(= IIS as an adaptive response to personal loss)
Distinguishing the two kinds of cause	[Nothing further needs to be done to draw this distinction since the two kinds of cause are cleanly disjunctive.]		?	

This table foregrounds an obvious but notable point, namely that manifestations of coronary dysfunction do not, by definition, ever signify normal functioning. Intense, immobilizing sadness (IIS), by contrast, is not an inherently pathological notion in this way—as Horwitz and Wakefield are keen to emphasize. This is reflected in the table by the fact that IIS appears as a common factor that straddles columns B1 and B2, and it is why Horwitz and Wakefield are faced with the task of delineating the scope of pathology with respect to IIS. This, recall, is a task the correct execution of which is debated: Kendler sees *all* instances of IIS as potentially pathological; Horwitz and Wakefield deem only a subset of IIS to be pathological. The subset of IIS that is deemed to be non-pathological is conceptualized by Horwitz and Wakefield as an adaptive response to loss, and discerning that there is an adaptive value in the experience of IIS turns upon focusing on the ways our basic needs are met through this experience.

The claim that an experience of IIS can be adaptive seems fine as it goes. However, Horwitz and Wakefield's judgement of (non)-pathology that rides directly on this claim is far shakier in light of the BED framework because what the latter shows is that there are at least five dimensions of evaluation that look to be germane to delineating the scope of pathology with respect to IIS, and our basic needs are registered in only two of them. To exemplify this point in the terms used in Section 1.6, Horwitz and Wakefield characterize IIS

as a loss response mechanism exclusively by reference to the basic evaluative dimension of (low-level) volitional agency, omitting reference to personal-level agential efficacy. Kendler's appeal to the disruptiveness of IIS to our intentional goals and concerns is therein left unaddressed in Horwitz and Wakefield's account of IIS (a subset thereof) as non-pathological.

How do Horwitz and Wakefield justify the exclusion of negative appraisals of IIS in the dimension of personal-level agential efficacy? Simply, it would seem, by holding a prior commitment to reading into the notion of IIS a sense that denotes biological function. This leads to their focus squarely on the way our basic needs are met through the experience of IIS, yielding the claim that IIS is biologically adaptive. In this way, Horwitz and Wakefield carve out a subset of normal (non-pathological) instances of IIS, from the total set of instances of IIS:

Table 1.2 Delineating the boundary of disorder: the evolution-theoretic characterization of IIS as pathological

	I. Impediments to engaging in vigorous physical activity		**II. Inability to engage intentional-level concerns**	
Common factor	[None]		**Intense immobilizing sadness (IIS)**	
Cause(s)	A1. Pathological causes(e.g. coronary/ respiratory dysfunction)	A2. Non-pathological causes(e.g. time scarcity, financial limitations)	B1. Pathological IIS(= IIS in the absence of any personal loss, i.e. caused by a dysfunction in the subject's loss response mechanism)	B2. Non-pathological IIS(= IIS as an adaptive response to personal loss)
Distinguishing the two kinds of cause	[Nothing further needs to be done to draw this distinction since the two kinds of cause are cleanly disjunctive].		**Restricting** the characterization of loss response **to just a subject's basic needs** (i.e. low-level volitional efficacy), excluding all reference to personal-level agential efficacy, delivers B2 (i.e. non-pathological IIS).	

This restriction in the characterization of loss response to just a subject's basic needs looks plausible only in light of the prior commitment to reading into the notion of IIS a sense that denotes biological function. Remove this commitment and the characterization of loss response seems radically incomplete

because when we think *intuitively* about Mary or Ms. S's response to loss, the fact that they are immobilized to the point of being unable to execute their intentional goals seems to be a significant aspect of this response—one that is relevant to the concept of emotional disorder (thus, in Table 1.2., giving us grounds for moving these instances of IIS from B2 to B1). That is, without this restriction to just basic needs, the distinction between pathological and non-pathological IIS threatens to collapse since *all* instances of IIS would potentially be pathological on account of their disruptiveness to personal-level agential efficacy—or so Kendler would contend, since he apportions a significant weight to this basic evaluative dimension. Indeed, Kendler takes this disruption to agential efficacy as sufficient grounds for judging the IIS suffered by Ms. S to be pathological. Whilst it may be true that Kendler is too swift in taking these grounds as sufficient, his stance nonetheless alludes to an important consideration: negative appraisals of an emotional experience along the dimension of agential efficacy are in general accepted as grounds for judging the experience to be pathological, and IIS's disruption to intentional agency summons just such a negative appraisal.

Including this aspect of IIS *within* their characterization of the loss response mechanism will, however, threaten Horwitz and Wakefield's delineation of non-pathological and pathological forms of IIS since what looked like a perfectly adaptive response when viewed from the angle of lower-level volitional efficacy—neatly offering an account of the proper functioning of the loss response mechanism, herein anchoring the concept of non-pathological IIS—now appears to be disruptive, i.e. when viewed from the perspective of intentional-level agential efficacy. This disruptiveness puts pressure on the claim that IIS, construed as a loss response mechanism, is functioning well.[36]

[36] A point of clarification about this last claim: evolutionary theorists adopting Horwitz and Wakefield's approach insist that we use the concept of biological function in a specific way, i.e. a way that pegs mechanisms to the biological advantages they have putatively conferred on a species living in a particular ecological niche—or range of niches—in evolutionary time, rather than on whatever consequences they happen to produce today. However, in insisting on this point, the evolutionary theorist has to be mindful of the possibility that rather too large a gap could have opened up between the (hypothesized) evolutionary context in which a biological mechanism's function is specified and the context in which the mechanism operates today. If the functioning of a biological mechanism becomes divorced entirely from an organism's current circumstances, this threatens to undermine the mechanism's current explanatory power and its usefulness, today, in helping us to demarcate the boundary of the pathology concept.

It is this very consideration that leads some evolutionary theorists to foreground instead the distinction between the notion of biological dysfunction and the notion of a response-environment mismatch (i.e. an emotional response that was functional in the EEA but is deemed no longer adaptive in our present environment), and to allow for the possibility that the concept of mental disorder may be rooted in judgements of mismatch, rather than judgements of dysfunction. In making this move, evolutionary theorists of this stripe hope to close the gap between the evolutionary and present-day contexts. But precisely because it entails an agnosticism about whether emotional disorders are necessarily biological dysfunctions, this move comes at the price of having to relinquish the use of the notion of biological dysfunction to delineate

Excluding all reference to agential efficacy and restricting the characterization of IIS as a loss response mechanism to just lower-level volitional efficacy allows Horwitz and Wakefield to bypass this difficulty. In this way, the restriction comes to play a vital role in helping Horwitz and Wakefield to delineate non-pathological and pathological forms of IIS.

Given this trade-off (i.e. given that taking IIS to denote a biological function thereby excludes from the characterization we give of this function any reference to intentional-level concerns), it behoves us to ask why we should want to commit ourselves to reading into the notion of IIS a sense that denotes a biological function, and in doing so, to accord the idea of IIS as biological function the primacy that Horwitz and Wakefield do. That is: why should we use the notion of biological function as an organizing principle for carving up all experiences of IIS into two kinds—those that signify good biological functioning, and those that do not—when doing so entails treating our intentional-level concerns as irrelevant to the most basic characterization we give of our emotional experience of IIS?

Here seems to be a problematic circularity inherent in Horwitz and Wakefield's delineation of the scope of pathology with respect to IIS, a circularity that doesn't arise with respect to the pathology concept of coronary dysfunction. The BED framework helps us to see why there is no circularity in the characterization of the latter concept: simply, we do not appraise manifestations of cardiovascular function in our lives in any way other than biologically, i.e. these assessments are made squarely with reference to the impersonally specified apical goals of survival and reproductive viability. This pattern of assessment grounds the fact that, in this case, the scope of pathology is coextensive with the scope of biological (dys)function, and hence why no circularity arises in delineating the scope of pathology by using the notion of biological function. By contrast, given that from a theory-neutral perspective it is not obvious that IIS *qua* pathological state is adequately characterizable by reference to just the apical goals of survival/reproductive viability and thus to our basic needs exclusively, Horwitz and Wakefield need a justification for making this restricting move, i.e. for claiming that there exists a sense of IIS that denotes biological function and thus for limiting our conception of IIS-as-pathological in this way. The need for this justification is made pressing by the fact that IIS intuitively has a significant impact not just on our basic needs but also our intentional-level

the scope of pathology and herein looks to threaten the naturalistic programme in analysing the concept of emotional disorder, at least as this is understood by Horwitz and Wakefield (since judgements of mismatch look to be considerably more evaluative as these judgements are no longer taken to be constrained wholly by the notion of biological function, determined strictly in evolution-theoretic terms).

concerns (concerns that may very well transcend biological imperatives). But that justification seems to consist ultimately in appealing to the idea that privileging our basic needs is what will deliver a biological construal of IIS (according to which IIS can be adaptive), which in turn will enable us to discern a non-pathological form of IIS.[37]

Summarizing, the evolution theoretic framework for understanding the concept of emotional pathology requires us to endorse a tight circle between (a) construing IIS as a biological notion that allows us to discern a non-pathological form of IIS and (b) privileging basic needs exclusively. As we shall see in Section 5.2.1, this tight circle reflects a systemic limitation in the evolution-theoretic approach to analysing the concept of emotional disorder: it threatens to disregard a wide and significant range of appraisals that we would ordinarily (pre-theoretically) take to bear directly upon the judgement of emotional disorder. The approach struggles particularly to bring into view appraisals of our emotional experiences, agentially construed. Sticking for the time being to the tight circle just described, we may note the following point: outside this circle imposed by the evolution-theoretic approach, the negative appraisal of IIS along the dimension of agential efficacy—on account of its disruptiveness to intentional-level concerns—does in fact seem potentially to ground the judgement of IIS as disordered. Recognizing this everyday use of the concept of emotional disorder motivates a very different conception of IIS, i.e. with all instances of the inability to execute intentional goals on account of IIS potentially being construed as pathological in form. (A reminder: the qualifier 'potentially' denotes the possibility that this appraisal may be trumped by a competing appraisal in another basic evaluative dimension.) The BED framework enables us to register this fact about the way the concept of emotional disorder is put to use in daily life.

[37] A more general version of this difficulty will be presented in Section 5.2.1. To be clear: it isn't open to Horwitz and Wakefield to suggest that the detrimental impact of IIS on agential efficacy, and our intuitions about its relevance to judgements about pathology, can be taken care of within the *harm* condition of the HD analysis of the concept of disorder—and thus for Horwitz and Wakefield to try and capture our pre-theoretical intuitions about IIS in this way. This is because, in determining whether or not a condition is disordered, the harm condition becomes relevant and applicable only *once* the dysfunction condition has been satisfied. In the set-up that is being sketched out here, IIS in its non-pathological form would *not* satisfy the dysfunction condition, and thus even if IIS is disruptive at the level of agential efficacy and is said to be harmful for this reason, the harm that ensues would have to be of a non-disordered kind (according to the terms set out by Horwitz and Wakefield's HD analysis of the concept of disorder). And what Kendler's stance draws our attention to is the fact that there is an expectation that the concept of emotional pathology is sensitive to the disruptions to agential efficacy caused by emotional upheavals.

1.9.3 Comparing the Two Approaches: Scope

All this is not to exonerate Kendler's response to the boundary problem, according to which every instance of IIS may be judged to be pathological. Within the BED framework, a negative appraisal of an emotional experience in the dimension of agential efficacy (which potentially anchors the judgement of pathology) can be trumped by a positive appraisal in another dimension—emotional rationality, or existential significance, say. In considering Mary's experience of IIS, the fact that the IIS is an expression of bereavement entails a positive appraisal of her experience in exactly these latter dimensions. Alluding to both appraisals, the anthropologist Arthur Kleinman describes the grief he experienced when his wife died:

> My grief, like that of millions of others, signalled the loss of something truly vital in my life. This pain was part of the remembering and maybe also the remaking. It punctuated the end of a time and a form of living, and marked the transition to a new time and a different way of living. The suffering pushed me out of my ordinary day-to-day existence and called into question the meanings and values that animated our life. The cultural reframing—at once subjective and shared with others in my life-world—held moral and religious significance. (Kleinman 2012: 609)

The awful experience of grief reveals to us the value of our most intimate relationships in an experientially vivid and non-negotiable way. In grief, we feel our entire existence being plunged into the bottomless chasm of personal loss. This total decimation of the self is the measure of how deeply we loved. It is for this reason that the experience of grief is construed as an appropriate reaction to the loss of a loved one, meriting a positive appraisal in the dimension of emotional rationality.

It also merits a positive appraisal in the dimension of existential significance. In coming to see how deeply this relationship mattered to us, and the way in which it mattered (in the experience of grief), we come to understand something about ourselves too. Our identification of ourselves as spouses, parents, or children (for instance) moves beyond a mere acknowledgement of the biological facts when we feel a vulnerability and tenderness towards someone that can only be described as love, a vulnerability that leaves us raw with the searing pain of grief when we lose the person through separation or death. Love and grief transform mere biological and factual truths into existentially significant ones. In this vein, Robert Solomon notes:

> Grief is much more than a 'cognitive' response, if by this we mean only the recognition of a serious loss. For grief is the emotion, much more than Heidegger's celebrated angst, that puts us in touch with our mortality. It also puts us in closer touch with love. (Solomon 2004: 79)

It is in these ways that Mary's experience of IIS may be positively appraised along the dimensions of emotional rationality and existential significance. Further, these appraisals may be deemed to be of greater importance than the negative appraisal of this experience in the dimension of agential efficacy. If it is so viewed, then this constitutes grounds for judging Mary's experience of IIS to be normal (non-pathological). Generalizing, the BED framework enables to see that negative appraisals in the dimension of agential efficacy can be trumped by positive appraisals along other dimensions. This is why Kendler's assumption—i.e. that the immobilizing impact of intense sadness on a person's intentional agency is sufficient for the judgement of pathology—seems too simplistic.

We are now in a position to see that the BED framework provides a comprehensive view of the concept of emotional pathology, within which the dispute between Horwitz and Wakefield and Kendler can be situated and better understood. Kendler focuses just on the dimensions of agential efficacy and affective valence, in judging Ms. S's experience of IIS to be pathological; the BED framework registers the negative appraisals that Kendler makes of this experience along these two dimensions. But the BED framework *also* brings into view the dimensions of emotional rationality and existential significance—and both the latter dimensions of evaluation are just as relevant to the judgement of emotional pathology as are two dimensions Kendler privileges. The evolution-theoretic approach deployed by Horwitz and Wakefield leads them to foreground the dimensions of representational accuracy (a component of emotional rationality) and low-level volitional efficacy, where the view of this latter evaluative dimension is limited to just an organism's basic needs, to the exclusion of intentional-level agency. Again, the BED framework allows us to register this move, but *also* to see that this restriction prevents Horwitz and Wakefield from acknowledging and incorporating an aspect of intense sadness that is intuitively important to the concept of emotional pathology, i.e. its devastating impact on our ability to engage our intentional concerns and projects.

This comprehensive view of the terrain—offered by the BED framework—helps us understand why it is so difficult to know how to respond to the dispute between Horwitz and Wakefield and Kendler. Any direct response that simply endorses either position looks to be arbitrary: both parties seem, after

all, to have latched on to considerations that are salient to the concept of pathology. But we can now also see that such a response would be arbitrary because each party adopts a partial view of the concept of emotional pathology, making it difficult for us to know how to relate the two positions to each other. Horwitz and Wakefield and Kendler appear to be talking past one another. In asserting that the experience of IIS is pathological, Kendler doesn't engage Horwitz and Wakefield's appeal to the representational accuracy of IIS or its volitional efficacy (at the level of basic needs); he simply insists that the disruption of intentional agency and the negative valence is more important to the judgement of emotional disorder. Conversely, in rejecting the claim that IIS is (necessarily) pathological, we've seen why Horwitz and Wakefield cannot engage Kendler's emphasis on the disruptiveness of IIS to an agent's intentional concerns and goals. The BED framework enables us to situate both positions within an integrated conceptual space. This gives us a way of thinking about *how* we might move forward in this dispute: the crucial further step is for each party to provide a justification for prioritizing their favoured basic appraisals over the basic appraisals favoured by the other party to the dispute.[38]

If this is the right way to think about the dispute over the boundary problem with which Horwitz and Wakefield and Kendler are engaged, then an important conclusion follows: the judgement of emotional pathology cannot rest on a value-neutral base. It necessarily involves deciding between a range of values, expressed in the appraisals of our emotional experiences in the five basic evaluative dimensions. In the example of grief, we have to choose between fully recognizing the significance of personal loss (along the dimensions of emotional rationality and existential significance) and continuing to be able to exercise our intentional agency effectively (along the dimension of agential efficacy). The idea that our emotions serve biological functions gives rise to the appearance of a value-neutral base upon which the concept of emotional pathology can be based. But this appearance turns out to be illusory. It may be true that at least some of our emotional capacities have evolved as part of the functional machinery that has enabled us to survive and to reproduce (a claim that is reinforced by the fact that emotions are observed in all but the most primitive sentient organisms). But the foregoing discussion shows that it would be hasty to conclude on this basis alone that the concept of emotional *pathology* ought to be understood exclusively in evolution-theoretic

[38] In Chapter 4 (pp. 99–162), I suggest that the way we negotiate the appraisals made of our emotional experiences in the basic evaluative dimensions—and our application on this basis of the concept of emotional disorder—is informed and guided to a large extent by which background conception(s) of human flourishing we adopt.

terms—or that it is even possible for it to be so understood. For this concept plays a rich and complex role in human life, deriving from a wide range of appraisals that we make of our emotional experiences, not all of which are readily capturable in terms of the notion of biological function. Summarizing, the main claim in this section has been this: we obtain a more complete understanding of the concept of emotional pathology by recognizing its thoroughly evaluative character.

1.9.4 Comparing the Two Approaches: The Case of Phobias

I turn finally to discuss the example of phobias briefly, because this is one aspect of our emotional lives that the evolution-theoretic approach is taken to be especially useful in illuminating. Accepting this claim, I want to suggest that even here, the evolution-theoretic response to the boundary problem draws on the terms set out by the BED framework.

Phobias are deemed to be irrational, but whether they should be judged as pathological is a subject of debate, especially given how relatively common they are. Using the terms of the BED framework, we can say that there appears to be a gap between the negative appraisal of phobias in the dimension of emotional rationality and the meta-evaluative judgement of emotional disorder: the appraisal of a phobic experience as irrational does not necessarily entail the judgement that the experience is disordered. Currently the most influential way of accounting for this gap is in evolution-theoretic terms. In view of the conceptualization of emotions as mechanisms that serve a function in enhancing our adaptive fitness, it is suggested that although a phobic response may seem to us to be an irrational one, it may nonetheless have given its subject an advantage in survival and/or reproduction. If the phobic response occurs with sufficient frequency to be deemed statistically normal, the claim that it is adaptive is reinforced by the further claim that the response had a role in promoting species fitness. According to a concept of emotional pathology that is defined in terms of the notion of a dysfunctional mechanism, the phobic emotional experienced is judged not to be pathological. Horwitz and Wakefield develop this idea using the example of the fear of snakes:

> Fear of snakes, to the degree that it is a biologically prepared sensitivity transcending reason and a quickly triggering fear based on naturally selected mechanisms, seems to have no built-in distinction between poisonous snakes that are reasonable to fear and non-poisonous snakes that are not. The prepared fear seems biologically designed to be of snakes, period—biological design appears in this case

> to have erred on the side of caution, in accordance with the 'smoke detector principle'—so such fears are not inherently abnormal despite their frequent unreasonableness in light of the actual nature of the feared snake . . . . This is not to say that there cannot be snake phobias that are true disorders; even a naturally unreasonable fear can inflate into a disabling fear that, for example, keeps one from moving about even in snake-free environments. But the unreasonableness is not evidence of a failure of biologically designed functioning, thus is not sufficient for disorder. (Horwitz and Wakefield 2012: 124)

The suggestion that snake phobia is an emotional response that serves a biological function consisting in enabling us to respond adaptively to snakes sounds plausible, especially given the concrete and specific (non-abstract) stimulus that elicits our fear response, i.e. snakes (and things that look like snakes). However, I want to suggest that this evolution-theoretic account of the gap between irrationality and pathology is in fact a particular instantiation of a more general form of explanation of this gap—and that it is worth recognizing as such, in order to appreciate the generality and the primacy of this explanatory form.

According to this more general form of explanation, made lucid by the BED framework, the irrationality of a snake phobia is not yet sufficient grounds for judging the phobia to be disordered because the phobia is positively appraised in the dimension of lower-level volitional efficacy. The action-tendencies inherent in the experience of fear (flight or fight) enable its subject to avoid a potentially lethal snake, or if this is not possible, to attack it. It is interesting that Horwitz and Wakefield recognize that the fear can become so disabling that it prevents its subject from exercising her agency as she'd wish to even in snake-free environments, and that this may provide us with grounds for judging the fear to be pathological. The BED framework explains why. As the phobia increases in intensity (to the point where it could become immobilizing and therein becomes maladaptive), this increases our tendency to appraise it negatively in the dimension of volitional and agential efficacy. This negative appraisal competes with the earlier positive appraisal in this dimension (according to which the phobia of snakes was understood to be adaptive). The greater and more crippling the phobia, the more significance we will place on the negative appraisal (relative to the competing positive appraisal), and this may carry on all the way to the point at which we judge the fear to be disordered.

In this way, the evolution-theoretic account can be understood as drawing on the terms set out by the BED framework. In the case of phobias, this

deployment of the evolution-theoretic framework may come at no great cost. Given the relatively circumscribed nature of phobias (i.e. the specificity of the eliciting stimulus and behavioural response), the move to identify a biological function for the phobia seems fairly uncontentious. This approach also has the advantage of promising to explain in naturalistic terms why phobias exist. But with reference to the concept of emotional pathology at large, and in the task of determining where to set the boundary between normal and pathological emotional experience, it is important to recognize this evolution-theoretic account of the gap between irrationality and pathology for what it is, i.e. a particular instantiation of the more general explanatory form made lucid by the BED framework. As a particular instantiation, it may well be that the evolutionary account does not generalize too well, i.e. it may not work as well in explaining the gap between irrationality and pathology in other types of emotional experience besides phobias. The more complex and multifaceted an emotional kind is, the harder it will be to squeeze it into the mold of biological function, and thus the harder it will be to articulate a plausible account of the gap between irrationality and pathology in these terms.

Having set out the claim that judgements about emotional disorder are meta-evaluations that consist in weighing up and negotiating the appraisals made of our emotional experiences in more basic evaluative dimensions, the next two chapters will look at attempts to uphold the pathologization of distressing emotional experiences in the face of this claim. Chapter 2 examines the attempt to defend the pathologization of distressing emotions by circumventing the BED framework altogether. Then in the third chapter, I will turn to consider the attempt to defend the pathologization of distressing emotions within the BED framework.

2
The Medicalization of Emotions (I)

Circumventing the BED Framework

> Because our field deals with fundamental questions of what it means to be human, psychiatry is particularly susceptible to preconceptions that can strongly color the value we assign to differing methodological perspectives. With the growth of neuroscience and molecular biology, psychiatry is set to inherit rich insights into the basic workings of the human brain. To maximally use this new information, however, will require that we have our conceptual house in order.
>
> **Kendler (2005): 433**

2.1 Introduction

In the last chapter, I argued that judgements of emotional disorder are meta-evaluative judgements that are based upon appraisals we make of our emotional experiences along five more basic dimensions. I referred to this as the BED analysis of the concept of emotional pathology. This framework explains the opacity of judgements of emotional disorder: it doesn't seem possible directly to judge an emotional experience to be disordered; it is only through prior appraisals of these experiences—appraisals that do not themselves invoke the concept of pathology—that we arrive at this judgement.

In this chapter and the next, I want to consider two strategies for defending the pathologization of distressing emotions in light of this claim. These strategies consist respectively in negotiating the basic evaluative dimension (BED) framework, and seeking to circumvent the framework altogether. I will explore the negotiation strategy in the next chapter. In the current chapter, I explore the circumventing strategy for defending the pathologization of distressing emotions. To defend pathologization in this way, the proponent of pathologization will have to establish that disordered emotions wear their

The Concept of Emotional Disorder. Gloria Sibson Ayob, Oxford University Press. © Oxford University Press 2024.
DOI: 10.1093/9780198909620.003.0003

pathology on the sleeves, i.e. transparently. I will set out this circumventing strategy more fully in Section 2.2 below. Having noted the importance of the claim of transparency to this defence of pathologization, in the remainder of the chapter I will then explore two ways of upholding this claim of transparency. To anticipate: we will see that both approaches fail to do this. This in turn threatens to undermine the circumventing strategy for defending the pathologization of distressing emotions.

2.2 Circumventing the BED Framework: The Role of the Claim of Transparency

When an emotional experience is appraised negatively along all four of the intrapersonal basic evaluative dimensions—if it is deemed to be irrational, unpleasant, immobilizing, and viewed as bearing no existential significance—then the judgement of emotional pathology is uncontentious. The meta-evaluative judgement of disorder only becomes problematic when an emotional experience is appraised negatively in fewer than all four intrapersonal basic evaluative dimensions. That is, if there is one or more dimension in which the experience is positively appraised, this positive appraisal forms the basis for resisting the pathologization of the emotional experience. Authors resisting pathologization want to say that even if it has a deeply unpleasant valence and/or it is immobilizing, the experience is not pathological since they deem the emotional experience to be a rational response to its subject's circumstances and/or they take it to bear existential significance for the subject. By contrast, those like Kendler who insist on pathologizing these emotional experiences have to contend with the fact that they appear to be meaningful, both in the sense that an emotional experience of this sort may have an intentional content, and also in the sense that it appears to be a rational response to its subject's circumstances (and/or in the sense that the experience bears existential significance). One way of responding to this datum is to try to quash the appearance of meaningfulness, and a way of doing this is by according primacy to the judgement that the emotional experience is pathological, and on this basis asserting that the appearance of meaningfulness is misleading. The appearance of meaningfulness is thus overridden by the judgement that the experience is pathological. Both steps in this response need to be unpacked.

2.2.1 Step 1: According Primacy to the Judgement of Emotional Disorder

The BED analysis of the concept of emotional disorder helps us to see that the dispute between Horwitz and Wakefield and Kendler turns out to be a dispute about which basic evaluative dimensions to prioritize in judging our emotional experiences to be disordered. Kendler places the emphasis on the dimensions of valence and agency (at the intentional level), whilst Horwitz and Wakefield emphasize the dimensions of rationality and agency (at the volitional level). If Kendler accepts the BED analytical framework and seeks to argue against Horwitz and Wakefield's position within this framework, he will need to give us reasons for favouring the negative appraisals issued along the dimensions of valence and agency. This response accepts that the emotional experience is a meaningful one. (I will explore this response in the next chapter.) But he might do something else instead: he could try to step outside the BED framework altogether by rejecting its central claim that judgements of emotional disorder are meta-evaluative. If it can be shown that pathological emotional experiences display their disorderedness transparently, this would mean that we can judge an emotional experience to be pathological directly, without having to base this judgement upon any more basic evaluation of the experience. This judgement of disorderedness can then be treated as primary, i.e. it can form the basis upon which to override the emotional experience's appearance of being meaningful (step 2 below). The crucial point here is that the pathology status of the experience is fixed prior to, and independently of, all consideration of its appearance of meaningfulness.

How might a pathological emotional experience display its disorderedness transparently? The two most frequently cited candidates here are the neurophysiological and the phenomenological features of our emotional experiences. The idea, simply put, is that we can read off the pathology status of an emotional experience (both at the level of types and at the level of tokens) from its neurophysiological profile, or its phenomenological profile, or both. This would allow those defending the pathologization of distressing emotional experiences to circumvent the BED framework, and therein to override any appearance of meaningfulness stemming from positive appraisals of such an experience in the dimensions of rationality and/or existential significance. So the claim of transparency carries substantial import: this defence of pathologization rests entirely on it. Whether this claim can be upheld will be the subject of investigation in what follows.

2.2.2 Step 2: Denying the Appearance of Meaningfulness

Once an emotional experience is judged to be disordered directly on the basis of its neurophysiological or phenomenological profile, it is possible to combine this judgement with a conceptual package that is commonly associated with the general concept of pathology, in order to override the experience's appearance of being meaningful. The plausibility of this conceptual package may vary depending on the context in which it is employed, but for the sake of clarity, it is worth setting it out in its strictest formulation. For brevity, I shall refer to this configuration of ideas associated with the concept of pathology as 'the strict construal of pathology'. For immediate purposes, we will need only the first idea, so I shall discuss this now, introducing the remaining ideas in the next chapter (in Section 3.2).

The first pivotal idea in the strict construal of pathology then is that when a condition of a person—whether physical or mental—is said to be disordered, the fact of the subject's having this condition is determined wholly by the course of a pathological process. This condition may contain mental states that appear to be meaningful, e.g. doxastic or experiential states. For instance, the condition of epilepsy can generate belief-like states (i.e. delusions) that have, on the face of it, intentional content. But this appearance of meaningfulness is misleading because these states are in fact symptomatic of the pathogenesis of epilepsy. Or again, obsessive thoughts in obsessive-compulsive disorder (OCD) appear to have content, but if we take these thoughts to be symptomatic of the pathogenesis of OCD, then the fact of the subject's having them looks to be meaningless. This latter reflective judgement is underpinned by the observation that these mental states do not appear to be 'hooked up' in the appropriate ways to what they purport to be about, being generated instead by an independent process of disorder. To capture this idea, it will be helpful to draw a distinction between (1) having a mental state that appears to bear content, and, at the reflective level, (2) the fact of a subject's instantiating that mental state as being meaningful or meaningless. When our cognitive, perceptual, and affective capacities are exercised normally, our perceptual beliefs provide us with reliable information about our immediate surroundings. Such beliefs thus not only appear to bear content—they do in fact bear content—we can also be said meaningfully to have these beliefs. Conversely, the fact that a subject has a mental state which appears to bear content (a belief

or an emotional experience, say) will be meaningless if the mental state in question is in fact symptomatic of a pathological process.[1]

Applying this idea to emotional pathology, the proponent of pathologization could argue as follows: a distressing emotional experience may *appear to be* meaningful both in the sense that it has an intentional content, and also in the sense that it appears to be a rational response to its subject's circumstances and/or in the sense that the experience bears existential significance. However, appearances are misleading when the experience is disordered. In spite of the experience appearing to be meaningful, the fact of the subject's having the experience is determined entirely by the course of pathology, thereby lying outwith the register of meaningfulness altogether. In short, the fact of the subject's having the emotional experience is meaningless. Pathological anxiety, for instance, in no way indicates danger or the presence of a threatening stimulus condition because it is being generated by a pathogenic process, rather than tracking actual danger in a reliable way as normal anxiety does. Or to take an example that will be discussed at some length later in the chapter: the author Leo Tolstoy suffered from existential despair, which he grappled with and wrote about in his *Confessions*. If this existential despair is viewed as symptomatic of the pathogenesis of (pathological) depression, then what appears to be a (meaningful) existential crisis in which Tolstoy is wrestling with the ontological truths of human life—*his* life—will be judged in fact to be an experiential state that is symptomatic of a blind pathogenic process. Tolstoy would thus be fundamentally mistaken about the nature of his state of existential despair: the fact of his suffering from this state lies outside the register of meaningfulness altogether. To restate the point in terms of the BED framework, what appears to be an emotional experience that merits a positive appraisal in the dimension of existential significance turns out to be altogether undermined by the prior judgement that the experience is in fact disordered.[2]

The primacy of the judgement of pathology is all-important here: the pathology status of the emotional experience must be fixed prior to, and independently of, all consideration of its appearance of meaningfulness, in order

[1] There is another sense of the notion of meaning according to which the fact of having this mental state is meaningful, namely that of factive meaning. The fact that the subject has this mental state means, on this understanding of the notion of meaning, that the subject is psychotic (or suffers from OCD). But this notion of factive meaning is rooted purely in a brute association, in a manner similar to the fact that when smoke is visible it (usually) means there is a fire nearby. See Travis 2005 for a discussion of the notion of factive meaning.

[2] This circumvention strategy is less plausible when dealing with a distressing emotional experience that appears to be rational response to the subject's circumstances. Kendler, recall, recognizes that Ms. S has good reason to feel depressed, given the imminent threat to her life. Given this recognition, Kendler's defence of pathologization is better understood to involve negotiating the BED framework, rather than circumventing it. The negotiation strategy will be the subject of the next chapter.

for it to be able to circumvent and undermine the positive appraisal of the experience in the dimension of existential significance. This in turn presupposes that the disorderedness of the emotional experience is transparent: this would enable us to judge the experience to be pathological primitively, i.e. its pathology status being read off directly from observing the experience. What grounds are there for holding that judgements of emotional disorder are transparent? It is the neurophysiological and phenomenological features of emotional experience that are most frequently appealed to, in order to assert that an experience wears disorderedness on its sleeve (when it is indeed judged to be disordered). In the discussion that follows, I will focus mainly on the appeal to phenomenology, but I want to begin by considering the appeal to neurophysiology.

2.3 Circumventing the BED Framework: The Appeal to Neurophysiology

The move to circumvent the BED framework by appealing to the neurophysiological profile of emotional disorders consists in assuming, firstly, that pathologies of emotion can be distinguished and identified directly by their unique neurophysiological profile and on this basis, secondly, asserting that once the pathology status of an emotional experience has been fixed in this way, the strict construal of pathology entails that we disregard the *appearance* of meaning inherent in the emotional experience (manifest through the appraisal of the experience as rational, say, or as bearing existential significance). Its pathological nature undermines any putative meaning that the emotional experience may be thought to have.

This strategy for defending judgements of emotional disorder isn't conclusively available to us today since no unique neurophysiological profile has yet been definitively established for any category of emotional pathology. But what of a future state-of-affairs? Could this strategy become available? The hope that it will be is what motivates the Research Domain Criteria (RDoC) programme in psychiatry, and the huge financial and resource investment being made in neuroscientific research into mental disorders more generally.[3] But I want to suggest here that even if unique neurophysiological profiles

[3] The Research Domain Criteria (RDoC) framework was launched in 2009 by the US National Institute of Mental Health with the aim of resolving the perceived unwieldiness of the nosological framework set out in the DSM. The RDoC seeks to achieve this by deploying a biological-based—*rather than a symptom-based*—architectonic for understanding mental disorders. For an introduction to this framework, please see Cuthbert and Insel 2013.

are eventually discovered for the various categories of emotional pathology, this will only be possible on the basis of a *prior* resolution of the boundary problem, and that an adequate resolution can emerge only through a negotiation of the BED framework. If this is right, then it would be a mistake to think that it is possible to appeal to neurophysiology as a way of *circumventing* the framework.

Neurophysiological research findings are not self-interpreting with respect to the boundary problem, i.e. the problem of how to distinguish between disorder and ordinary functioning. The example of 'the homosexual brain' illustrates this point simply. Even if it turned out that the brain structures of homosexual individuals differ in systematic ways to the brain structures of heterosexual individuals, this would not be sufficient for determining that either the homosexual brain or the heterosexual brain (or neither) is disordered. We would need to be guided by ideas about 'healthy' sexual functioning in order to interpret this hypothesized neurophysiological difference as having a boundary-delineating significance, i.e. as amounting to a difference between disorder and ordinary functioning. Contrast this with the neurophysiology of dementia: in the Alzheimer's brain, abnormal levels of the naturally occurring protein—beta-amyloid 42—clump together to form plaques that collect between neurons and disrupt cell function.[4] The distinction between disorder and ordinary functioning is, in this instance, formulated in terms of neurophysiological theory, and this is what allows us to interpret this neurophysiological pattern (abnormal levels of beta-amyloid 42) as disordered, and on this basis to judge the neurophysiological structures that are distinctive of the Alzheimer's brain to be pathological (cf. Graham 2010).

While lowered serotonin or dopamine in the brain may be associated with depressed mood, to interpret a certain level of serotonin as disordered (or otherwise) in terms of biochemical theory would yield a different judgement of pathology from the judgement of emotional disorder that is rooted in the first instance in the observation of depressed mood. The former would be a judgement of biochemical health (whatever this might consist in); the latter would be a judgement of emotional health. Of course, 'biochemical health' may not be altogether unrelated to emotional health: dysfunction in the former domain may have an impact on emotional functioning. The important point to notice, though, is that we can understand this impact as bearing an emotional health-related significance only against a background set of ideas about what constitutes emotional health and emotional disorder (in other words: a

[4] Source: https://www.nia.nih.gov/health/what-happens-brain-alzheimers-disease.

resolution to the boundary problem with regard to the emotions). Lowered serotonin (a biochemical property) becomes relevant to the judgement of emotional disorder only through its association with depressed mood, and only on the condition that depressed mood is placed on the pathology side of the boundary between emotional health and pathology.

This means that the primary challenge is to determine the status of depressed mood—i.e. as pathological or otherwise. Until this prior challenge is addressed, the fact that depressed mood might have a unique neurophysiological profile (one that includes deficient activity of monoamine transmitters, for instance) cannot help us any further in fixing the status of depressed mood as pathological. So what goes into the decision about whether depressed mood should be understood to be pathological? Depressed mood is deeply unpleasant and often cripples our agency. Kendler contends that this unpleasantness and immobilization is sufficient for judging depressed mood to be disordered. But Horwitz and Wakefield insist that if depressed mood is experienced in the face of significant personal loss, then this is in fact a natural (thus: non-pathological) emotional response to this loss. Here, we return to the basic evaluative dimension (BED) framework. Deciding on the pathology status of experiences of depressed mood involves negotiating the five basic evaluative dimensions. So we need to have resolved the boundary problem—which in the case of the emotions involves negotiating the BED framework—in order to interpret the (putative) unique physiological profile of depressed mood as marking out a condition of disordered emotion. It is for this reason that the appeal to neurophysiology, as a way of circumventing the BED framework in the defence of judgements of emotional disorder, fails.

2.4 Circumventing the BED Framework: The Appeal to Phenomenology

I now want to turn to consider the appeal to the phenomenology of disordered emotion as a way of circumventing the BED framework. To anticipate: this appeal seems to me to be equally unpromising, and the principal difficulty here is that depression (construed as a pathological state) does not seem to have a unique and distinctive phenomenological profile that would specifically serve to distinguish it from emotional distress appraised as a rational response to challenging life circumstances and/or as being existentially significant.[5] I will

[5] The appeal to phenomenology will be discussed in the review of Ratcliffe's position below.

establish this point by drawing upon the phenomenological characterization of depression that has been presented by Matthew Ratcliffe (Ratcliffe 2015). I have chosen Ratcliffe's view because of its thoroughness. This is one of the best worked-out phenomenological accounts of depression available today, so it presents as strong an opportunity as any (I believe) for anyone hoping to appeal to a detailed phenomenological profile of depression to establish its pathological status directly.

For such an interlocutor, Ratcliffe's account starts off promisingly because Ratcliffe does think that depression is distinguishable from intense sadness on phenomenological grounds. It is unique in this sense. However, in spite of his insistence that depression is phenomenologically distinct, Ratcliffe does not assume at the outset that depression is pathological in nature. When he does eventually address the boundary problem, Ratcliffe concedes that depression, in spite of its distinct phenomenological profile, is not always or necessarily pathological. To indicate his open-ended stance towards the boundary problem, I will refer to Ratcliffe's concept of depression as 'depression$_{\text{Ratcliffe}}$'. The way in which Ratcliffe deems it to be possible for depression$_{\text{Ratcliffe}}$ not to be pathological seems to me to closely match the reasons why various authors have resisted the pathologization of emotional distress that is taken to bear existential significance. Restating this point in terms of the language of the five basic evaluative dimensions: an emotional experience that is negatively appraised along the dimensions of affective valence and volitional/agential efficacy (and on this basis described as an experience of depression), but which is appraised positively along the dimension of existential significance—that is, the experience is viewed as having existential significance—may not be pathological, according to Ratcliffe. Note that although Ratcliffe's view converges on the analysis of the concept of emotional disorder as a meta-evaluative concept that I am setting out here, it is distinct from it because Ratcliffe characterizes an experience as 'depressive' not directly on the basis of negative appraisals in any of the five basic evaluative dimensions, but rather on the basis of the observation that the experience instantiates a distinct phenomenological profile. Nonetheless, the crucial point is that although it has a distinct phenomenological profile, Ratcliffe recognizes that depression$_{\text{Ratcliffe}}$ does not—indeed, cannot—serve to delineate the boundary of emotional disorder. If this is right, it means we shan't be able to use the phenomenology of depression as a basis for denying the opacity of judgements of disorder, and this in effect means that we shall not be able to quash the appearance of meaningfulness of a depressive experience by appealing to the prior judgement that the experience is pathological.

Having explored one of the best worked-out phenomenological accounts of depression and found that it doesn't offer us a basis for demarcating the boundary of emotional disorder, I conclude that the appeal to the phenomenology of depression will fail to defeat the claim of opacity. The implications of this for the attempt to defend the pathologization of distressing emotional experiences, by circumventing the BED framework, will be discussed in Section 2.7 below.

2.5 Ratcliffe's Phenomenological Characterization of Depression

Ratcliffe holds that depression is distinct phenomenologically from experiences of sadness, even intense sadness. In the broadest terms, what makes depression distinct is that it involves a shift in 'existential feelings', whereas sadness does not. Existential feelings, Ratcliffe states, constitute 'a variable sense of the possibilities that the world incorporates' (Ratcliffe 2015: 2). To understand what this means—to understand how our sense of the world consists, at least in part, in possibilities it incorporates (indeed, to understand what Ratcliffe means by 'possibilities')—we need to take note of an idea that is of fundamental importance in Ratcliffe's framework. According to Ratcliffe, although many of our emotions are intentional states, there are a group of feelings that comprise *pre*-intentional ways of 'finding oneself in the world'. These feelings form a kind of background against which intentional states (including the emotions referred to a moment ago) are experienced. This background forms the *existential structure* of experience. Examples of existential feeling provided by Ratcliffe include 'unfamiliarity', 'strangeness', 'detachment', or conversely, 'at home in the world'. In experiencing the world through these feelings, we experience the world as affording—or foreclosing—possibilities. A shift in existential feeling is thus a change in the kinds of possibility that are experienced (ibid.: 42). What are these kinds of possibility? Ratcliffe presents us with a diverse list here: the possibilities for experiencing objects and spaces perceptually, these arising within specific perceptual modalities; the possibilities of events, opportunities, threats etc. that are anticipated; the possibilities that one can actualize through an exercise of one's agency; the possibilities of experiencing things as significant or as mattering; and the possibilities of relating to ourselves and to others (ibid.: 52).

Depression, claims Ratcliffe, essentially involves a change in the kinds of possibility that are experienced—and so should be understood to consist in

a shift in existential feelings (ibid.: 42). Elaborating on this schematic point, Ratcliffe states:

> Depression, I maintain, involves a change in the kinds of possibility that are experienced as integral to the world and, with it, a change in the structure of one's overall relationship with the world . . . . [In further detail, the changes involved include] altered bodily experience, loss of hope, feelings of guilt, a diminished sense of agency and self, altered experience of time, and isolation from other people. All of these are to be conceived of as inextricable aspects of a unitary experience—a shift in existential feeling or 'existential change'. (Ratcliffe 2015: 2)

Filling out this schema, Ratcliffe says that in depression, this change in 'the structure of the [depressed] subject's overall relationship with the world' may be captured using the image of eternal incarceration. The depressed subject feels globally imprisoned, without hope of reprieve (ibid.: 64). The prison or enclosure is perpetually oppressive and suffocating. It is also solitary and it is static. Ratcliffe:

> Despite superficial differences in how the 'world' of depression is described, consistent themes are easily discerned. The enclosure is always oppressive and suffocating. Styron (2001, p. 49) compares it to 'the diabolical discomfort of being imprisoned in a fiercely over-heated room', whereas Alvarez (2002, p. 293) found himself in a 'closed, concentrated world, airless and without exits'. It is also solitary and inescapable: one is irrevocably alone, cut off from the rest of humanity. Another theme is that of stasis: the world of depression is bereft of even the *possibility* of change. One watches the dynamic lives of other people from inside a solitary, unchanging bubble. Alvarez (2002, p. 103) thus describes a severe depression as 'a kind of spiritual winter, frozen, sterile, unmoving'. This emphasis on the lack of movement suggests that the sense of incarceration is temporal more so than spatial in character. It involves a feeling that things will not and cannot change. If nothing can change, then one cannot escape. (Ibid.: 65)

The 'loss of possibility' is taken by Ratcliffe as the primary element in terms of which the existential feeling of being imprisoned is to be understood. This loss of possibility shows up in a range of ways. For instance, in the domain of affect, this loss of possibility is experienced as the sense that one is no longer able to find happiness in anything ('It is not just that specific things cease to make the person happy in the way they once did. Instead, "happiness" is no longer part of her emotional repertoire . . . . She gradually loses the sense that

anything in the world could offer happiness: she ceases to experience its possibility' (ibid.: 65–6). In the domain of agency, this loss of possibility may be experienced in ever more general ways: it is possible that the depressed subject retains her hold on her concerns, but simply finds herself unable to act because of the loss of enticing possibilities in action. More fundamentally, though, the attractiveness of those concerns themselves may recede. In this latter loss of possibility, which Ratcliffe terms 'loss of practical significance', everything that the depressed subject encounters is 'stripped of the possibilities for action that it was previously imbued with. So she not only feels unable to act in such a way as to bring about significant outcomes; no sense remains of there being *any* significant outcomes' (ibid.: 166). Summarizing: the notion of imprisonment—which is fleshed out as the loss of possibilities for the subject—is Ratcliffe's basic schema for the phenomenology of depression. Just as he does with affect and agency, Ratcliffe applies this schema to other domains (e.g. interpersonal functioning and temporal experience), and in this way builds up a comprehensive characterization of the phenomenology of depression.

On this basis, Ratcliffe claims that depression$_{\text{Ratcliffe}}$ manifests a distinct phenomenological profile. My aim here is neither to endorse nor to reject Ratcliffe's phenomenological characterization of depression (I remain open-minded as to whether depression and intense sadness are altogether distinct from each other), so I will spend no time critically inspecting it. Rather, the purpose of reviewing Ratcliffe's account is to have it before us in order to understand how it might be used by someone wishing to establish the pathological status of depression directly on the grounds of its unique phenomenological profile. This is one of the best accounts available to which the envisaged interlocutor could appeal.

2.6 Is the Unique Phenomenological Profile of Depression$_{\text{Ratcliffe}}$ Sufficient for Delineating the Boundary of Emotional Disorder?

If Ratcliffe's claim is granted that depression$_{\text{Ratcliffe}}$ is phenomenologically distinct, can we say after all that pathology does wear itself on the sleeves of disordered emotion? Ratcliffe himself stops short of drawing this conclusion, asserting that depression$_{\text{Ratcliffe}}$ is not always or necessarily pathological. His argument for this position is worth setting out because it identifies an important reason why the pathologization of (at least some) depressive experiences

has been resisted on the grounds that such experiences can bear existential significance.

Addressing the boundary problem requires us to have a clear understanding of what the term 'pathology' means. So what is Ratcliffe's view on this? Ratcliffe starts out by noting Jennifer Radden's observation that there are two conceptions of depression:

(i) The disease conception. Depression consists of one or more as yet unidentified pathological processes, to be detected by identifying characteristic symptoms. Here, depression is conceived of as an underlying disease processes that causes the existential changes that are definitive of depressive$_{\text{Ratcliffe}}$ experiences.
(ii) The descriptive conception. Depression just *is* a 'syndrome' or cluster of symptoms. Here, depression is identified with the existential changes directly. (Ratcliffe 2015: 266)

Ratcliffe thinks that the first construal of the concept of pathology is not viable (at least not currently), not just because the relevant pathological processes have yet to be identified, but more fundamentally, because depressive experiences are presently very heterogeneous. According to Ratcliffe, it is improbable that pathological processes will be found to underpin every kind of depressive experience. A more probable view is that at least some depression experiences arise due to disease processes—but adopting this view will mean having to (a) distinguish between kinds of depression experience and then to (b) determine which depression experiences do arise in this way, and which do not. The empirical task of identifying underlying disease processes will have to be informed by such discriminating phenomenological analyses, on pain of the task's being marred by a lack of focus that results in 'conflicting findings and confusion' (ibid.: 267). Until these hurdles are overcome, Ratcliffe thinks that the disease conception of depression will remain something of a regulative ideal, rather than the conception of depression that underpins the actual judgements we make about emotional pathology in this domain.

Having set aside the disease conception, Ratcliffe then turns to the second of the two conceptions of depression, i.e. the descriptive conception. Here, Ratcliffe notes that we'll need to identify criteria for pathology that can be applied to the existential changes that mark depressive$_{\text{Ratcliffe}}$ experiences, in order to judge that these existential changes are pathological. He suggests that the only (obviously) available criteria are pragmatic: in judging a depressive$_{\text{Ratcliffe}}$ experience to be pathological, we are in fact evaluating the effects

that depression has on a person's life—more specifically, we are evaluating the *harmful* and *beneficial* effects of depression on our lives (ibid.: 267). Insofar as a depressive$_{\text{Ratcliffe}}$ experience has harmful effects, it is judged to be pathological; beneficial effects pull against this judgement, leading us to view the depressive$_{\text{Ratcliffe}}$ experience as part and parcel of ordinary (i.e. non-pathological) human experience. In sum, Ratcliffe resolves the boundary problem pragmatically, i.e. in terms of harmful and beneficial effects on a person's life.[6]

Now, if depressive$_{\text{Ratcliffe}}$ experiences were categorically deemed to be harmful, Ratcliffe thinks that this would make it possible for us to take the distinctive phenomenological profile of depressive$_{\text{Ratcliffe}}$ experiences as sufficient for fixing the pathology status of these experiences (i.e. as disordered). However, Ratcliffe observes that while depression is 'overall . . . a bad thing to have', it is not altogether clear that depression is exhaustively and categorically harmful, especially when we bring into view its *epistemic* benefits and harms (ibid.: 268). The key question for Ratcliffe here is whether depressive$_{\text{Ratcliffe}}$ experiences can be revelatory of important truths (thus beneficial), or whether such experiences only ever distort its subject's grasp of reality (thus harmful). If the latter, depression can be judged to be an 'epistemic pathology'. Otherwise, not.

To investigate this, Ratcliffe embarks on an inquiry into depressive experiences of a particular sort, which he terms 'existential despair'. He explores this kind of depressive experience through Leo Tolstoy's (1882) and William James' (1902) writings on the subject, and summarizes it as follows:

> Tolstoy's despair [in his *Confessions*] involves an especially profound loss of existential hope . . . . It is not just that he takes all human life to be without value. He cannot even contemplate the possibility of its being otherwise, and the experience has a *feeling* of irrevocable certainty to it. The capacity to take pleasure in anything, to be drawn in by situations, or to engage in meaningful activity is altogether absent . . . . Why does all human activity appear but a futile distraction? A heightened and/or altered awareness of mortality seems to be largely responsible, and Tolstoy couches existential despair in terms of a negative response to the question 'is there any meaning in my life that the inevitable death awaiting me does not destroy?' . . .

[6] What the five basic dimensions of evaluation suggest is that evaluations of benefit and harm are vague as they stand, since it is possible to judge an emotional experience to be harmful/beneficial in one (or more) of the basic evaluative dimensions without thereby also judging it to be pathological. Although Ratcliffe's response to the boundary problem is not explicitly informed by the BED framework, we'll see in the final paragraph of this section (Section 2.6.) that he recognizes that a depressive emotional experience, distressing though it may be, may bear existential significance.

> So the experience is something like this: a heightened sense of mortality comes to light when the capacity for effortless, pleasurable immersion in activity is blocked, and this renders worthwhile activity unintelligible. An agitated need to achieve something lingers on, with no possible outlet. (Ratcliffe 2015: 270–1)

William James, drawing in part on Tolstoy's *Confessions*, also explores the state of existential despair, describing it as 'an intense awareness of human mortality and the inevitability of suffering, which develops into an experience of the world as fundamentally evil, a place in which we can never be safe or feel at home. Our projects crumble, given that they rest upon a hope or faith in the goodness of life that reveals itself as utterly unfounded' (ibid.: 271).

Ratcliffe now asks: is this view of reality—and specifically, of the ontological situation we find ourselves in—revelatory or distorting? One way of interpreting this question is in terms of assessing the impact that depression has on the cognitive processes that underpin our capacity for evaluative judgement. However, Ratcliffe rejects this way of approaching the question on the grounds that there is likely to be more than one capacity exercised in the myriad evaluative judgements we make, so that it would be hasty to generalize from (a) findings of impaired capacity regarding evaluative judgements on matters the truth of which is easier to establish (e.g. a subject's assessment of her self-worth, or her feelings of excessive guilt) to (b) the capacities exercised in judging whether the quotidian activities we engage in are ultimately futile. Existential despair could very well involve an accurate evaluation of our ontological predicament, even if it interferes with thought and activity more generally (ibid.: 274).

So Ratcliffe suggests that we should instead interpret the question about whether existential despair distorts or reveals by reference to his general analysis of the existential changes in depression. According to this interpretation of the question, existential despair is distorting—it is epistemically pathological—insofar as it involves 'losing kinds of possibility that play an essential epistemic role'. Ratcliffe elaborates:

> The way the despairing person evaluates her predicament (and that of others too) is symptomatic of her inability to contemplate alternatives. In order to competently evaluative a state of affairs as *p* rather than *q*, one must be able to first comprehend the possibility of *q* and then rule it out. If an ability to even entertain the possibility of *q* were lost, then one's commitment to *p* would reflect incapacity rather than *p*'s relative plausibility. (274)

Applying this more concretely to the despairing person, Ratcliffe states:

> Feelings of certainty that arise in depression are often deceptive. For instance, the belief that recovery is impossible stems from an ability to contemplate something that is not only possible but probable . . . . If access to alternatives were restored, it would again reveal itself as a contingent evaluation, and not a very enticing one either. (274)

From this we might try to conclude that the depressive experience of existential despair is distorting in this sense: 'Those who are not stuck in it have access to kinds of possibility that reveal its certainties as misguided.' However, while this may be a satisfactory way of dealing with certain aspects of a broader depressive$_{\text{Ratcliffe}}$ experience within which the specific state of existential despair may be embedded (e.g. Ratcliffe's example of the prospects of 'recovery'), the pathology status of the despair specifically doesn't lend itself to being easily determined on this interpretation of the question either. What relevant possibilities is Tolstoy unable to entertain, what possibilities are sealed off him in this way, when he views his daily efforts as futile and empty as a certainty? The obvious relevant possibility is that the world is a fundamentally good place that accommodates worthwhile human activity. But the difficulty here, Ratcliffe notes, is that it is implausible to say that being certain about the futility of our daily efforts is indeed misguided, or at any rate, more misguided than being open-minded about whether our efforts are futile or worthwhile, because we cannot conclusively demonstrate that the despairing view is mistaken. (Compare: it is far easier to establish the falsity of claims such as 'things cannot get better', which is why it is more plausible to judge the certainty invested in such claims—by depressed subjects—to be misguided.)

Even when the certainty in other regards recedes, a person who has a depressive$_{\text{Ratcliffe}}$ experience may continue to view her ontological predicament through the lens of despair. 'How,' asks Ratcliffe, 'do we respond to someone who continues to place the flag of truth in the world of existential despair?' (277). Here, we seem to face an impasse. Accepting that existential despair and its converse—existential hopefulness—are partly constituted by existential feeling, Ratcliffe states that each respective stance may only be fully graspable if we have the required feeling. This is because these feelings may ' "offer our only mode of access to certain values" ' (Ratcliffe 2015, quoting Wynn 2005). So those who enjoy existential feelings of the sort that anchors hopefulness simply do not have epistemic access to the world inhabited by the despairing. This leaves open the question: have those who suffer from existential despair stumbled upon a truth that the existentially hopeful have the good

fortune of being unable to access? As each stance is encapsulated from the other—this, given the ineliminable epistemic role of existential feeling (over which we have little choice)—there seems to be no way of settling the matter conclusively in favour of one stance. So in the end, we cannot rule out the possibility that existential despair is revelatory. If this is right, then in light of Ratcliffe's pragmatic resolution of the boundary problem, it is not possible to say that existential despair is always and necessarily pathological. So even though depressive$_{\text{Ratcliffe}}$ experiences have a distinct phenomenological profile, this alone is not sufficient for grounding judgements of emotional disorder.

2.7 The Collapse of the Claim of Transparency: Implications

The purpose of sketching out Ratcliffe's phenomenological characterization of depression has been to survey an instance of a well-worked-out account that would be available now to someone hoping to use the unique phenomenological profile of depression to defend the claim that judgements of emotional disorder are transparent. The upshot of this discussion is that even if depressive experiences have a unique phenomenological profile, as Ratcliffe suggests, this profile does not yet offer us a basis for demarcating the boundary of emotional disorder. So the appeal to phenomenology fails to uphold the claim of transparency. There may be no discernible phenomenological difference between someone who is pathologically depressed and someone who experiences (non-pathological) existential despair. This means that it is not possible to take the claim of pathology as primary, appealing directly to the unique phenomenological profile of depressive experiences in order to do this, and then using this claim of pathology as the basis upon which to reject the apparent meaningfulness of such experiences, i.e. experiences that have been appraised positively in the basic evaluative dimensions of rationality or of existential significance. In light of this, all that the pathology overlay does is to misconstrue an emotional experience—that would otherwise be meaningfully engaged with—as meaningless, viewing it (without sufficient reason) as the product of a blind pathogenic process.[7]

[7] It is worth bearing in mind that this risk of mis-construal doesn't arise if it is universally accepted that the emotional experience in question is negatively appraised in all four intrapersonal basic evaluative dimensions—for in that case, everyone agrees that there is nothing of significance to be discovered in the experience. The risk is germane only when the claim of pathology rests on negative appraisals in fewer than all four evaluative dimensions.

The fact that there is no clear basis upon which to directly rule out the apparent meaningfulness of distressing emotions carries an important implication: it places the burden of proof on those who seek to pathologize such experiences. How so? The insufficiency of phenomenology (or physiology) to establish the pathology status of contested emotional experiences entails that judgements of emotional disorder remain opaque. We arrive at these judgements only through more basic evaluations of our emotional experiences that do not themselves invoke the concept of pathology. In view of this opacity, it doesn't seem possible to circumvent the BED framework. Instead, those seeking to defend the pathologization of distressing but apparently meaningful emotional experiences will have to *negotiate* the basic evaluative dimensions (as set out in the BED framework). This means that the decision to pathologize a distressing or immobilizing emotional experience that seems to bear existential significance (e.g. existential despair) will need to be justified. In particular, it is essential that we are told why the dimensions of valence and agential efficacy along which the experience is negatively appraised are given priority, specifically since doing so would negate the apparent meaningfulness of the experience. By contrast, the decision not to pathologize the experience preserves this appearance of meaningfulness: we can take the appearance of bearing existential significance at face value. In this way, the onus of proof lies with the decision to pathologize, rather than the decision not to pathologize, the emotional experience in question. In the next chapter, I will explore the significance of this choice.

3
The Medicalization of Emotions (II)

Negotiating the BED Framework

> A self is the thing the world is least apt to inquire about.
>
> **Kierkegaard (1941: 301)**

3.1 Introduction

The central claim in this book is that judgements of emotional disorder are meta-evaluative and thus opaque. In the first chapter, a framework was set out that elucidated this claim. According to this framework (the 'BED framework' for brevity), judgements of emotional disorder are made on the basis of negotiating five more basic dimensions along which we appraise our emotional experiences. Negative appraisals tend to yield judgements of emotional pathology, whilst positive appraisals tend towards the opposite. In the second chapter, we looked at the attempt to circumvent this framework altogether by rejecting the claim of opacity. If the disorderedness of an emotional experience can be taken as primitive, on the basis of its transparency, then even if the experience appears to be a meaningful one (on account of appearing to be a rational response or of appearing to bear existential significance), the primary judgement of disorderedness overrides this appearance since the claim of pathology entails that the fact of having the emotional experience is inherently devoid of meaning (appearances to the contrary notwithstanding). In this way, all basic evaluative dimensions can be circumvented. To probe this rejection of the claim of opacity, two most common ways of fleshing out the claim of transparency were considered, i.e. by appealing to the physiology and to the phenomenology of emotional experiences respectively. However, both these aspects of emotional experiences were shown to fail to establish the transparency of emotional disorder. The upshot of this discussion is that it does not seem possible to circumvent the BED framework.

The Concept of Emotional Disorder. Gloria Sibson Ayob, Oxford University Press. © Oxford University Press 2024.
DOI: 10.1093/9780198909620.003.0004

This means that any attempt to defend the pathologization of distressing emotions that are positively appraised as rational and/or as bearing existential significance must do so by negotiating the basic evaluative dimensions. That is to say, although such a defence of pathologization will be anchored in negative appraisals in the dimensions of valence and agency, it has to address the positive appraisals along the dimensions of rationality and existential significance. One way of doing this is to water down the strict construal of pathology by asserting that an emotional experience can be both disordered *and* meaningful. This response tries to take the meaningfulness of the experience as seriously as its disorderedness. However, the difficulty with watering down the strict construal of pathology is that it threatens to render vacuous our understanding of what it means to suffer from a pathological condition. As the claims we make about pathology become more insubstantial, their meaning and significance threatens to become ever more obscure, and we are in danger of becoming incapable of drawing important distinctions. Some of these distinctions will be examined in greater detail in this chapter and the next two.[1]

A different approach to defending the pathologization of distressing but rational and/or existentially significant emotional experiences is to negotiate the BED framework within the parameters set by the strict construal of pathology. Here, the guiding idea would be that, whilst an emotional experience may appear to bear existential significance and/or be a rational response to the subject's circumstances, these positive appraisals are *outweighed by* the negative appraisals of the experience along the dimensions of valence and/or agency. This defence therefore broadly consists in taking the dimensions of (negative) valence and (immobilizing) agency as sufficient for justifying the pathologization of emotional experiences.

This latter response seems to be the strategy that Kendler adopts in his dispute with Wakefield and Horwitz. Recall from the Introduction that in arguing against and Horwitz and Wakefield's insistence that we should not be pathologizing intense sadness that is a response to significant personal loss, Kendler accuses Horwitz and Wakefield of 'over-romanticizing' distressing emotional experiences (Kendler 2008; cf. Introduction, Section I.2). By this, Kendler must mean either that there is nothing of significant value in the suffering wrought by these distressing emotional experiences, or he might be making the more qualified claim that, whilst the experiences themselves may

[1] In Section 5.3, in the final chapter, I will discuss a few considerations that have encouraged the watering down of this strict construal in the domain of our emotional lives. By that point in the book, we will have a better understanding what is at stake in this watering-down—what price we pay for it—and therein to appreciate the value of holding on to the strict construal.

be inherently valuable, this matters less to the judgement of emotional disorder than the fact that they are intensely distressing, perhaps to the point of being immobilizing. Either way, according to Kendler, Horwitz and Wakefield are mistaken in thinking otherwise and are thus also mistaken in arguing on this basis for restricting the concept of emotional pathology. This accusation is the central plank upon which rests Kendler's defence of the pathologization of distressing emotional experiences.

The aim of this chapter is to examine this latter strategy for defending the pathologization of distressing emotions. I will begin in Section 3.2 by laying out the strict construal of pathology mentioned in the previous chapter. This provides a vital background both for understanding what might be meant by the charge that we over-romanticize emotional suffering by (wrongly) failing to judge it as disordered, and for understanding why one might seek to repudiate this charge. Setting out this strict construal will thus enable us to obtain a clearer view of the values at stake in negotiating the basic evaluative dimensions as Kendler does.

I will then set Kendler's accusation of over-romanticization in relief, doing so by exploring an instance of the contrary view that emotional suffering can be inherently valuable. Specifically, according to the view to be explored, certain distressing emotions can bear existential significance, and this provides us with a sound reason for resisting the meta-evaluative judgement that such emotional experiences are pathological.[2] The point of setting out this view is to have before us a clearer sense of what has to be defeated—i.e. the claims about distressing but existentially significant emotions that have to be shown to be groundless or mistaken—by those wishing to uphold the pathologization of emotional experiences on the basis of the sufficiency of negative appraisals in the dimensions of valence and agency alone. The example I have chosen of the view that emotional suffering can bear existential significance is presented by James Hollis, a psychoanalyst who writes about depressive experiences. I will sketch out Hollis' account of these experiences in Section 3.3 below.

Having set out Hollis' view, I will then argue in Section 3.4 that this view provides us with strong grounds against pathologizing depressive experiences of the kind Hollis discusses. To establish this, I'll compare the two ways of

[2] I will focus on positive appraisals made in the basic evaluative dimension of existential significance, rather than in the evaluative dimension of emotional rationality. I have chosen to emphasize the dimension of existential significance in this chapter since disputes concerning the dimension of rationality are more familiar. Indeed, this latter dimension (in particular the strand of representational accuracy) is a focal point in Horwitz and Wakefield's pushback against the wholesale pathologization of distressing emotional experiences. The relatively unfamiliar evaluative dimension of existential significance thus merits further exploration.

viewing these experiences, i.e. through the lens of pathology and without it. The relationship in which we stand to these experiences emerges as a question of central importance here: it is a defining axis along which the dispute between those resisting and those advocating the pathologization of these depressive experiences may be understood. For this reason, the decision to pathologize the depressive experiences that Hollis discusses turns out to be a morally and personally significant one, as I'll suggest in Section 3.5. Altogether, the discussion in this chapter maps out the space that has to be negotiated by those who seek to defend the pathologization of distressing but existentially significant emotional experiences on the grounds that their distressing and immobilizing quality is sufficient for anchoring the judgement of pathology.

3.2 The Strict Construal of the Pathology Concept

Our concepts of pathology are multifaceted and flexible, used in a wide range of ways. Whether it will ever be possible to achieve consensus on an informative analytical or even prototypical definition of these concepts remains to be seen, and there is much debate about whether (and if so, how) they might be extended to include mental disorders (see Pickering 2006). There is, however, a conceptual package that is commonly associated with the general concept of pathology, which I have referred to in the previous chapter as 'the strict construal of pathology'. The term 'association' is used deliberately: elements in this conceptual package are deployed variously by different authors, and as the contexts in which we use the notions of health and illness shift, so too do the ideas we associate with these notions. This is a fast-moving target.

However, to enable us to think carefully through the meaning of the concept of emotional pathology, and to understand the actual and potential implications of the judgement of emotional disorder, it will be useful to fix this target. The strict construal of the pathology concept preserves the integrity and substance of this concept, so using it as a basis for examining the concept of emotional disorder will help give us a clearer understanding of the stakes that are involved as we negotiate the five basic evaluative dimensions in delineating the boundary of emotional disorder. The strict construal seems to capture the concept of pathology in practical rather than theoretical terms, but it is precisely this fact about it that enables the strict concept to play a *comprehensive* anchoring role in the present inquiry into the concept of emotional disorder in view of the latter's opacity, for the following reason. Drawing on the discussion in Section 1.9 illustratively, we have seen that the concept of biological

dysfunction within which Horwitz and Wakefield's evolution-theoretic analysis of disorder is rooted looks to track appraisals of our emotional experiences in just the dimensions of representational accuracy and low-level volitional efficacy. In so doing, the analysis seems to presume that states of emotional disorder are constituted transparently by the features of an emotional experience as manifest within these dimensions exclusively. But this presumption overlooks the fact that there are further basic evaluative dimensions that are germane to the concept of emotional disorder and, importantly, that judgements of emotional disorder are meta-evaluative judgements that involve negotiating the appraisals made of an emotional experience along these various basic dimensions. Horwitz and Wakefield's evolution-theoretic analysis thus seems to offer us at best a partial understanding of the concept of emotional disorder, as noted in Section 1.9.2. If a purely theoretical concept of pathology is ever to be articulated that can adequately shed light on the idea of emotional disorder, it will need to encompass this feature of opacity (i.e. that judgements of emotional disorder are meta-evaluative).

Because the claim of opacity is a pivot upon which the present analysis of the concept of emotional disorder turns, and because it remains to be seen whether—and if so, how—a theoretical analysis of the concept of pathology might be articulated that accommodates the opacity of the concept of emotional disorder, we will need a way of fixing our grasp on the general notion of pathology other than by means of a substantive theoretical analysis of the concept. The strict construal of the pathology concept does just this. It gives us a handle on what we are speaking about when we talk about emotions being disordered; crucially, it gives us a way of contrasting different kinds of judgement we make about our emotions (one such contrast will be explored later in this chapter; then in Section 5.3, I will touch upon a more general version of this contrast, namely the judgements we make of our emotions rooted in moral and in medical discourse respectively). In other words, the strict construal is a regulative assumption that gives our medical discourse its distinctive shape. It is for this reason that I will now lay out the key elements in the strict construal of the concept of pathology more fully and use it to inform the inquiry that follows.[3]

[3] It may be objected that the strict construal of the pathology concept as set out here works well for the concept of physical disorder only, and not for mental disorder, because it is far from obvious that experiences judged to be psychopathological should be taken to be devoid of meaning. To this I wish to make the following observation: it may very well turn out to be the case, at least in the domain of human emotion, that we are tempted to say this because we've drawn the boundary between disorderedness and non-disorderedness in the wrong place to begin with. That is, it may turn out to be the case that we should resist judging an emotional experience to be pathological *precisely* in view of the fact that it is deemed meaningful, in spite of the fact that we had been initially led to the view that the experience is pathological on account of its being negatively appraised in certain dimensions (the dimensions of affective valence and of

In the previous chapter I began by presenting the following principal idea contained within this construal: any condition experienced by a person—this can include mental states—that is symptomatic of a pathogenetic process stands outside the register of meaningfulness altogether. Even if the mental state appears to be content-bearing, the fact of the subject's having that mental state is in fact meaningless, being generated and determined as it is by processes of pathology. Closely related to this idea is another prevalent idea contained within the strict construal of pathology: this condition is external to the subject's core or real self. The notion of the real self has been explored extensively in the philosophical debate about personhood and personal autonomy and, in a nutshell, it aims at capturing the intuitive distinction between experiences, motives, intentions, etc. that are mine in a robust sense—from which stems the idea of acting in a voluntary and self-directing or self-governing way—and experiences, motives, etc. that are not mine in this robust way (even though they originate from within the physical boundary of my organism). Unbidden thoughts or involuntary movements that I do not reflectively endorse are examples of the latter sort. The notion of reflective endorsement plays a central role in helping to fix the sense of ownership in the more robust sense: the more an experience is aligned with the subject's most deeply held values and concerns, the easier it is to avow it, to see it as integral to oneself in this sense. The term 'real self' is used to denote our sense of personal identity, our sense of which experiences are ours, in this robust sense.[4] A fuller characterization of the idea of our robustly owning certain emotional experiences—and therein of these experiences being partially constitutive of our real selves—will be offered in Section 4.6.3, but it should suffice for now to note the following point. An experience the having of which is viewed as meaningless on account of its being seen as the product of an independent, pathogenetic process will by definition fall outwith the subject's core or real self, for to reflectively endorse an experience presupposes the subject's considering it to be a meaningful experience to have. The subject sees it as emerging out of her personal history, in conjunction with her ever-evolving engagement with her present environment; it is an experience that can be integrated in

volitional agency, say). Appreciating that the judgement of emotional disorder is meta-evaluative helps us to see that the latter negative appraisals do not settle the matter conclusively, and that the meaningfulness of an emotional experience may in fact be taken to constitute valid grounds for rejecting the claim that this experience is disordered. I hope that the value in approaching things this way will become clearer in this chapter and in the fifth.

[4] This way of thinking about what it means to be an autonomous, self-directing agent enjoys broad influence today (though it is not without difficulties). For the articulation of this way of conceptualizing the notion of self-governance, and personhood in this robust sense, see Frankfurt (1971).

meaningful ways with her deepest concerns and values. Indeed, a key driver of the debate about the pathology status of experiences such as psychosis, in individual instances, is the question of whether the fact of having a psychotic experience is meaningful, and this matters precisely because it determines how the experience is taken to be situated relative to the subject's sense of who she feels herself to be (Jackson and Fulford 1997 cf. Sibson Ayob 2019). If the fact of the subject's having the experience is judged to be meaningful, this allows us to view it as expressive of who she really is. If on the other hand the experience is taken to be heteronomous, generated and determined by an external process of pathology, this precludes viewing it as being meaningfully had, and is placed outside the boundary of the subject's real self. This is the second idea contained within the strict construal of the concept of pathology.[5]

[5] In the debate about the pathology status of autism, the relationship in which the (contested) experience stands to self is invoked to place pressure on the pathologisation of autism. Against this, and in defence of pathologisation, Wakefield and his co-authors contend: 'A sense of identity as an autist does not imply the nondisorder status of autism, any more than the ego-syntonicity of a personality disorder implies no disorder.' (Wakefield et al 2020: 512). As a general response, this seems to me to move too swiftly because of the close connection between the first two ideas just outlined: it is in virtue of lying outwith the register of meaningfulness—on account of being caused by a pathogenic process—that the experience is situated externally to the core self. So to say that a disordered experience can be taken by its subject as internal to the self, as Wakefield does, is *also* to say that an experience that lies outwith the register of meaningfulness (i.e. an experience that is devoid of meaning) can be taken as internal in this sense. Setting aside the discussion of autism and focusing just upon the domain of emotion (as it pertains to the concept of emotional pathology), this latter assertion entails that an altogether novel account has to be given of how a subject comes to regard an emotional experience as internal to her core self, one that doesn't relate at all to the content-bearing status of the experience. This novel account threatens to trivialise if not altogether distort our attempt to understand ourselves, to form and to sustain a reflective view of who we are, in light of our experiences. As we engage in self-reflection, what exactly would we be tracking in our emotional experiences if not their purported content? The meaningfulness of an experience looks to be at least part of the basis upon which we come to regard the experience as internal to our core selves; the two ideas thus seem to be tightly connected with each other. If we aren't to jettison this connection, the only other option for asserting that a disordered experience can nonetheless be a constituent of a subject's core self would be to reject the idea that experiences generated by a pathogenic process necessarily lie outwith the register of meaningfulness. It seems to me that the success of this manoeuvre will turn upon the sense we can make of the notion of a biologically dysfunctional emotional mechanism that is insulated from the way(s) by which emotional experiences come to actually bear content or be meaningful, for it is only by being thus insulated that an emotional experience could be the product of a pathological mechanism and yet still be meaningful. Being meaningful, it can come to experienced as internal to self, its pathological nature notwithstanding. However, the coherence of this notion of independence looks doubtful given that it is presumably on account of failing to be hooked up in the appropriate ways to the environment that the emotional mechanism in question generates experiences that systematically fail to be representationally accurate and volitionally efficacious—thus casting these experiences outwith the register of meaningfulness—therein reducing its subject's biological fitness in evolutionary time and is on these grounds deemed dysfunctional. In other words, it seems integral to the very idea of a dysfunctional emotional mechanism that it generates emotional experiences that lie outwith the register of meaningfulness.

Whilst these difficulties may ultimately be surmountable, what they show us is that the ideas that together constitute the strict construal of the pathology concept are tightly interconnected and thus that we should be cautious about assuming that the concept of pathology is straightforwardly consistent with the rejection of any one (or more) of these ideas. (As for the analogy that Wakefield draws with personality disorders, the discussion of personality disorders in Chapter 5 (Section 5.3.3.) suggests that, far from settling the matter, the analogy only raises further questions about the putative consistency between the concepts of disorder and self-identity).

From the first two ideas we get a further idea: the condition of pathology is something we are passively struck by, and to the degree that this detrimentally affects our ability to exercise our voluntary and autonomous agency, the condition is viewed as an obstacle or an interference to, or a distortion of, our agency in this sense. This is why the condition of pathology is generally taken to be an excusing condition: when an act is the product of pathology, the subject of the act is seen as acting thus *in spite of herself*, and so cannot be held responsible for the act. This picture of the relationship between pathology and agency most obviously informs our understanding of what happens in physical conditions that cause involuntary movements, such as Parkinson's disease, but the picture is also present in our thinking about mental disorders, e.g. as is evident in the insanity defence in the law. The second idea noted above—that the pathological condition is viewed as external to the subject's real self—is what helps us to formulate this third idea, i.e. that a pathological condition that affects our agency does so as an obstacle, a hindrance, or as a distorting influence thereto.

Another idea embedded within the strict construal of pathology is that a condition of pathology is negatively valued. Illnesses may range from being unpleasant and inconvenient, all the way to being life-threatening. While this fourth idea is so obvious as to seem truistic, it is worth mentioning because it gives rise to a fifth: a pathological condition is something we would normally seek to remove. Furthermore, we usually try to remove it by interventionist means. By this is meant that we try and manipulate key variables that are believed to be are causally efficacious in removing the condition. This applies fairly straightforwardly in the case of conditions of physical pathology, but the idea becomes more contentious in the case of mental pathology since these involve mental states and experiences that appear to have content. There is a *prima facie* call to engage with these experiences hermeneutically, by plumbing their depths and inquiring into their meaning. In light of the first idea, though, it would be a mistake to answer this call directly. If the experience is symptomatic of a pathogenetic process, then there is no meaning (in the relevant sense) to inquire into: the having of the experience is determined by processes that lie outwith the register of meaningfulness altogether. The right response, therefore, would be to remove it interventionistically, as we would a condition of physical pathology.

I have presented five ideas that form part of a conceptual package which is widely associated with the concept of pathology—a package that delivers a strict construal of this concept. Drawing them together, the five ideas are as follows.

A condition of pathology or illness is a condition:

i. the having of which lies outwith the register of meaningfulness;
ii. that is external to the core or real self;
iii. to which its subject is passively related;
iv. that is negatively valued;
v. that calls for its removal by interventionist means.[6]

So according to this strict construal of pathology, when an experience of intense anxiety (say) is judged to be disordered, the fact of the subject's having the experience is deemed to be meaningless, for pathological anxiety in no way indicates danger or the presence of a threatening stimulus condition (or else it does so but to a degree that is wildly disproportionate to the actual threat present). It is being generated by a pathogenetic process, rather than tracking actual danger in a reliable way as normal anxiety does. Furthermore, because the experience is said to be meaningless in this way, it is external to the subject's real self, and it is something by which she is passively struck. The fact that this experience of intense anxiety is distressing (whilst also being deemed meaningless) naturally leads to its being negatively valued, and because the experience is deemed to be meaningless and externally situated, the appropriate response to it would be to seek to remove the experience by interventionist means.

Although Kendler himself neither explicitly endorses nor rejects this strict construal of the pathology concept, in what follows I will reconstruct the position he represents in the dispute between himself and Horwitz and Wakefield in light of this construal in order to bring into sharp focus the contrast between viewing distressing emotional experiences through the lens of pathology and without it. That is to say, in prioritizing the negative appraisals made of a distressing emotional experience in the dimensions of valence and agency over the positive appraisals made of the experience in the dimensions of rationality and/or existential significance—and on this basis asserting that such experiences are rightly deemed to be disordered—an advocate of Kendler's position would be understood as saying that the subject's having of the experience is meaningless (even though the experience itself may *appear* to be a rational response to her circumstances or to bear existential significance); that the experience is externally situated relative to the subject's real self, that the experience is something she is passively struck by, etc. To reiterate, the point of understanding the judgement of emotional disorder in these terms is that the strict construal gives the concept of pathology

[6] The assumptive stance towards illness described in point iv. is what grounds the very enterprise of healthcare, giving it its foundational goals of treating illness and alleviating suffering.

substance: in this way, it helps us to better understand the significance of each option, in deciding which basic evaluative dimensions to prioritize as we work out how best to delineate the boundary of emotional disorder. This will be the subject of investigation in Sections 3.4–3.5 below, and in the next chapter. To embark upon this inquiry, I will begin by presenting the psychoanalyst James Hollis' account of distressing emotional experiences—specifically, depressive experiences—according to which certain experiences of this kind are viewed as bearing existential significance. I will argue that this fact about them provides us with good reason to resist the meta-evaluative judgement that such emotional experiences are pathological.

3.3 Case Study: 'Depression's Therapeutic Gift'

It is a long-standing idea (one we find in Greek antiquity) that emotional distress, at least in certain forms and under certain conditions, is an inherently valuable part of the human experience.[7] This idea has been explored and developed in a multitude of ways. To try to comprehensively enumerate these ways would be an extensive task, one that lies well beyond the bounds of the current inquiry (but see Davies 2011 for a discussion). It is important, though, to gain a reasonably vivid understanding of this idea as it applies to our emotional experiences, in order to achieve a clearer understanding of what is in danger of being dismissed or overlooked by the move to pathologize distressing emotional experiences. It is in order to obtain this understanding that I shall now sketch out the psychoanalyst James Hollis' account of the value of certain depressive experiences (Hollis 2005).

In his discussion of depressive experiences, James Hollis distinguishes between the following three kinds of experience we might think of as 'depressive':

1. Biologically based depression.
2. Reactive depression (a reaction to loss).
3. Depression rooted in intrapsychic conflict ('DRIC' for brevity, thus: 'depressive$_{\text{DRIC}}$ experiences').

Hollis' discussion focuses on depressive experiences of the last kind. It would be a mistake to interpret Hollis' depressive$_{\text{DRIC}}$ experience as an exotic kind of human experience, separate and distinct from the kinds of experience

[7] In the *Oresteia*, Aeschylus writes: 'Zeus, who guided mortals to be wise, has established his fixed law—wisdom comes through suffering' (Aeschylus 2007).

referred to in the DSM's section on depressive disorders, for Hollis states explicitly that he is offering an alternative construal of depressive experiences that lie within the scope of psychiatric diagnostic manuals. In the register of DSM language, depressive$_{DRIC}$ experiences are according to Hollis known as dysthymic disorder (in the latest edition, this category has been termed 'Persistent Depressive Disorder').[8] So while Hollis is in agreement with the DSM about the pathological status of biologically based depression (the first of three sorts of depressive experience identified above), Hollis holds that both reactive depression and depressive$_{DRIC}$ experiences are inherently meaningful experiences. In each case, the experience has something important to tell its subject, and it is in this 'telling' that the experience is said to bear existential significance. On Hollis' view, therefore, it would be a mistake to pathologize the latter two kinds of depressive experience. Insofar as the DSM pathologizes all three kinds of depressive experience, Hollis and the DSM should thus be understood as making competing claims about the same region of our emotional lives.

Having clarified this point, we can now turn to Hollis' account of depressive$_{DRIC}$ experiences. Hollisian depressive$_{DRIC}$ experiences are classified as dysthymic disorder, and the key diagnostic features of dysthymic experiences listed in the DSM include depressed mood (this should persist for at least two years), loss of appetite, low energy, low self-esteem, and feelings of hopelessness (APA 2013). Hollis summarizes the condition being described here by stating that depressive$_{DRIC}$ experience is marked by 'an absence of or disturbance of strong affect for the conduct of one's life.' Then he puts forward his main claim: this form of depression is a manifestation of the autonomy of the psyche. Hollis explains this claim in the following way:

> The ego, the conscious sense of who we are, wishes to invest energy in a certain direction, perhaps in service to economic goals, but the soul has another agenda. [So] it autonomously withdraws the invested energy, inverts it, and as it withdraws into the psyche it often pulls the ego in after it . . . . A close cousin to this form of depression is boredom, or ennui, which means that the object or the goal that has carried our projections of psychological energy thus far no longer sustains the agenda of the soul. Even what may have been a good choice at one point has now been served, the task exhausted, and the psyche demands renewal, or greater balance, through investment in other values. (Hollis 2005: 74)

[8] 'Actually, this garden variety of depression is today called *dysthymic disorder* in the psychiatric manuals' (Hollis 2005: 74).

The terms 'ego' and 'soul' call for brief clarification. Hollis uses the term 'ego' to refer to the self that shows up and functions in daily life—the self that goes to work; interacts with family, friends, and colleagues; pursues pastimes; etc. This ego self tends to be shaped by conditions we intuitively think of as heteronomous: parental expectations, habits of upbringing, insecurities concerning our basic needs (these include social needs, such as the need for approval). The term 'soul', by contrast, is used to refer to that aspect of ourselves from which originates our more authentic values and concerns. To live in harmony with the soul's direction or calling (or more poetically, 'longing') would be to live the life we'd truly wish to live—for instance by pursuing a vocation we find genuinely interesting and fulfilling, this being the main or only basis for deciding to pursue it (thus: without this decision being overly influenced by worries about whether the job we do brings us enough money to maintain our current social status, say); choosing a life partner who inspires our love (thus: without this decision being overly influenced by concerns about parental approval, say).

If this sounds a little idealistic, it is because Hollis is setting out an ideal that he thinks is worth pursuing and which he takes to be essential for human flourishing. Hollis acknowledges that, in reality, pursuing this ideal is far from easy. One challenge we face is to distinguish between the various things that we are influenced and shaped by, i.e. between those things that genuinely matter to us and those things that we mistakenly think matter to us. A more fundamental difficulty in pursuing this ideal is that we may not know what our soul's calling is. That is, we may not know what it is we value most deeply. The conditioning of our ego selves may be so entrenched that it seems to us as if there is nothing else to our sense of self other than the dictates of the ego. This predicament is what makes the identification of a depressive experience as an instance of DRIC so challenging. If we haven't the faintest notion of what we value most, when we have no sense of ourselves beyond the daily functioning of our ego selves, then a depressive experience that is an embodiment of an intrapsychic conflict can look straightforwardly to be an instance of a biologically based depression. It is far easier to recognize a depressive$_{\text{DRIC}}$ experience as such when we have a sense, however vaguely, of what we value most, and when we suspect that we are living in a way that compromises these values.

In any case, Hollis contends that not to pursue this ideal of harmony is to put oneself on course to landing in intrapsychic conflict, in which one's ego-self stands in conflict with one's soul-self. This is a life that 'constricts meaning', and depressive$_{\text{DRIC}}$ experiences are the embodiment of this intrapsychic conflict. Hollis provides the reader with a range of vignettes from his practice as

a psychotherapist to illustrate this understanding of the depressive$_{DRIC}$ experience. Here is one:

> I think of a woman who, the child of two psychiatrists, grew up to be a psychiatrist in order to win their approval, neglecting the fact that her soul had another plan. Her true talent and calling was found in the arts, and while she was a caring and competent psychiatrist, her midlife depression deepened with each passing year. One might say that with each year her soul was further exiled from her constructed world, her depression grew as a sign of the psyche's protest. She was living in constrictive service to parental complexes, as we mostly do, and not in service to the larger summons of her talent. Why would she not be depressed? She was very good at dispensing medication to others who suffered biochemical depressions, but was so close to her own problem that she could not recognise intrapsychic depression when she saw it. (Hollis 2005: 75)

Distressing as the depressive$_{DRIC}$ experience is, it is a 'therapeutic gift', according to Hollis. What does he mean by this? Of another client in a similar situation (this time, choosing a direction in life—including her profession—in opposition to her parents' instructions, but in this way still being defined by their plans for her), Hollis remarks:

> Why wouldn't she be depressed at such a predicament? Why would we not expect the body to revolt, and the psyche to withdraw energy from the place where the complex-driven ego wishes to invest it? Yet this troubling insurgency in the psyche is a friend, for it brings to [this client] an accounting, and the possibility of greater consciousness. At this writing, she is conscientiously sorting through her choices, learning which ones are really hers and which ones derivative. It is a discernment process that must continue for the rest of her journey, as must ours. (Ibid.: 20)

On the view Hollis is presenting, depressive$_{DRIC}$ experiences are the embodiment of a kind of intrapsychic conflict, i.e. one that emerges from serious mismatches between our outward circumstances and our most deeply held values. So these experiences can serve as the means by which the depressed subject comes to know that her outward circumstances are grossly misaligned with what matters most to her. In this way, the depressive$_{DRIC}$ experience is taken to bear existential significance.

Construing depressive$_{DRIC}$ experiences in this way presents the subject with the following choice (a choice that might not have been apparent before): either to alleviate her emotional suffering immediately by interventionistically

removing the distressing affect that characterizes her depressive$_{DRIC}$ experience (through pharmacological means, say), or else to discern its existential significance and to engage the experience hermeneutically, i.e. by seeking to understand the meaning of the unease she suffers from.[9] The availability of the latter option—made visible by the conception of depressive$_{DRIC}$ experiences presented by Hollis—provides a key source of resistance to the meta-evaluative judgement that a depressive experience of this kind is pathological. To see why, I will further unpack this view of the depressive$_{DRIC}$ experience by juxtaposing it against the claims made about this experience by the judgement of pathology (in accordance with the strict construal of the concept of pathology).

3.4 Discussion of Hollis' Account of Depressive$_{DRIC}$ Experiences

Depressive experiences in general tend to get pathologized in psychiatry because of negative appraisals made of these experiences in the evaluative dimensions of affective valence and agential efficacy. When a depressive experience is also negatively appraised in the dimensions of emotional rationality and of existential significance, the move to pathologize this experience is unproblematic since its being irrational and being deemed to bear no existential significance is consistent with viewing it in the terms set out by the strict construal of pathology in Section 3.2 above, i.e. the having of this experience is viewed as meaningless and as being located externally to the subject's real self.

But a depressive$_{DRIC}$ experience is not like this. The having of such an experience is viewed by Hollis neither as meaningless nor as external to the subject's real self. Instead, the experience is understood to bear existential significance in virtue of being an expression of an intrapsychic conflict. The term 'expression' is chosen carefully. It seems a little too strong to claim that the content of the depressive$_{DRIC}$ experience is that the subject is suffering from serious intrapsychic conflict, since the experience doesn't refer to it or make an assertion to this effect. Yet at the same time, the connection between the intrapsychic conflict and the depressive$_{DRIC}$ experience seems to be stronger than that of mere indication, e.g. as akin to the connection between smoke and fire. Rather than being just a consequence of the intrapsychic conflict, the experience seems to be an embodiment of it. The subject lives out this

[9] I am drawing on Charles Taylor's use of the term 'hermeneutic' (Taylor 2007). More on this shortly.

conflict through the depressive experience. So although the way in which a depressive$_{DRIC}$ experience bears existential significance may differ from the way in which an experience of intense fear bears the content 'This situation is extremely dangerous', it would nevertheless be a mistake (on Hollis' view) either to say that the depressive$_{DRIC}$ experience is devoid of meaning or to view the having of the experience as a meaningless occurrence. On the contrary, an experience of this kind serves as an essential means through which we learn to recognize that our outer circumstances are grossly misaligned with our most deeply held values, when indeed we find ourselves in a predicament of intrapsychic conflict.

Making a similar point, the philosopher Charles Taylor suggests that the depressive$_{DRIC}$ experience can be understood as a perception of the misdirection in the subject's life. Contrasting what he calls a 'spiritual perspective' on this experience with a 'therapeutic perspective', Taylor states:

> In one case [the spiritual perspective], they ['our experiences of un-ease, anguish, emptiness, division, and the like'] may be telling us something important; they may be revealing some lack or misdirection in our lives. In the other [the 'therapeutic' perspective], they are akin to illness, and as such may be *symptomatic* of some mistaken direction; . . . they don't constitute a . . . *perception* of this mis-direction. (Taylor 2007: 621, emphasis in original)

A quick terminological point: slightly confusingly, Taylor refers to the perspective on depressive$_{DRIC}$ experiences that is similar to the one Hollis offers us as the 'spiritual perspective'. In fact, given that Hollis (writing as a psychotherapist) contends that the experience of anguish and emptiness can be revelatory, what Taylor calls the 'therapeutic' perspective may for our present purposes be better termed the 'pathology' perspective.[10] With this point clarified, we see in the passage just quoted a statement of the idea that depressive$_{DRIC}$ experiences aren't merely a consequence of suffering a misdirection in our lives—a misdirection that consists for instance in a misalignment between one's outward circumstances and one's most deeply held values—but rather that such experiences stand in a significance-bearing relationship to this misdirection. We come to be aware of this misdirection through the experience of anguish, emptiness, and division (say). It follows from this, according to Taylor, that

[10] The source of this confusion is that the notion of psychotherapeutic engagement is ambiguous between engagement that is informed by the spiritual and the pathology perspectives on depressive$_{DRIC}$ experiences respectively, both of which may be deployed within the practice of psychotherapy. I will return to this important point in Section 5.3.

'In one case [viewed from the spiritual perspective], the unease needs to be further understood, worked through, perhaps in prayer or meditation; in the other [viewed from what I am now calling the pathology perspective], it needs to be got rid of, or at least rendered mild enough to be lived with' (ibid.). Once we are in the grip of the pathology perspective, it can then seem reasonable to seek to remove such distressing emotions entirely by interventionist means.

Because of what it is said to be an embodiment of—namely, a misalignment between a subject's outward circumstances and the concerns and interests that matter most to her—the depressive$_{DRIC}$ experience is placed firmly within the boundary of the subject's real self when the experience is regarded from the spiritual perspective. How so? The misalignment is an integral part the subject's real self since it is constituted by the conflict between her most deeply held values (values that define this self) and the outward circumstances of her daily life. So in contrast with the strict construal of pathology that places depressive$_{DRIC}$ experiences outside the boundary of the subject's real self, the depressive$_{DRIC}$ experience on this view is situated right at the heart of the real self, giving it its distinctive shape (even if the subject of the experience is yet unaware of this).[11] Relatedly, if depressive$_{DRIC}$ experiences are an important means by which at least some people begin to undertake the task of working out which of the values they currently live by are authentically theirs, then it is through hermeneutically engaging such experiences that they come to be in a position to exercise greater personal autonomy.

All this has an important bearing on the question of passivity. By now it should be clear that on Hollis' view, the depressive$_{DRIC}$ experience is not a heteronomous phenomenon fixed by pathogenetic processes the having of which is devoid of meaning, a (literally) senseless pathogen that strikes the subject as a bout of flu might. The flu, as it immobilizes, stands dumbly in the way of our trying to get on with our daily tasks, and a response that consists in removing it in an interventionist manner looks to be of the right shape for this reason. While the anhedonia and immobilization in depressive$_{DRIC}$ experience may be experienced in a similarly passive way, the fact that the depressive$_{DRIC}$ experience bears existential significance (by embodying an intrapsychic conflict or some other existentially significant misdirection) and is an integral part of the subject's real self entails that the subject is actively poised in at least the following way. One's real self is revealed in the depressive$_{DRIC}$ experience (compare: it makes no sense to say the same of flu, i.e. that it reveals our real selves to us); and the grasp of this possibility marks a profound shift in the subject's

[11] This point is developed further in Section 4.6.3.

orientation towards this experience. It reanimates her emotional life in the specific sense that this understanding casts even the flattest anhedonic states in a new light: it becomes possible to regard it as a valuable source of self-knowledge. In this way, it is an experience about which curiosity, for instance, becomes an apt stance; more than that, it is an experience that comes to be recognized as being worthy of our curiosity. The depressive$_{DRIC}$ experience thereby becomes a summons to the subject to discover her real self—a summons the subject can only recognize as such by adopting a hermeneutic stance within and towards the experience (cf. Taylor 2007: 621). It is the availability of this stance—even if it is in fact not adopted by a particular subject—that transforms the depressive$_{DRIC}$ experience from one that is undergone passively to one that can be actively navigated. In the next chapter, I will situate this work of self-discovery within a broadly Aristotelian framework for thinking about human flourishing, and on this basis I'll suggest that the adoption of a hermeneutic stance towards our depressive$_{DRIC}$ experiences within a psychotherapeutic context forms part of the process of *enriched habituation* (see Section 4.6.1).

3.5 The Significance of the Choice to Pathologize

What emerges from the foregoing discussion is that the decision to pathologize or not to pathologize depressive$_{DRIC}$ experiences amounts to a decision about whether to place these experiences within, or outwith, the boundary of the subject's real self. A depressive$_{DRIC}$ experience is by definition positively appraised in the dimension of existential significance by Hollis, and because this appraisal consists in viewing the experience as situated firmly within the boundary of the real self—indeed, it lies right at the heart of this self, on Hollis' account—a positive appraisal in this basic evaluative dimension constitutes strong grounds for resisting the pathologization of depressive$_{DRIC}$ experiences (on the strict construal of pathology, according to which conditions of disorder are understood as being externally situated in relation to the self).

As it determines how we view the relationship between the depressive$_{DRIC}$ experience and the self (i.e. as an external or an internal one), the decision to pathologize turns out to be a personally and morally significant one for the following reason. In choosing between (i) negative appraisals in the dimensions of valence and agency (pathologizing depressive$_{DRIC}$ experiences on this basis), and (ii) a positive appraisal in the dimension of existential significance (resisting pathologization on this basis), what we are effectively choosing

between is pleasant affect and effective agency, on the one hand, and on the other, growing more psychologically mature by (amongst other ways) coming to know oneself better and becoming psychically more unified and, through this, perhaps also becoming more personally autonomous.[12] The strategy for defending pathologization mentioned in Section 3.1 above, which consists in claiming that the negative appraisals made in the dimensions of valence and of agency are sufficient to ground the judgement of emotional disorder, can now be understood within the context of this choice. The claim of sufficiency, recall, is anchored in the assertion that there is either nothing of significant value in the suffering wrought by depressive experiences, or that although they may be inherently valuable, this matters less to the judgement of emotional disorder than the fact that these experiences are intensely distressing, perhaps to the point of being immobilizing. We can now specify these evaluative claims in more concrete terms, understanding the proponent of psychiatry to be saying one of the two following things:

1. Adopting an open-minded stance towards Hollis' account of depressive$_{DRIC}$ experiences, he might question the value of psychological maturation per se.
2. Accepting the claim that psychological maturation is an inherently valuable process, he might nonetheless reject Hollis' suggestion that depressive$_{DRIC}$ experiences are a *necessary* path towards psychological maturation.

I will examine each of these moves in turn.

3.5.1 How Important Is Psychological Maturation?

According to the first strategy just identified for defending the pathologization of distressing emotions that, the proponent of psychiatry is understood as denying that psychological maturation is ever *more* valuable than the value inherent in living a life permeated by pleasant affect and effective agency. This defence of pathologization rests upon on a commitment to a substantial hierarchy of values, according to which greatest value is placed on the preponderance of pleasant affect and effective agency. Such a value hierarchy ultimately prescribes a particular vision of the ideal human life. According to this ideal,

[12] The theme of psychological maturation will be explored more extensively in the next two chapters.

the best life it is possible to live is one according to which we enjoy pleasant emotions and in which we exercise our agency effectively and seamlessly, most if not all of the time.

With this vision of flourishing in place, it is possible to reinforce the defence of pathologization of distressing emotions by arguing that we make progress in our response to these states by pathologizing them, as follows. By alleviating suffering—suffering that turns out in light of this value hierarchy to be of limited and questionable value—the pathologization of distressing emotions moves us ever closer towards the ideal of flourishing, as set out by this particular hierarchy of values. In the next chapter, I will explore in much greater depth the way that our conceptions of flourishing inform and shape the relative weightings we assign to the appraisals we make of our emotional experiences along the basic evaluative dimensions. For the moment, I want to note the following point: this ambitious claim commits itself to a normative view about the kinds of creature we are, namely that we are subjects for whom psychological maturation is of secondary importance relative to living a life permeated by pleasant affect and effective agency. Those who hold the contrary view—namely, that psychological maturation is at least as important as the preponderance of pleasant affect—precisely challenge this conception of human flourishing. Sympathy with Hollis's account of depressive$_{DRIC}$ experiences motivates the worry that the alleviation of emotional distress (wholly) by pathologizing it—and then seeking to remove it through pharmacological means—threatens to eliminate an essential chunk of our humanity. This is done by externalizing the distressing emotion(s) we experience, distancing them from our sense of who we are, and on this basis, encouraging us to adopt an interventionist stance towards these emotions (thereby aiming exclusively to remove them mechanically). Viewed from the Hollisian perspective, the move to pathologize depressive$_{DRIC}$ experiences is fraught with moral and existential difficulties because the chunk being eliminated (i.e. the depressive$_{DRIC}$ experience) relates essentially to something we would usually recognize as having paramount value: psychological maturation as manifest through deepening self-knowledge.

Those hoping to defend the pathologization of distressing emotions in this manner are thus presented with the challenge of defending this normative view of human life. The legitimization of psychiatry on this route cannot be defended from within the scope of psychiatry itself (by which I mean psychiatry as it is practiced, and psychiatric science as a medical research programme) since there is no single substantive conception of flourishing that psychiatry overtly subscribes to, and it is only in light of just such a specific

substantive conception of human flourishing (according to which psychological maturation is of secondary importance) that it makes sense to claim that distressing emotions are, on account of their distress alone, pathological in nature. Relatedly, the claim that psychiatry makes progress by pathologizing distressing emotions is made intelligible only by reference to this conception of flourishing. I will fill out one influential version of this conception of flourishing in Section 4.2 of the next chapter.

3.5.2 Are depressive$_{DRIC}$ Experiences Necessary for Psychological Maturation?

A less ambitious response to Hollis' account consists in agreeing with Hollis that psychological maturation is an inherently valuable process but rejecting the idea that depressive$_{DRIC}$ experiences are a *necessary* path towards psychological maturation. In this way, the distress in such experiences is understood to be superfluous, containing nothing of value in its own right. This establishes the sufficiency of the negative appraisals in the dimensions of valence and agency in anchoring the judgement of emotional disorder.

Like the previous response, this response may also be bolstered by an appeal to the notion of progress. If distressing emotions have no intrinsic value (because the distress is unnecessary for psychological maturation), then to treat them as if they do have such value is to adopt a fundamentally misguided view about the role these emotional experiences play in our lives. Pathologization, insofar as it eradicates misguided or superstitious beliefs about distressing emotions, constitutes progress.

The claim that psychological maturation may be achieved in ways that do not need to involve emotional distress may be rooted in a suspicion of depth-psychology. The charge here would be that the depth-psychological framework is itself dubious (lacking robust scientific credentials), so that any claim it makes about the value of emotional distress must also be doubted. It is with this qualm in mind that I have selected James Hollis' account of depressive$_{DRIC}$ experiences. The account is 'Jung-lite': there is no need to buy into the whole apparatus of depth-psychology to see what Hollis is getting at. The situations he describes are intuitive, familiar, easily recognizable. The greater the misalignment between our outward circumstances and our deepest concerns, values, interests and passions, the stronger our feelings of listlessness and our

sense of despair. We will naturally feel increasingly anhedonic, trapped, and despondent.[13]

Alternatively, the denial that depressive$_{DRIC}$ experiences are a necessary path towards psychological maturation for at least some people may be meant in a more absolute sense. The proponent of psychiatry may grant that Hollis has identified a phenomenon that occurs, perhaps even with considerable frequency, but which happens merely contingently. This is shown by the conceivability of living in a life in which all these values are simultaneously instantiated—that is, a pleasant and happy life that is lived authentically, in which we exercise our agency effectively *and* autonomously. Some of us may indeed be fortunate enough to inhabit this possibility. However, the claim that this is a real option for absolutely everyone seems implausible since it depends on us all being blessed with felicitous outward circumstances: loving and wise parents, financial security, an enlightened culture that fosters psychological maturation and encourages us as a matter of course to discover what our deepest values are, etc. Most of us, even in materially comfortable settings, do not enjoy altogether favourable outward circumstances in this way. Friction between outward circumstances and our inner lives is the more predictable route many of us will have to take. Where there is friction, the ground is laid for conditions of intrapsychic conflict to grow, and Hollis' point is that when we do find ourselves in this predicament of intrapsychic conflict, our emotional suffering and distress becomes an essential means by which we may grow psychologically more mature. That this suffering is contingent in some absolute sense matters little here.

In noting that intrapsychic conflict is a predicament many of us will likely undergo at some point in our lives, it is worth remarking upon a peculiar feature of depressive$_{DRIC}$ experiences. Such experiences are experiences we nearly always find ourselves thrown into, rather than being experiences we actively go in search of, let alone choose to have. This thrownness into a depressive$_{DRIC}$ experience suggests that, while there may be paths to self-discovery other than through such an experience, in practice it would be to miss the point to suggest that the suffering wrought by a depressive$_{DRIC}$ experience is optional (and thus superfluous) in view of the hypothetical availability of these alternative paths. If a person is suffering from a depressive$_{DRIC}$ experience (rather than a depressive experience with no existential significance), then that is the predicament she already finds herself in. The intrapsychic conflict is, *ex hypothesis*, showing up for her at this point in time in the form of a depressive$_{DRIC}$

[13] See for instance Johann Hari's personal testimony in Hari 2018.

experience. Simply to say that a pharmacological intervention is what is called for in response to this experience, to alleviate the negative valence and immobilization brought on by it, and that a hypothetical alternative path may be pursued to resolving the intrapsychic conflict instead, is to misconstrue the actual existential predicament with which the depressed subject is grappling. Experiencing the negatively valenced emotions that constitute a depressive$_{DRIC}$ experience is the only direct means available to the subject of feeling the weight and the gravity of the misalignment between her daily activities and what she most deeply values. To dim down the intensity of these emotions too drastically would be to threaten to disconnect her experientially from the reality of her present existential predicament. Kendler's claim about over-romanticization of emotional suffering begins, from this perspective, to look like a trivialization of this suffering—where this means, specifically, that Kendler's judgement of emotional disorder wrongly distances the depressive$_{DRIC}$ experience from the subject's sense of her (real) self, and on this basis, misleads her into adopting an interventionist stance towards it. Therein the subject fails to grasp the existential import of her depressive$_{DRIC}$ experience and its proximity to her real self. All this is not to deny that pharmacological interventions may have an auxiliary role to play in mitigating the subject's suffering. It is rather to underscore the point that the suffering is not something accidental to the call towards self-discovery and therefore to be thought of as something optional that may be removed altogether whilst the subject still claims to be responding to this call by meaningfully engaging with it.

So far, I have been examining the attempt to defend the pathologization of distressing emotional experiences from within the BED framework, i.e. by negotiating the competing appraisals made of these experiences in the dimensions of valence and agency, and in the dimension of existential significance, respectively. To this end, I have probed Kendler's move to prioritize the former two dimensions over the latter dimension on the grounds that to do otherwise would simply be to over-romanticize emotional distress. I have drawn on James Hollis' account of depressive$_{DRIC}$ experiences to articulate two interpretations of this charge of over-romanticization. The first interpretation consists in assigning less weight to the value of psychological maturation than to the value of pleasurable affect and effective agency, the second consists in acknowledging the primary importance of psychological maturation but denying that emotional suffering is necessary for achieving this. The upshot of this discussion is that neither interpretation conclusively disposes of the claim that emotional suffering can be inherently valuable in virtue of bearing existential significance. This means that we do not as yet have grounds for taking

negative appraisals in the dimensions of valence and agency to be sufficient for anchoring the judgement of emotional disorder.

There is one further move that the advocate of pathologization might make, however, and that is to invite us to reflect on the ultimate value of psychological maturation. Granting its importance, the pathologization advocate might argue that we value psychological maturation ultimately because such maturity enables us to lead happier, subjectively more satisfying lives—at the very least, by reducing the sources of friction, frustration, despondency, and distress we would otherwise encounter. This way of thinking about the value of psychological maturation returns us to the valence quality of our emotional experiences as our basic reference point, and having this as our primary reference point encourages the prioritization of this evaluative dimension over all others (including the dimensions of existential significance and of rationality). This in turn leads easily to the thought that it matters little, ultimately, whether we pathologize a depressive$_{\text{DRIC}}$ experience and respond to it in an interventionist manner, or whether we resist pathologizing it and engage with it hermeneutically, since our ultimate goal *either way* is the one that guides the psychiatric enterprise from the outset: to alleviate the suffering wrought by negatively valenced emotions and to increase the preponderance of positively valenced emotional experiences in our lives. In the next chapter, I will scrutinize this understanding of what is at stake in negotiating the five basic evaluative dimensions, doing this by contrasting two very different conceptions of human flourishing: the one being presupposed in this interpretation (which I shall refer to as the health-based conception of flourishing), and a conception of flourishing that has its roots in Aristotelian philosophy. Once we have both conceptions in view, we will be better placed to appreciate that it is possible to view each of the two ways of negotiating the basic evaluative dimensions explored in this chapter as irreducible to the other.

3.6 Conclusion

In this chapter, I have identified some of the challenges that confront those who seek to defend the pathologization of distressing but existentially significant emotional experiences on the grounds that their distressing and immobilizing quality is sufficient for anchoring the judgement of pathology. One way of trying to address these challenges is by loosening the concept of pathology—for instance, by asserting that an emotional experience can be both disordered *and* meaningful—but this approach threatens to diminish

the clarity and the substance of the concept. We should thus be willing to do this only with good reason, i.e. revisions to the strict construal of pathology should be earned, not presumed. A different way of responding to these challenges is to address them head-on, i.e. by accepting the parameters set by the strict construal and then arguing within these parameters for the sufficiency for the judgement of emotional disorder on the basis of negative appraisals made of distressing emotional experiences in the dimensions of valence and agency exclusively. Either way, the discussion in this chapter has shown that, as we negotiate the various appraisals made in the basic evaluative dimensions and try to work out where best to place the boundary of emotional disorder, we inevitably find ourselves navigating a rich evaluative space.

4
Emotions and Human Flourishing

> We see that we have a way of speaking of goodness in human beings not corresponding to anything in the other cases. I am thinking now of the thoughts expressed by sentences in which the word 'good' is joined to the name of the species itself, as when we speak of 'a good human being' or, more colloquially, 'a good person'. There is no equivalent to this in the language in which we evaluate plants and animals. Firstly, in so far as we do speak of 'a good S' in these other cases (where, incidentally, we tend rather to say 'a healthy S' or 'a good specimen of an S'), we are thinking about the plant or the animal as a whole; whereas to call someone a good human being is to evaluate him or her only in a certain respect. And . . . this particular evaluation can only be of human beings. For to speak of a good person is to speak of an individual not in respect of his body, or of faculties such as sight and memory, but as concerns his rational will.
>
> **Foot (2001: 66)**

4.1 Introduction

It might be argued that it doesn't ultimately matter which way we weigh up the appraisals made of a depressive$_{\text{DRIC}}$ experience along the basic evaluative dimensions, and whether on this basis we come to view the experience as pathological or otherwise, because the desired outcome of whatever intervention adopted is the same in both cases: the alleviation of the subject's suffering and the promotion of her happiness. Given this supposition, it can seem tempting to suppose that both routes explored in the previous chapter ultimately come to the same thing: the subject of a depressive$_{\text{DRIC}}$ experience wishes to become happier, to lead a more fulfilling life, and it matters little in the end whether she achieves this through a process of self-examination leading to psychological growth or by externalizing the experience and manipulating it by means

The Concept of Emotional Disorder. Gloria Sibson Ayob, Oxford University Press. © Oxford University Press 2024.
DOI: 10.1093/9780198909620.003.0005

of a pharmacological intervention or a talking therapy. The outcome either way is that, if successful, the subject comes closer to flourishing.

This apparent lack of a substantial difference between the two responses is rooted, I suggest, in the predominance of a single, specific conception of flourishing, one that controls our sense of what possibilities there are with respect to the ultimate regulative ideals in terms of which we appraise our lives at a global level. The aim of this chapter is to demonstrate that the full significance of the weightings we assign to the respective BEDs, in appraising our emotional experiences, emerges only when we view them in light of a given conception of flourishing. As I will show, the particular conception of flourishing we presuppose will shape the weightings we assign to the various appraisals we make of our emotional experiences along the basic evaluative dimensions. That this should be so isn't surprising since our emotions are taken to figure centrally in our flourishing—albeit in radically different ways—at least according to two influential conceptions of flourishing.

According to the first of these two conceptions, which I shall call 'the health-based conception of flourishing' (HF), what matters principally for our flourishing is whether we are enjoying a preponderance of pleasant emotional experiences both at a given time and over longer stretches of our lives. The HF has come to dominate much of our public discourse, and nowhere is this more evident than in the widespread tendency today to view unpleasant emotional experiences such as anxiety and sadness as disordered or pathological. But there is a very different way of understanding what it means to flourish, one that also takes our emotional experiences to be of central importance for our flourishing, although in a way that has little to do with their valence quality. According to this second conception—which I refer to as 'the excellence-based conception of flourishing' (EF)—it is primarily through the exercise of our epistemic, moral, and deliberative agency that our emotional experiences contribute to our well-being.[1] I'll now look at each of these conceptions of flourishing in turn.

[1] Before I begin, the following clarification is worth making. Whilst it is both important and interesting to ask what role if any emotions play in our best worked-out conception of human flourishing, I am examining a separate question in this book, namely this: what is the structure of our concept of emotional disorder? In the course of the inquiry, it emerges (as we shall see) that certain conceptions of flourishing decisively shape the weightings we assign to the appraisals we make of our emotional experiences along the basic evaluative dimensions, which in turn determines where we draw the boundary between pathological and non-pathological emotional experience. The specific conceptions of flourishing that I will explore below have been selected because they place emotions at the centre of their respective accounts, and thus serve as useful illustrations of the structure of the concept of emotional disorder (as this is revealed through the present inquiry). As their function is illustrative, I do not take myself to be arguing for the intrinsic worth of either of these conceptions of flourishing (doing so would take me too far afield into the debate about the nature of human flourishing, which lies outside the scope of this work). Ultimately, I suspect the

4.2 The Health-Based Conception of Flourishing (HF)

According to the health-based conception of flourishing, to flourish is to be healthy, where the term 'health' has expanded significantly, so that it no longer refers merely to physical health but now also includes psychological well-being. At the heart of the latter idea is the relative preponderance of pleasant emotional experiences and the relative paucity of unpleasant ones in a person's life.[2] This is a curious artefact of the HF, i.e. that psychological well-being, so understood, should come to be subsumed under the concept of health (with which the idea of flourishing is then equated). Why, for instance, do we not instead tend to think of flourishing as comprising two constitutive but irreducibly distinct elements, physical health and psychological well-being (without associating the latter with the idea of health)?

We find an answer to this question in a story told by an author in this tradition, a story that bears testament to the influence of psychiatric thinking in shaping this currently dominant understanding what it means to flourish. The psychologist Martin Seligman reports that he began researching what he calls 'positive emotions' in a deliberate attempt to move away from the emphasis in psychological research hitherto placed on mental illness, and specifically on conditions of disordered emotion such as depression and anxiety. Embarking from this starting point, Seligman sets off to find out how we can 'come to live in the upper reaches of our set range of happiness.'[3] In so doing, Seligman in effect assumes that emotional disorder and human flourishing signify the opposite ends of a single continuum, and that if emotions deemed negative are placed at the mental illness end of this spectrum, then what we should expect to find at the human flourishing end are emotions deemed positive. This is one way in which the idea of flourishing comes to be associated with that of health, i.e. by presupposing a binary set-up in which mental illness is juxtaposed against a concept that is taken to be its antonym, 'mental health'. The latter, being construed superlatively (and not just in terms of the idea of the absence of mental illness, as Seligman's talk of 'the upper reaches of our set range of happiness' indicates), is then used interchangeably with concepts

two inquiries are related (the arguments in this book establish a *prima facie* case for this claim), but further work needs to be done to connect them to each other.

[2] As one commentator observes: 'Within empirical psychology . . . the dominant model of well-being is subjective and hedonic. Well-being is indexed to pleasure attainment and pain avoidance. It is defined in purely subjective terms as a characteristic 'presence of pleasant emotions, the absence of unpleasant emotions, and the belief that life is satisfying in general' [in David et al. 2014, p. 216]' (Elliot 2016: 5).

[3] Seligman 2002: xii.

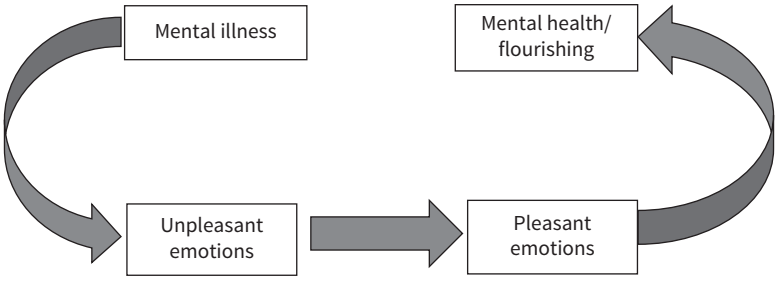

Diagram 4.1 The health-based conception of human flourishing. Note that this diagram represents nodes along a single continuum.

such as well-being and flourishing. It will be helpful to trace out this line of reasoning diagrammatically as follows:

The key move I wish to note through Diagram 4.1 is the extrapolation from whatever property is taken to be the essential mark of states located on the mental illness end of the continuum and projecting its imagined opposite onto the flourishing end. Doing this is what sets substantive parameters on the ensuing conception of flourishing. For Seligman, the states to be found along the entirety of the continuum that interests him are emotional states, and the property that Seligman thinks determines where along the continuum any emotional state should be positioned is that of its valence quality.[4] Thus, since mental illness (or more specifically, emotional) illness is taken to consist in suffering from unpleasant emotions—deemed to be negative on account of their unpleasantness—it follows that their counterparts on the opposite end of this spectrum, namely pleasant emotions—deemed to be positive on account of their pleasantness—are taken to be constitutive of mental health. Mental health, construed superlatively, is then treated as synonymous with flourishing. This is how the concept of flourishing comes to be understood as consisting in the enjoyment of a preponderance of positive emotional experiences and a relative dearth of negative ones.

Seligman is a founding father of the modern 'Positive Psychology' movement, and his way of framing the idea of flourishing has enjoyed wide acceptance. This may in part be due to Seligman's own influence in psychological research into human flourishing, but it is likely also due to the fact that it taps

[4] Martha Nussbaum, summarizing Seligman's thinking on this point, states: 'For Seligman, positive emotions . . . are those that feel good. So love would be positive, anger and grief negative, and so forth' (Nussbaum 2008: S92). Later Seligman adds character traits, but this addition subserves the basic story that accords our emotional states a central position since character traits are emotional dispositions. Slightly more on this in n. 6 below.

into an idea that is now deep-rooted in our thinking, i.e. the association between flourishing and feeling good. Indeed, the term 'positive psychology' itself was coined by Abraham Maslow as long ago as 1954. Maslow, like Seligman after him, was moved by a sense of disquiet about the focus in psychological research on 'man's shortcomings, his illnesses' (the 'negative side'), to start thinking about what might be found on 'the positive side' instead (Maslow 1954: 354). Summarising this general trend in the way the idea of flourishing is approached as a subject of investigation, Seligman writes:

> Psychology after World War II . . . concentrated on repairing damage using a disease model of human functioning. This almost exclusive attention to pathology neglected the idea of a fulfilled individual and a thriving community. . . . The aim of positive psychology is to catalyze a change in psychology from a preoccupation only with repairing the worst things in life to also building the best qualities in life. To redress the previous imbalance, we must bring the building of strength to the forefront in the treatment and prevention of mental illness. The field of positive psychology at the subjective level is about positive subjective experience: well-being and satisfaction (past); flow, joy, the sensual pleasures, and happiness (present); and constructive cognitions about the future—optimism, hope, and faith. (Seligman 2005: 3)

The conception of flourishing that is predominant today thus seems to be constrained in a fundamental way by the concept of mental illness. This is a not insignificant constraint because, by thinking about human flourishing as the antonym of mental illness, much else of what we value in life isn't going to naturally appear in the resulting picture of flourishing—domains of human evaluation that have little if anything to do with the domain of psychopathology, for instance evaluations pertaining to moral, deliberative, prudential, epistemic, or aesthetic qualities, all of which we may in fact think are relevant to flourishing (independently of the HF).[5]

Because flourishing is one of the key concepts (if not the main concept) we use to appraise our individual lives at the most global level, this constraint ramifies across all levels of concern and shapes our most basic understanding

[5] Seligman has subsequently refined his account of flourishing by seeking to incorporate 'strengths and virtues', reflected in his shift from talking about 'happiness' (as such) to speaking of 'authentic happiness'. Seligman describes this shift as being in part a response to Kahneman's empirical work on how people *recall* their happiness, but crucially, it is a response to people's rejection of the (imagined) opportunity to inhabit Nozick's (pleasure-inducing) 'experience machine'. Seligman interprets this rejection as expressing the view that 'It is not just positive feelings we want, we want to be *entitled to* our positive feelings' and that we are 'entitled to these feelings by the exercise of personal strengths and virtues' (Seligman 2002: 7–8, italics in original). Strengths and virtues, then, insofar as they make an appearance in Seligman's modified account of flourishing, do so on account of serving and fostering positive feelings, which remain the primary value in terms of which flourishing is understood.

of ourselves. We make sense of much of what we do in terms of flourishing thus construed.[6] What for instance does the subject of a depressive$_{DRIC}$ experience go to a therapist for? The HF suggests that he does this in the hope of feeling better—and this is consonant with why most people seek out psychotherapeutic or psychiatric help. If the depressive$_{DRIC}$ subject's therapist operates with the HF conception of flourishing too, then she (the therapist) will also understand her task to consist in helping her client feel better. The overarching concern of those who presuppose this conception of flourishing may be summed up by the following question: 'How can we reduce human suffering in individual lives and increase human happiness', where happiness is taken to consist in the enjoyment of subjective feelings with a positive valence?[7] It is against this benchmark that we come to view individual human lives either as flourishing or as languishing. Once flourishing is framed in these terms, then it can become difficult to see why it ultimately matters whether we weigh up the negative valence of a depressive$_{DRIC}$ experience more heavily than its positive appraisal along the dimension of existential significance. In other words, it becomes unclear why it ultimately matters whether or not we pathologize the depressive$_{DRIC}$ experience. It comes to the same thing eventually: construed as a pathological state or as a meaningful one, the goal either way is to shift the depressive$_{DRIC}$ experience (whether by pharmacological means or by engaging with the experience meaningfully), so that the subject is able to start feeling better, i.e. to begin to flourish.

It is not accidental both that this meta-evaluative judgement of emotional disorder tends to prioritize the BED of emotional valence *and* that the currently dominant HF conception of flourishing equates our global well-being with positively valenced emotional experiences, for the latter stems from grappling with mental illness (as we've seen in the summary above of Seligman's research strategy). One consequence of this, though, is that the HF-conception of flourishing makes it easier to argue in favour of weighting the basic evaluative dimension of valence more heavily against that of existential significance—where appraisals along the two dimensions are at odds with each other—thereby encouraging the application of the meta-evaluative concept of emotional disorder to the experience. At the very least (so the thought goes), we lose nothing of great importance by *not* giving priority to the positive appraisal of the experience along the dimension of existential significance since this is simply one further means of achieving (or trying to achieve) the

[6] Aristotle, *NE* 1097a15–1097b21.

[7] In the particular forms that suffering and happiness take that are germane to the psychiatrist and the psychotherapist, rather than the public policy-maker.

same end-goal of maximizing good feeling. In effect, the prioritization of the dimension of valence over that of existential significance is bootstrapped by the HF.

Here, then, we begin to see that there are three evaluative levels in terms of which we make sense of, and appraise, our emotional experiences. There is firstly the evaluations of an emotional experience made along the five basic evaluative dimensions (BEDs) set out in the first chapter. The concept of disordered/normal emotion is a meta-evaluative one consisting in negotiating the appraisal of the emotional experience along the five BEDs, as we've seen in the third chapter. But this meta-evaluation itself turns out to be influenced by a yet higher-level evaluation that pertains to the global assessments we make of our lives, i.e. assessments we make in terms of the notion of flourishing. For brevity, I shall refer to the concept of flourishing as an *apical* concept to indicate that it is an organizing principle that shapes the evaluations we make of our emotional experiences at the first two levels (i.e. negotiating the BEDs and applying the meta-evaluative judgement of disorder on this basis).

4.3 Thinking about Flourishing: An Altogether Different Starting Point

There was once a very different starting point for thinking about our lives as a whole, anchored in an altogether distinct question, namely this: 'How ought I to live?' The first-person formulation of this question is of vital importance, for the question's value was taken to lie in each person's exercising their deliberative agency in reflecting on his or her own life by asking this question. When Socrates asserted that the unexamined life is not worth living, no one then took him to mean that someone else—researchers in psychology or psychiatry, say—could do the examining for us, and deliver a set of prescriptions for us to abide by, in the hope that we might live as we ought to. The point of departure into a life of flourishing was taken to be each individual's exercise of their own reflective and deliberative agency. The act of reflection itself was read as decisive.[8] This way of approaching the notion of flourishing is one we've inherited from classical Greek antiquity.

[8] This is not to rule out the need for, and relevance of, empirical research into factors that conduce to our well-being. We could profit from knowledge of general patterns and tendencies, but the question is how this knowledge is related to our own exercise of deliberative agency—our participation in the quest to understand how we ought to live.

This starting point shapes everything else that follows—all the answers offered within this tradition of antiquity were understood as contributing to a dialogue the participants of whom are assumed to be self-reflecting agents actively grappling with this question, 'How should I live?' And to comprehend them, it is in this spirit that we have to approach any of the answers given by the various schools of ancient Greek thought. We have to be wondering for ourselves how we are to spend our lives, how we are to decide on whom or what to commit ourselves to, and on which projects we should invest ourselves in (and the manner by which we should conduct ourselves in thus engaging), how we are to value the opportunities and obstacles that we encounter along the way. Overlooking our own role as self-reflecting deliberative agents—missing the point that these answers emerge from a dialogue in which we are ourselves grappling with the question of how to live—would be to risk fundamentally misunderstanding the claims made about the nature of human flourishing by the Stoics, the Epicureans, the Sceptics, and of course the Aristotelians.

But this all-encompassing starting point is all too easily overlooked from our contemporary vantage-point. We have become used to studying, thinking, and talking about human flourishing in an oddly impersonal way—driven in part by the quest to study flourishing in a scientific manner (where this in turn is taken to require the inquirer to adopt a detached, observational standpoint), and driven in part by the health-based conception of flourishing sketched out in the previous section, which recall, takes as central the question, 'How can we reduce human suffering and increase happiness?' Nothing is lost by reformulating this question impersonally, i.e. 'How can human suffering be reduced and happiness increased?' It is perfectly possible for someone to read a list of prescriptions that suggest ways of reducing suffering and increasing happiness and to then try to follow these prescriptions regardless of whether or not she takes herself to be doing this in the context of reflecting on her life as a whole.[9] With this mindset in place, it becomes easy to misread a Stoic or an Aristotelian response to the question, 'How ought I to live?', as offering us just such a list of prescriptions—and worse still, to misread these responses as aiming at offering suggestions about how to increase our happiness-as-good-feeling.

[9] The principal contrast here is between two kinds of stance that we might adopt towards findings in the empirical study of human happiness: an observational-interventionist stance vs deliberative-reflective stance. Someone could observe and then seek to adhere to the prescriptions made by positive psychologists, and therein be exercising her agency in an interventionist sense, without at all adopting a deliberative-reflective stance towards her own flourishing—that is, without seeking to understand what doing so would mean for her as she takes sum of her life as a whole. Thus she might, for instance, presume that to flourish is to enjoy a preponderance of good feeling, without realizing that she does indeed presume this—and thus without questioning whether this particular conception of flourishing is worthy of her allegiance.

To guard against this misreading, it might help to notice that in stark contrast to the question, 'How can human happiness (read: good feeling) be increased?', the question, 'How should I live?' does not have written into it the presupposition that human flourishing is to be understood in terms of any one particular substantive value (be that love, wealth, fame, health, or happiness-as-good-feeling). It is also worth noticing that the responses the ancients give may not take the form of proposing a substantive value either (e.g. that I should live so as to increase my material wealth, or that I should live so as to love and be loved). I explore this idea more fully in the next section.

Having noted our susceptibility to this misreading, I will now turn to an Aristotelian answer to the question, 'How should I live?' Aristotle is of special interest to us because he held that our emotional experiences play a decisive role in determining whether we flourish or languish, but his understanding of how our emotions do this is strikingly different from the way our emotions are taken to contribute to our flourishing according to the HF. So a comparison of the two conceptions of flourishing promises to yield interesting and important insights that bear upon our present inquiry into the concept of emotional disorder. More specifically, the point of turning now to consider an Aristotelian conception of flourishing is this: I suggested earlier that we make sense of, and evaluate, our emotional experiences at three levels. The third of these is where we find the apical concept of flourishing, which informs the way we negotiate the basic evaluative dimensions and make the meta-evaluative judgement of emotional disorder. An Aristotelian conception of flourishing provides us with a contrast to the HF, thereby giving us a vivid sense of how very differently our emotional experiences would be construed at the lower two levels of evaluation when informed by this alternative conception of flourishing.[10] A preliminary way of appreciating this is by noticing that the argument sketched earlier in favour of prioritizing the BED of valence in a depressive$_{\text{DRIC}}$ experience would be undercut if we were not to presuppose the health-based conception of flourishing. That is, if the end-goal in responding to depressive$_{\text{DRIC}}$ experience is not assumed to be the reduction of emotional distress and/or increasing the depressive$_{\text{DRIC}}$ subject's happiness (happiness-as-good-feeling), then perhaps it will turn out to be the case that it does matter decisively whether we prioritize the BED of valence over the BED of existential significance, or vice versa. Indeed, an Aristotelian conception of

[10] Because Aristotle doesn't start out in the way that Maslow and Seligman do when pondering the nature of human flourishing—which is by assuming that flourishing is the opposite of mental illness—it is not obvious how prominent, if at all, the second level will be in the Aristotelian picture. The concept of emotional disorder could very well be much less *interesting* to Aristotle than it is to us.

flourishing shows us one way in which it precisely does matter, and decisively so, how we weigh up the contrary appraisals made of a depressive$_{\text{DRIC}}$ experience in the BEDs of valence and existential significance respectively. I turn now to explore this claim further. I shall refer to the Aristotelian conception of flourishing that I present below as 'the excellence-based conception of flourishing', or 'EF' for brevity.

There are many ways one can embark on the task of characterizing the Aristotelian conception of flourishing. The way I have chosen to do so in the next section is informed by the susceptibility to misreading Aristotle in the way I have just noted.

4.4 The Excellence-Based Conception of Flourishing (EF)

Aristotle situates the question, 'How ought I to live?', within the context of a particular schema for thinking about flourishing. According to this schema, all living things can be said to flourish or to fail to flourish, and the criteria for determining whether an individual flourishes is set by the species to which it belongs. So a willow tree, a swallow, a cat-fish, and a tiger can all be said to flourish and to thrive—by which Aristotle means they can all be excellent specimens of their species. This conception of flourishing stems from the notion that for all things that have a function (which all living things do, according to Aristotle), the good and the 'well' is thought to reside in the function specific to any particular individual, as set by its species membership (*EN* 1097b26).

Aristotle locates the function of humans in the form of life that he takes to be distinctive of being human (and it is in excelling in this function that we can be said to flourish). This form of life, says Aristotle, is 'an active life of the element [in the human organism] that has reason' (*EN* 1098a2). In other words, while the human organism is composed in part of nutritive, perceptive, and appetitive functions (sharing the first with plants and animals, and the latter two with animals), we uniquely possess reason and exercise thought. Crucially, the possession of reason enables us to engage our perceptive and appetitive functions in a manner that is distinctive (different, that is, from the way non-rational animals do): we can engage them in a way that is informed by, and which in turn informs, our capacity for reason (*EN* 1097b35–1098a5). This mutual interdependence is useful to note at the outset in a study of the role of emotions in human flourishing because Aristotle takes our emotional states to emerge from, and to unfold through, a complex interplay of our exercise of all three functions (perceptive, nutritive, rational).

Asking of ourselves, 'How ought I to live?', necessarily involves an exercise of our rational capacities—perhaps it is *the* exercise of our rational capacities *par excellence* because it manifests our capacity to understand that our lives as such can be valued, making it apt to normatively assess the way we spend our lives. We begin to see that it is possible to steer our lives in one direction rather than another, to decide in one way rather than another. So in asking this question, in turning a reflective gaze onto our own lives, we begin to approach our life as a distinctively human life, and to inhabit it in a distinctively human way. To manifest our capacity for reason in this self-reflective manner sets us on the path towards flourishing as human beings, according to the Aristotelian conception, because (recall) it is in exercising well the function that makes us distinctively human—namely, reason—that we flourish.

Asking this self-reflective question is a necessary though not sufficient condition for exercising our reason well. How we grapple with the question, and the answers we give in response to it throughout our lives (not just intellectually but in our affective and volitional dispositions), are just as important. Here, a point that is central to Aristotle's conception of flourishing (but which appears obliquely, if at all, in the HF) can be introduced. Once we begin to grapple with the question of how we ought to live, however inchoately, we start to consider what it would be *good* for us to do and how it would be *good* for us to live — and to desire to live accordingly. We come to value things and to evaluate them. A basic fault line appears in our psychology between two very different kinds of motivational state: we are motivated by states of our appetitive function, and we can also be motivated by what we rationally conceive to be good for us, a conception that emerges from our grappling with this question about how to live.[11] This furnishes humans with two distinct kinds of object of desire: the pleasant or the pleasurable (*epithumia*), on the one hand, and the good (*boulesis*), on the other. Describing the emergence of the desire for the good as signalling a distinctly human form of life, Alisdair Macintyre states:

> The key moment in distinctively human development occurs when someone first makes use of their linguistic powers to pose the question 'What is the good of doing this or that, of making this or that happen or allowing this or that to happen?' and is understood as inviting from others or from himself some statement of reasons for and against any particular answer, reasons which can then be evaluated. From then on human projects, human responses to good and bad fortune, and human

[11] There is a third kind of motivational state that Aristotle identifies (*thumos*) which I will not touch upon here.

> relationships were taken to be intelligible in terms of the good and the bad, the rationally justifiable and the rationally unjustifiable. (Macintyre 2016: 225)

On the Aristotelian picture, it is by successfully steering our lives in accordance with the good that we excel in our function as rational animals, and it is in excelling in this way that we can be said to flourish.[12] Thus to excel is, suggests Aristotle, our final or ultimate end—the end for the sake of which we choose everything.[13] This gives us the start of a reply to the defence of pathologization sketched out at the end of the last chapter: here is a way of construing our final end that is not spelt out in terms of the notion of happiness-as-good-feeling or of enjoying a pleasant life. It will take us the rest of this chapter to grasp more completely this alternative construal of our final end, and so I will return to this reply in the concluding section of the chapter.

Whilst we are able to desire—to be moved to feel and to act—in these two ways thanks to our capacity for reason, we may not initially be very skilful at identifying what is good for us. So it is possible that something can appear to us to be good, even though it may turn out actually not to be so. It is through the cultivation of virtues, both virtues of character and virtues of intellect, that we become more adept at identifying what is genuinely good for us, desiring it for this reason (i.e. that it is good for us), and acting accordingly. In doing all this well, we would in effect excel in the exercise of our epistemic, moral, and deliberative agency. In thus excelling, we would be instantiating the distinctively human form of life in an excellent manner, and therein we flourish.

4.4.1 The Role of Nature in Aristotle's Account of Flourishing

Although the question of how we ought to live is approached by Aristotle from the perspective of thinking about human life in a species-based manner, there is one reading of his response to this question that I wish to put aside, namely that Aristotle is offering us a substantively prescriptive answer. What Aristotle does not appear to do (at least, not until the tenth and final book of the *Nicomachean Ethics*) is to suggest that there is one substantively specified way of life—a life guided by a specific value or one consisting in engaging in a particular kind of activity—that humans must try to adopt if they are to flourish.[14] Indeed, as various commentators have pointed out, in living our

[12] It mustn't be forgotten that at the heart of this steering is our exercise of the capacity to understand and to appreciate the force of the question, 'How ought I to live?'

[13] *EN* 1097b5–7.

[14] Not everyone agrees, however. See Richard Kraut (1991) for a dissenting view.

lives as human beings, we are presented with a wide range of goods—not all of which are commensurable to each other—and central to the Aristotelian picture of human flourishing is a self-reflective agent who exercises her reason well as she deliberates and chooses between goods, at all levels of scale in life. (Nussbaum 1985 and 2008; Macintyre 2016).

What Aristotle's appeal to the idea of 'human nature' does, rather, is the following two things (and here I follow Julia Annas). First, in approaching the question of human flourishing from this angle, Aristotle can be understood as seeking to offer us an account of flourishing that is liveable and humanly achievable, i.e. it is flourishing of a kind that we can hope to realize in a human life, with all its characteristic needs, desires, vulnerabilities, and so on (cf. Annas 1993: 440). Second, and particularly instructive to our present inquiry, is this: while it seeks to be liveable, Aristotle's account of flourishing also sets out an ideal that we can use to help identify, scrutinize, revise, and possibly discard the merely conventional elements in our evaluative beliefs and practices, including our beliefs about what it means to flourish (ibid.: 440). One such conventional idea is that flourishing consists in enjoying subjective well-being, where this is fleshed out in terms of the idea of experiencing a prevalence of positively valenced emotional experiences (as embodied in the health-based conception of flourishing sketched out in Section 4.2 above). In contrast to this, Aristotle's thinking about human nature results in his placing at the centre of his account of human flourishing our capacity for reason, and in so doing, Aristotle can be understood as articulating an ideal that humans would do well to live up to. That Aristotle's view of flourishing can be understood in these terms may be appreciated by considering the many ways we may fail to realize this ideal. One basic way of failing to exercise our reason well, or to barely exercise reason at all, would be to be insufficiently curious or concerned about how one ought to live, and thus to fail to form on this basis a serious conception of the good that one seeks to realize in one's life. From the Aristotelian perspective, such a life, *qua* human life, would be a defective one—and it would be so precisely because we hadn't fully appreciated all that our lives afford us, i.e. in particular that, as the life of a rational animal, it is capable of being valued by the subject whose life it is and that, recognizing this, the subject becomes concerned about how to spend it well, how to make the best use of it.

And yet this defective life is one that is all too easily lived given currently conventional ideas about flourishing. The focus in the HF is almost entirely on satisfying our desire for the pleasant (richly construed as this can be). On this picture, there is no recognition of the good as a distinct, separate object

of human desire. Insofar as any reference is made to the idea of the good, it is ultimately cashed out in terms of subjective well-being, i.e. the pleasant. This occurs because we have failed to recognize that in addition to desiring the pleasant because we have nutritive and appetitive drives, we are also rational beings—beings for whom our nutritive, perceptive, and appetitive nature is amenable to being structured and organized by our rational capacity to desire the good and thus to value and to evaluate. Seen from this angle, Aristotle's appeal to nature can be understood as an attempt on his part not to miss out anything of significance to human flourishing, that is, to have the question, 'How ought I to live?', posed in a manner that is informed by a clear-sighted recognition of all that we—as humans—are capable of being (and of becoming).[15]

The roles that emotions play in determining human flourishing, as conceived of according to the EF, has to be understood in light of this basic distinction between desiring the pleasant and desiring the good. The reason for this is simple: once the good is identified as an irreducible object of human desire, it becomes possible to countenance the possibility that emotions may contribute to our flourishing in ways other than by making life feel pleasant. This last remark shouldn't be read in such a way as to suggest that I am attributing to Aristotle a desire-satisfaction theory of flourishing. On the Aristotelian view, it is our desiring of the good itself (when we successfully desire what is genuinely good), and not merely the satisfaction of this desire, that is taken to be a constitutive part of our flourishing. For in desiring the good, we exercise our capacity for reason in a way that enables us to direct our own lives, at least to some degree. We begin to be able to steer our lives in accordance with what we value, and this sets us on the path of exercising our epistemic, moral, and deliberative agency in a manner that is apt for normative appraisal. In doing this well, a person excels at being human; she lives an excellent human life—and it is on account of this that she can be said to flourish.

4.5 Emotions and Flourishing According to the Excellence-Based Conception

Given this setup, we can anticipate that it will be through the exercise of our epistemic, moral, and deliberative agency (abbreviating this, I'll speak of our

[15] This is in fact reflected in his own reasoning on this point, which proceeds by eliminating relevant alternatives. That he is attempting to articulate an ideal is made clear in passages such as *EN* 11176b32 in which he dismisses the suggestion that amusement could constitute our ultimate or final end on the grounds that 'to exert oneself and work for the sake of amusement seems silly and utterly childish.'

rational agency) that our emotional experiences contribute to our well-being. In order to understand how the Aristotelian thinks our emotional experiences do this, then, we will need to understand how, for the Aristotelian, our emotional experiences figure in the exercise of our rational agency. The picture is a complex and intricate one because on the Aristotelian view, our embodiment—with all the nutritive, perceptive, and appetitive functions that this brings—forms the container within which we exercise our capacity for rationality. We are rational creatures, to be sure, but we are rational in a thoroughly embodied way (so we might call the form of rationality we enjoy *embodied rationality*). This means that the perceptive and appetitive dimensions of our embodiment continually interact with our reason throughout our lifespans—and this, for Aristotle, is simply the truth about being human (it is neither condemned nor glorified). The idea that we should seek through the use of reason to transcend or outgrow our appetitive states would be very alien to Aristotle. While it is true that he thinks that brute appetitive states, left uncultivated, can subvert our ability to exercise our reason well, what is more noteworthy is that Aristotle thinks that (a) our appetitive states can be cultivated through habituation and our use of reason, and that (b) a full and excellent use of reason essentially involves our cultivated feelings and emotions. So as well as not having our reason dominated brutishly by uncultivated emotions, living well as a human being also means being properly affected by our circumstances. As one author puts it, 'Having non-rational feelings, such as anger, grief, fear, etc., appropriate to the circumstances we find ourselves in is part of our fullest perfection,' on the Aristotelian view (Cooper 1988: 36). Indeed, Aristotle holds that having appropriate emotional dispositions is what provides us with the right starting point for exercising our practical agency; our emotional dispositions are what enable us to step into a life in which we are able to exercise our reason well (*EN* 1144a6–37).

Encapsulated in this view are at least four distinct ways in which our emotional experiences figure in the exercise of our epistemic, moral, and deliberative agency, and therein four ways that they contribute to our flourishing. I will turn now to consider each of these ways in turn. My discussion here will be selective as each of these four roles has been dealt with extensively by various authors elsewhere, and as my purpose here is not to introduce these ideas as such but rather to present them in juxtaposition to the currently dominant way of thinking about the relationship between emotions and human flourishing (as summarized in Section 4.2 above). Given this aim, the sketch that follows may look slightly unfamiliar because the foregrounding of emotional dispositions entails that other important notions that usually occupy centre

stage in discussions of Aristotelian flourishing—in particular, that of practical reason and the excellences of the intellect—will appear less prominently in the discussion.

Before I proceed, a quick point of clarification: while some of examples to be discussed below are straightforwardly identifiable as emotions (e.g. guilt), others are traits that contain an affective component (e.g. valuing truthfulness). My reason for including the latter in a discussion of the way in which emotions contribute decisively to our flourishing (in the Aristotelian sense) is that the concept of emotion as it appears in this claim must be construed broadly to include our affective lives as such, as this is how Aristotle himself frames the claim. For the claim is at the broadest level rooted in Aristotle's view that the human psyche consists in a rational element and an irrational element, where the latter is said to consist in two further parts: 'the vegetative element [which] in no way shares in reason', and 'the appetitive and in general the desiring element', which is said to share in reason 'in so far as it listens to and obeys it' (*EN* 1102a34–1103a4). It is in virtue of its sharing in reason that 'the appetitive and in general the desiring element' contributes decisively to human flourishing (recalling that to flourish is to exercise our rational agency excellently). So it is this element, and not emotions narrowly, that is said to contribute decisively to human flourishing, on the Aristotelian view.

4.5.1 Orientation and Practical Deliberation: The Role of Emotions in Our Forming and Sustaining *Boulesis* ('Rational Wishes')

> Everyone capable of living by their own decision ought to lay down some aim for living finely, be it honour or reputation or wealth or education, which they will look to in the performance of all their actions, since not organizing one's life in relation to some goal is a mark of great foolishness.
>
> **(Aristotle, *Eudemian Ethics* 1214b6–10)**[16]

Our emotional dispositions enable us to pursue the good in a wholehearted and committed manner—and this is perhaps the most fundamental way in which our emotional dispositions figure in the excellent exercise of our

[16] Cottingham's translation: 'Not to have your life planned towards some end is a sign of great folly' (Cottingham 2010: 8).

epistemic, moral, and deliberative agency. To flourish in this sense, it is not sufficient that we are able to conceive of what is good for us, nor even to desire in some general way the realization of the good (thus identified). It is having the relevant particular emotional dispositions that practically orients our entire organism towards the realization of the good in our lives; these dispositions are what embed our conception of the good firmly into our lives. So our emotional dispositions have an orienting role in our flourishing: to the degree that we are able to realize the good in our lives through the exercise of our rational agency, our emotional dispositions are integral to our being disposed to do this (i.e. to realizing the good). An agent's life being orientated towards the good finds expression most directly and comprehensively through his emotional experiences.

The categorical difference our emotional dispositions makes here can be seen in the distinction Aristotle draws between the continent or strong-willed person and the fine (excellent) person. The continent person conceives of the good (i.e. what is truly and genuinely good) and even desires its realization in some general way. But he has contrary appetites—emotions or impulses, say—that make the realization of the good something he can only achieve through effort. The excellent person's realization of the good is by contrast seamless because he wholeheartedly (this includes *affectively*) embraces the good. This holds true from the most localized decision-making ('Go out for my daily exercise or sit on the sofa to watch another hour of television?') right through to the most global conception of the good in terms of which we organize our lives, i.e. our conceptions of what it would mean for us to live well.

It is because he recognizes the centrality of our affective dispositions in enabling our conception of the good to take root in our lives that Aristotle situates our desire for the good within the appetitive faculty of desire, in spite of the fact that only creatures endowed with reason are capable of forming this kind of desire (as discussed in Section 4.4 above). So our desire for the good is described, seemingly paradoxically, as a rational desire or wish (*boulesis*).[17] However, this is not so paradoxical when we view our appetitive functions as being permeable to reason. Our appetitive dispositions, when cultivated and sculpted by reason, are altogether transformed. In the good or fine person, reason-informed appetitive dispositions are not at all like appetitive dispositions that have not been thus cultivated (or where cultivation has been partial). In a fine person, the resulting state of being signifies a genuine integration of her appetitive and rational functions, hence it becomes possible to

[17] *De anima* 3.10.433a21–6.

ascribe to such a person the wholehearted rational wish (now *prohairesis*) to realize the good.[18]

4.5.1.1 Thinking of Our Lives as a Whole

Just as our aiming at (our conception of) the good ramifies throughout our lives at all levels of scale, so there are a multitude of ways in which our desire to live well—our desire to realize the good in our lives (as we conceive of it)—manifests in our lives. At one extreme, our moral and social sensibilities shape the way we conduct ourselves in the minutiae of daily living, often unconsciously and without much reflective thought; at the other extreme, we are from time to time pulled up on ourselves, circumstances inviting or compelling us to deliberatively reflect on our lives as a whole—to take stock and to ask the question in earnest, 'How shall I live?' (These various ways in which we are guided by our conception of the good do of course mutually inform and shape each other: we come to learn something of ourselves through our actions, as much as our conception of the good shapes the way we act).

Not everyone may as a matter of fact dwell on this question or take it very seriously, but it is mark of a person's exercising their rational agency well that she does this. In reflecting on the question, 'How ought I to live?', the fine or excellent person recognizes that she is guided by a final or an ultimate end.[19] This end embodies her conception of the good at the most general level, i.e. it frames her understanding of her life as a whole. This end has several features that make it unique and distinctive, so that it structures her life in a quite different way from the way that our various activities and actions are goal-directed. The most important characteristic of this end is that it is not pursued and achieved by means of engaging in any one particular activity (as studying for an academic degree or baking a cake for a child's birthday party is). Rather, it is realized by means of, and through, a person's engaging in all the activities, and more generally in living, as she does. The good that is our final end encompasses our lives as a whole, and all the various things that we do are partially constitutive of this all-encompassing end (Macintyre 2016: 229; cf. Urmson 1988: 11). So the excellent or fine person's final end—expressed as her conception of the good at the most general level, a conception that emerges from her grappling with the question of how she ought to live—is perhaps better understood as an organizing principle, rather than as a goal or aim understood in its traditional sense (as that which a particular action or activity is directed).[20] Successfully realizing this good necessarily involves the excellent

[18] On the use of the term *prohairesis*, see Macintyre 2016: 38.
[19] The *summum bonum* (the highest good or the final end).
[20] For a fuller discussion of the Aristotelian notion of the final end, see Macintyre 2016: 52–3.

exercise of the fine person's epistemic, moral, and deliberative agency—and in exercising one's agency in this way, one therein successfully realizes one's final end. It is for this reason that the claims that (a) flourishing consists in realizing one's ultimate end, and (b) flourishing consists in the excellent exercise of one's rational agency, come to the same thing. The point to note here is that in realizing one's ultimate end, one flourishes not because it feels good to realize this end (although there is an interesting sense in which the flourishing agent does take heart in realizing her ultimate end (more on this in the sketch of Vasily Grossman's life below)); rather, one flourishes because, in realizing this end, one exercises one's rational agency excellently.

Once we appreciate that the final end is more of an organizing principle than a goal (in the usual sense of this word), we are in a position also to anticipate that the relationship in which we stand to this final end cannot be like the relationship that holds between our particular actions and goals. Our final end comes into focus most sharply and consciously when we are asked to give an account of ourselves to others or to ourselves in the course of our pondering the question, 'How ought I to live?' It is in holding ourselves accountable—to others and to ourselves—that we most consciously understand ourselves as oriented towards the ultimate good, the *summum bonum*, of our lives. As has already been noted, our affective proclivities and dispositions have everything to do with this orientation: in discussing the foundational Aristotelian idea that we rationally desire the good (*boulesis*/*prohairesis*), it was observed that our emotional dispositions are the means through which we are thus oriented (successfully, if our emotional dispositions are well cultivated; less successfully, otherwise, as we see in the strong-willed, in the weak-willed, and in the bad respectively). But it is possible also for an agent's final end to be defined partly in terms of an affective disposition (or a range thereof). Being a kind and loving parent, being a truthful journalist, being a just political leader, or being a courageous defender of a local natural ecosystem against the encroaching threat of industrialization are all pursuits it is possible to commit one's life seriously to. But the affective disposition in each of these cases—being loving, being courageous—is usually also a deep-rooted value in the person's life (or grows to become so), which is how a person is able to commit himself wholeheartedly to the pursuit in question and to find it a worthwhile one. The agent's commitment to the pursuit grows out of this (well-cultivated) affective disposition. He manifests certain emotional tendencies—in the context of spending his energy and effort, his focus, his time, and other personal resources—in nurturing his children lovingly, or serving his political constituents justly, because being loving or being just matters to this person in an

essential way: these affective dispositions embody the person's conception of the ultimate good.

With this account in place of how emotions contribute to our flourishing, the possibility becomes intelligible to us that a person could flourish in spite of living in challenging circumstances—circumstances that offer diminished opportunities for pleasant affective experiences—in virtue of a clear conception of the *summum bonum* crystallizing for this person, a final end towards which he is wholeheartedly oriented and exercises by this means his rational agency. We are also able to grasp the possibility that a person could flourish (or come closer to flourishing) *in virtue of* experiencing difficult, even deeply distressing, emotions if it turns out that being thus orientated consists (in part) in experiencing these emotions. Although Aristotle himself held that the best life is one that isn't beset by overwhelming hardship, at the core of this view of flourishing is the human agent who exercises her capacity for reason fully and excellently, and integral to the fine exercise of one's capacity for reason is having our lives oriented towards (our conception of) the good. Being thus oriented is expressed most directly and comprehensively through our emotional experiences, which may very well include experiences of emotional distress.

To animate this conception of flourishing, I will now briefly sketch out an account of two lives. The first life is that of Vasily Grossman, a Soviet writer who initially worked as a war correspondent for the Red Army in WWII and later wrote a two-volume novel set during the Battle of Stalingrad. The second is the life of Winthrop Cohen, a pseudonym given by the anthropologist and psychiatrist Arthur Kleinman to a 'patient' who was referred to Kleinman by his family because he appeared to be suffering the symptoms of major depressive disorder.

Vasily Grossman

Vasily Grossman lived and wrote in one of the most dangerous periods of Soviet history, i.e. he lived under the Stalinist regime when millions were sent to the gulags and during the Second World War when the Nazis—besides being embroiled with the Red Army in a battle to the death in a bid to conquer Russian territory—were engaged in a genocidal campaign against the Ukrainian and Belarussian Jewish population. Grossman was embedded with the Red Army as a war correspondent in the Battle of Stalingrad, and so witnessed the savagery of war first-hand, but it was the personal tragedy of losing his mother to the genocidal horror that took place in Ukraine in September 1942 that ultimately shaped his path as a writer in a thoroughgoing and decisive way.

Ever since he first began writing in the 1920s, Grossman manifested a deeply humane perspective on life and on human suffering. However, his ability to express this perspective lucidly in his writings had been hitherto compromised to greater or lesser extents in his effort to negotiate the dangerous Stalinist state machinery in order to get his works published and, ultimately, in order to save his life and that of his family. The Holocaust, and the personal tragedy he suffered in this, changed all this. It crystallized Grossman's resolve to write truthfully, both in his journalistic work and in the historical novel set in the Battle of Stalingrad,[21] which he began writing after the war and which took him seven years to complete. Alisdair Macintyre has described Grossman's life and the process by which his commitment to truthful self-questioning gradually came to guide Grossman in a steadfast and unequivocal manner (Macintyre 2016: 244–64). I am deeply sympathetic to Macintyre's portrayal of Grossman's character and, rather than rehearsing what Macintyre has already said, my concern in what follows will be specifically to underscore the role that Grossman's affective dispositions played in shaping his understanding of truthfulness and in orienting him towards it.

To begin, we need to know something about what Grossman understood by truthfulness. Grossman lived in a time when megalomanic totalitarian ideologies swept through large areas of the European continent, turning millions of people into mere numbers. Truthfulness, for a humanist like Grossman living in such destructive times, meant above all affirming the value of each individual human life and of the human bond of love that sustains and nourishes us all (the bond between mother and child figures especially prominently in Grossman's outlook).[22] Here, it is germane to note how Grossman's conception of the good is shaped by his affective dispositions. Someone dissociated and cut off from loving relationships, someone who wasn't moved by human kindness and generosity of spirit, couldn't be truthful in the way Grossman was. The eloquent and powerful way by which Grossman was able to convey the value of the maternal bond in human life was no doubt borne out of the grief he suffered from losing his own mother in the Holocaust. Grossman's understanding of what it meant to be truthful—of what it is he felt he had to be truthful about—grew out of his own deepest emotional experiences and his affective dispositions. I will return to this point (i.e. the way our affective

[21] This novel, considered to be Grossman's *magnum opus*, is published in two volumes in English as 'Stalingrad' (1952) and 'Life and Fate' (1980).

[22] This theme recurs in Grossman's writings, but see especially his non-fictional essay, 'The Sistine Madonna' (2011 (originally published 1952)).

dispositions themselves can inform our conception of the good) in Section 4.5.3 below.

As well as shaping his conception of the good, Grossman's affective dispositions also orientated him towards his final good. Being truthful in the manner just specified is no mean feat *intellectually* in the face of ideologies that seduce and entice people into passionately believing that individual lives may be sacrificed, and are worth sacrificing, for the greater social good, i.e. in order to establish 'universal harmony' fashioned in the mold of Communism or of National Socialism. But truthfully affirming the value of individual human life and of the human bond of love in one's writing—and thereby to critically question these totalitarian visions—was also *practically* perilous given the violent imposition by these regimes on society of their respective agendas, and the brutal intolerance they displayed towards any dissenting voice. Grossman's commitment to his *summum bonum* has to be understood in this context, because it is only then that are we able to appreciate that his commitment meant swimming ferociously against the prevailing current: it would have been much easier to become indifferent to truthfulness in these circumstances. Clearly, the circumstances meant that there was only really one way of pursuing truthfulness, i.e. bravely and resolutely. And it is in the courage that Grossman displayed that we see the depth of his commitment to truthful self-questioning. His courage underscores this depth: it is an orientation towards truthfulness that is not just intellectual, but crucially, also affective and desiderative. It is hard to picture how Grossman's pursuit of his *summum bonum* could be anything but fully embodied, finding expression not just in his beliefs and judgements but also through his affective dispositions. Had it been otherwise, then given the immensity of the countervailing external pressures he faced, Grossman's pursuit of the good will most likely have appeared more fragmented, e.g. his practical reasoning more wavering, his writing more faltering. It is because of his depth of commitment that we see in Grossman's realization of his final good a life that is *eudaimon*—that is, a life that instantiates, or that comes close to instantiating, perfection of the human form, for in this commitment Grossman engages and employs well *all* the facets of his reason, i.e. cognitive and appetitive.

Grossman's life highlights the pivotal role that our affects play in orienting us towards the good (as we conceive of this) because of the extremity of the socio-political circumstances in which he lived. But what is very evident in Grossman is discernible everywhere in people whose lives are similarly oriented towards their conception of the good. Still, not everyone comes to form a final end in the way that Grossman did, and if we exercise our rational and affective capacities through our daily activities without being guided by any

conception of the ultimate good in our lives, what we will not be doing is exercising these capacities in that distinctively self-reflective way that manifests an understanding that our lives as such can be valued. Valuing it in this way is exactly what Grossman did, and it is in his eventual wholehearted pursuit of his ultimate good that we see in Grossman a person who flourishes, i.e. someone who exercises well his epistemic, moral, and deliberative agency. Macintyre observes that Grossman died a 'dissatisfied and unhappy man,'[23] but continues: 'It was the pursuit of those ends [the ends mandated by his task of writing truthfully] that gave finality to his life, so making it unhappy in the modern sense, but in fact *eudaimon*' (Macintyre 2016: 264).

Now, it might seem tempting, in light of the claim that Grossman flourished in spite of dying a 'dissatisfied and unhappy man', to cast the Aristotelian view as an 'objective' conception of flourishing—construing it as setting out a set of objective conditions the satisfaction of which entails that a person flourishes, regardless of the way he feels in and about his own life—and to juxtapose this against the health-based conception of flourishing, construed as a 'subjective' conception according to which how a person feels determines whether or not he flourishes. But this would be to miss the point that it mattered decisively to Grossman himself that he spent his life pursuing his final end wholeheartedly, and that he took this pursuit to be intrinsically worthwhile, i.e. worthwhile regardless of whether or not it resulted in worldly success (that is, in his writings being published).[24] This self-reflective valuing of his own life, his affirmation of the way he spent it, is integral to the Aristotelian claim that Grossman exercised his rational agency well. Here, again, Aristotle recognizes the importance of the affective dimension: he holds that executing virtuous actions is inherently pleasant for the agent himself (*EN* 1153a13–15). The notion of the pleasant being invoked here is not the brutely appetitive one that was earlier contrasted with the good, i.e. in the claim that humans desire both the pleasant and the good. Rather, it is a notion of the pleasant that is *internal to* the agent's conception of the good (cf. Annas 1980: 293). It is pleasure in the sense of being satisfied in knowing that we've done the right thing, however

[23] Grossman's pursuit of truthfulness came at a huge personal cost: his final novel was viewed as subversive by the Soviet regime and all known drafts of his final novel were confiscated by the KGB in 1962. He pleaded with Khrushchev himself to have it returned and to allow it to be published (to no avail). Grossman was to die two years later of kidney cancer, aged 58, without seeing his novel again.

[24] Although Grossman yearned for his novel to be published, had he known that it wouldn't be published in his lifetime, I doubt this would have stopped him from continuing to write. Doing so would have been a betrayal of all that he stood for, and Grossman's orientation to his final end is manifest perhaps most lucidly in the fact that he simply wouldn't have been Vasily Grossman had he been capable of betraying himself in this way (i.e. we would have to imagine so much about his character changing that we would begin to lose our grip on the singular individual who was the historical Grossman).

awkward or difficult the consequences of this may be for us. Once Grossman wholeheartedly embraced his calling and pursued it in earnest, the task of writing truthfully became easier in a specific sense: there was less to wrestle with, less to hesitate about. Or to turn the picture on its head, to not have embraced this calling wholeheartedly, once Grossman had felt its tug as forcefully as he had, would likely have led to a sense of frustration and ultimately of regret about a life half lived, a life not coming into full fruition. It might very well have been a life permeated by a long-standing depressive$_{\text{DRIC}}$ experience. Grossman's fidelity to his conception of the final or ultimate good brings with it its own wellspring of pleasure, i.e. pleasure of the sort that derives from an agent's knowing that he is living—and has lived—with integrity.[25]

Summarizing, Grossman is *eudaimon* because his conception of the good became starkly clear to him, and he pursued the good thus conceived in an intelligent and wholehearted way. As he matured as a writer and as a human being, Grossman's appetitive and rational functions grew to be integrated in such a way that his rational wish to write truthfully came to be firmly embedded in his entire being—his affective attunement to his final end guiding and supporting his practical reasoning and his actions—and in this way, Grossman eventually manifests an excellent exercise of his rational agency. Situating Grossman's life within the Aristotelian framework enables us to see that it is possible for a human being to flourish in spite of living in challenging circumstances that offer diminished opportunities for enjoying (brutely appetitive) pleasant affective experiences, even suffering greatly as a result of being confronted by such circumstances. Indeed, Grossman's life enables us to grasp the possibility that a person could flourish (or come closer to flourishing) *in virtue of* experiencing difficult emotions that play a constitutive role in shaping his or her conception of the good.

Winthrop Cohen

'Winthrop Cohen' is a pseudonym given by the psychiatrist and anthropologist Arthur Kleinman to a man who was referred to him (Kleinman) by Cohen's concerned wife and adult daughter. Cohen's life, as described by Kleinman, illustrates the orientating role that our affects play in a slightly different way to Grossman's. Kleinman's story begins with Cohen joining the US army as an 18-year-old and, not long afterwards, being deployed to the Pacific theatre in

[25] It is also worth adding here that because flourishing in the Aristotelian sense consists in the excellent exercise of one's epistemic agency, to exercise one's agency well must include knowing that one is doing so. In this way, too, the agent's own perspective on his own life is encompassed within the Aristotelian conception of flourishing. My thanks to my colleague Brian Rosebury for underscoring this point.

the Second World War. After being demobilized at the end of the war in 1945, Cohen returned to the United States and went to law school, graduating and then working successfully as a lawyer for forty years. Outwardly, all seemed well in this period: he married, had a child, and became a senior partner in a large law firm. But on his sixtieth birthday, Cohen began to be tormented by guilt for a particular misdeed he had committed in the war forty years earlier: he had deliberately fatally shot an unarmed military doctor on the enemy side. The doctor was in the midst of treating a wounded soldier when Cohen shot him multiple times.

Forty years later, as his guilt began to eat away at his conscience, Cohen started to show all the classic signs of major depressive disorder: he was overcome by grief, he suffered from anhedonia, he lost his appetite and his sex drive, he became sleepless, he had difficulty concentrating, and he would spend hours brooding over what he had seen in battle and, most acutely, he fixated on that particular irreversible action: his deliberately killing an unarmed doctor at work (Kleinman 2006: 34). To address this, Cohen's psychiatrist Kleinman prescribes anti-depressant medication. This works: the medication does indeed succeed in alleviating the 'symptoms' of Cohen's depression. But Cohen tells Kleinman: [thanks to the anti-depressant medication] 'I can put it away again. I don't feel the same pressure. I can sleep, and eat, and fornicate again. But you know as well as I do that what's bothering me can't be treated or cured. . . . I lost my humanity as those around me did the same. You don't have any answers. Nor do I. Save to live with it. To realise I did the worst is to understand how ordinary men do bad things' (ibid.: 35).

Cohen's suffering is a marker of his moral development, i.e. it is arguably because he has matured morally in the intervening forty years that he suffers presently in this way. What Cohen comes to discover through his searing guilt, experienced acutely for the first time at age sixty, is that a deep compassion for human life forms part of his mature conception of the ultimate good. This compassionate disposition comes to be (at least partly) definitive of Cohen's final end. So at the very same time that the guilt he experiences takes him away from a life of happiness-as-good-feeling, it brings Cohen closer to flourishing in the Aristotelian sense, i.e. both in the sense that it manifests the fact that he has a final end (one that becomes apparent to him through his guilt), and that in having this final end, Cohen exercises his epistemic, moral, and deliberative agency well.

Of course, we could agree that Cohen would have been altogether better off never having encountered the enemy doctor. But given that life turned out the way it did—Cohen did in fact encounter the doctor in a battlefield hospital,

and he did in fact proceed to kill him—a counterfactual version of 60-year-old Cohen who remains blissfully oblivious to this past misdeed would be further away from achieving a life of flourishing than the actual version of Cohen who suffers bitter remorse and guilt. This would be for either one of two reasons: firstly, because the counterfactual version of older Cohen never comes close to forming *any* conception of the ultimate good; this would constitute a failure on counterfactual (older) Cohen's part to exercise his rational agency well.[26] Alternatively, although the counterfactual version of (older) Cohen does form a conception of the good, it would have to be a very different one—one, crucially, in which a deep compassion for human life is absent (I will return to discuss this aspect of Cohen's predicament in the discussion below on the role that emotions play in shaping our judgement). Importantly, within this conception of the final good, counterfactual (older) Cohen would be blind to a significant fact in his own agential history, and this ignorance would mark another way in which counterfactual Cohen would manifest a failure to exercise well his capacity for rational agency.

What the lives of Cohen and Grossman reveal in light of the Aristotelian conception of flourishing is the possibility that a person could flourish—or come closer to flourishing—in virtue of experiencing deeply distressing emotions, and that a person could flourish in spite of living in challenging times, times in which the opportunities for enjoying happiness-as-good-feeling are greatly reduced. These possibilities are realized in virtue of the person's coming to form a conception of the good that he is wholeheartedly committed to—a commitment that is constituted in part through experiencing a range of pertinent emotions, some of which may be distressing—by which means he exercises well his capacity for rational agency.

This suggests an altogether different way of looking at depressive$_{\text{DRIC}}$ experiences and other emotionally distressing predicaments arising from the very human concern with matters of existential significance: they are (or may be revealed to be, if we bring this conception of flourishing to bear on the matter) expressions of our sensitivity to the question, 'How should I live?' And the emotional distress through which this sensitivity is manifest, far from being pathological, turns out be an attempt on its subject's part to exercise her rational (read: distinctly human) agency more fully and completely. For instance, on this understanding, the emotional distress a person suffers in a depressive$_{\text{DRIC}}$ experience might be indicative of her sensitivity to the

[26] Aristotle clearly takes a person's failure to form any conception of the ultimate good to constitute a failure on her part to exercise her rational agency well, stating 'Not organising one's life in relation to some goal is a mark of great foolishness' (*Eudemian Ethics* 1214b9–10; cf. Berryman 2019: 150).

misalignment between her latent (and likely inchoate) conception of the ultimate or final good and the actual circumstances of her life as they currently stand. Externalizing the emotional distress by pathologizing it moves a person further away from grappling with that most human of questions, 'How ought I to live?' From the Aristotelian perspective, to shut down the question before we've even begun to recognize it and to take it seriously is to deny ourselves an opportunity to uncover and make more articulate our conception of the good. It is, in sum, to shift us towards living a defective or incomplete human life, a life in which we languish rather than flourish.[27]

4.5.2 Evaluative Perception: Emotions Enable Us to Perceive Value

We become capable of valuing things, and of being evaluators, once we are able—through the exercise of our embodied reason—to conceive of the good. Out of this emerges a second way in which our affective dispositions figure in the exercise of our epistemic, moral, and deliberative agency: they constitute an important means by which we discern or perceive the value of persons, situations, activities, concerns and commitments, and objects in our lives, and this evaluative perception shapes our understanding of both the daily and longer-term situations we find ourselves in.[28] It is in the (proper) discernment of these values that a person can be said to exercise his or her capacity for reason fully and properly.

The notion that emotional experiences are integral to evaluative discernment or perception is underpinned by a broadly cognitivist conception of emotion (as outlined in Section 1.3). In the context of the present discussion, it is worth underscoring the following two principal claims:

(a) Even if our axiological vision or outlook is framed by commitments and values that are specifiable in general terms, this vision is only ever enacted through our interaction and engagement with concrete particulars. This

[27] Even if thanks to a pharmacological intervention we now find ourselves in a pleasanter mood more of the time.

[28] It is in the *Rhetoric*, rather than the *Nicomachean Ethics*, that we get the most extensive discussion of the role of emotions in enabling humans to perceive value. See Sherman 1997: 52–3 for an examination of this point and of Aristotle's discussion of the role that emotions play in evaluative perception. In what is perhaps the most famous remark that Aristotle makes in this regard, he observes: 'For things do not seem the same to those who love and those who hate, nor to those who are angry and those who are calm, but either altogether different or different in magnitude' (*Rhetoric* 1377b30).

is because the objects of our evaluative perception, i.e. perception on the basis of which we decide and we act, are always particular.

(b) Our emotional experiences are selective (i.e. discriminatory) intentionally structured, motivating states that are responsive to articulated, particular features of our environment. This claim should be understood as being embedded within an inclusive view of evaluative perception, according to which such perception is not just an intellectual or cognitive apprehending of what is perceived. It is crucially also emotional and imaginative in nature.[29]

Drawing these two points together, the claim is that we evaluatively perceive objects, people, activities, and situations through our emotional experiences. A simple comparison illustrates this claim. Someone who is fearful and anxious of strangers (especially strangers who look, dress, and speak differently to oneself) and another who feels basically safe in the world and who therefore has room for feeling compassion readily will likely see very different things when a dishevelled-looking homeless person approaches them on the street. The fearful individual perceives the homeless person as a threat, an unpredictable agent who might drunkenly verbally assault him (or worse); the individual who feels essentially safe in the world sees a fellow human being who is vulnerable and whose hold on life is precarious, i.e. a person whose basic physical needs for food and shelter are neither easily nor consistently met. The object of each of these individual's evaluative perception is a particular person, perceived as threatening or as vulnerable, and the mode of perceiving in each case is emotional: one individual sees fearfully, the other sees compassionately. These modes of seeing arise from the two individuals' respective emotional tendencies and dispositions. Although this particular example relates to a morally significant situation, the notion that our emotional experiences inform the evaluative judgements we make extends well beyond the moral domain. Another person's jokes can seem to us to be witty or crass, depending on the mood we are in (as well as our sense of humour). A task at work can appear as daunting or as exciting, depending on our affective orientation towards the task, and perhaps towards our work more generally. A pastime such as baking or woodworking can seem to us either frivolous or rewarding (or something else), again depending on our affective orientation towards leisure time and what we value and find interesting.

[29] Nussbaum 1985: 189.

This view of evaluative perception is an inclusive one, as noted above, and it is a view best appreciated within the broader framework described in the introduction to this section (Section 4.5), i.e. that our embodiment forms the container within which we exercise our capacity for rationality. On the Aristotelian view, to grasp a feature of our environment that is axiologically salient in merely an intellectual way would be to see it imperfectly or defectively (Sherman 1989: 47). Our affective engagement brings us more fully and completely into the evaluations we make, and in this way embeds our axiological vision more thoroughly into our lives, making it effectual in shaping our decision-making and action tendencies.[30]

Martha Nussbaum's sketch of this view of evaluative discernment is worth quoting at some length:

> Without feeling a part of correct perception is missing . . . . Perception is not merely aided by emotion but is also in part constituted by appropriate [emotional] response. Good perception is a full recognition or acknowledgement of the nature of the practical situation; the whole personality sees it for what it is. The agent who discerns intellectually that a friend is in need or that a loved one has died, but who fails to respond to these facts with appropriate sympathy or grief, clearly lacks a part of Aristotelian virtue. It seems right to say, in addition, that a part of discernment or perception is lacking. This person doesn't really, or doesn't fully, *see* what has happened, doesn't recognise it in a full-blooded way or take it in. We want to say that she is merely saying the words, 'He needs my help,' or 'she is dead,' but really doesn't yet fully *know* it, because the emotional part of cognition is lacking. (Nussbaum 1985: 188, italics in original)

Echoing Nussbaum, Nancy Sherman writes of this Aristotelian view that, 'Often we see not dispassionately, but because of and through the emotions . . . . Without emotions, we do not fully register the facts . . . with the sort of resonance and importance that only emotional involvement can sustain. It is as if we could see, but only flatly and inertly, as if our perceptions were strung together in our minds but not fully understood or embraced' (Sherman 1989: 45–7).

At this point, we can note another important Aristotelian idea. The relationship between what we are capable of perceiving evaluatively, and what is there to be thus perceived, is a dynamic one that deepens over time, with educated (and educable) experience. The richer our emotional repertoire and the

[30] For more on the notion that the affective aspect of an evaluative perception contributes to transforming it into action readiness, as applied specifically to morally significant action, see Oakley 1992: 53–7.

more refined our emotional sensitivities, the more accurate and nuanced our evaluative perception will be. Perception in this sense is something we can get better at doing through pedagogy, enculturation, self-examination (with the aid of psychotherapeutic work), and practice. This normative dimension of evaluative perception introduces into the Aristotelian view the notion of an emotionally rational or reasonable agent. The emotionally reasonable agent is one whose emotional responses fit their eliciting circumstances. Taking fear and courage as an example: an emotionally reasonable agent feels fear to the appropriate degree when confronted by a genuine threat, and she feels courage to the appropriate degree when, in the face of this threat, she stands her ground and acts to uphold the value that is being threatened. This notion of the emotionally reasonable agent is expressed by Aristotle in his famous Doctrine of the Mean. He writes: 'Both fear and confidence and appetite and anger and pity and in general pleasure and pain may be felt both too much and too little, and both cases not well; but to feel them at the right times, with reference to the right objects, towards the right people with the right motive, and in the right way, is what is both intermediate and best, and this is characteristic of virtue' (*EN* 1106b15–23).

The picture of the flourishing human being that emerges here is one according to which a person is able *accurately* to evaluatively perceive relationships, opportunities, obstacles, and daily circumstances generally—from the most mundane moments right through to life-altering situations—through the rational exercise of well-cultivated emotional dispositions. The idea of subjective well-being (construed as a preponderance of positively valenced emotional experiences) has very little to do with flourishing, thus construed.[31] More to the point, if part of what it means to be an excellent rational agent is that an agent desires to know and to understand aright, and to succeed in being disposed to evaluate correctly (as it surely must do, cf. the virtue of prudence), and if it is only through our emotional experiences that we can hope to achieve this, then we must expect to develop and to draw upon a wide emotional repertoire (this would include both positively and negatively valenced emotions), given the richness and complexity, the unpredictability, and ultimately the finitude of human life.

[31] Aristotle did hold the view that contentment is an important albeit auxiliary ingredient in the good life. However, this is not foundational to flourishing in the way that the excellent exercise of our rational agency is. So it is important to keep the two claims separate. Given the vicissitudes of life, we want to be able to recognize the possibility that a person comes closer to approximating excellent rational agency even if (and when) this involves experiencing significant emotional distress.

4.5.2.1 Affective Irrationality: Different Varieties

Within this framework, a disposition that tends its subject towards emotional responses that are excessive (or diminished) relative to their eliciting circumstances is deemed to be irrational. However, this judgement doesn't immediately entail the further judgement that such a disposition signifies a state of disorder or pathology. In fact, rather than framing the tendency towards excessive/diminished affect exclusively in terms of the concept of pathology, Aristotle importantly recognizes a construal of such tendencies according to which they are taken to signify a character defect that impairs an agent's capacity to exercise well her capacity for epistemic, moral, and deliberative agency. For brevity, I shall refer to this construal as the *agential* construal of irrational affective dispositions, to capture the Aristotelian idea that in being disposed to respond emotionally to eliciting circumstances either excessively or insufficiently, we are nonetheless still exercising our rational agency, albeit badly, as opposed to failing to exercise our rational agency at all on account of suffering from a condition of disorder. So on this construal the emotional disposition is not taken to be the result of a pathological process.[32]

In addition to agential forms of affective irrationality, Aristotle recognizes and discusses two other forms of affective irrationality that are non-agential, i.e. affective irrationality that manifests (i) disease or organic injury, or (ii) a brutish nature which is either innate or due to poor enculturation (*EN* 1148b17–20; and 1149b24–30). Having a brutish nature consists in living a life in which reason is largely if not altogether absent (thus, a life that is more beastlike than human), and being beset by a disease can destroy one's capacity for reason. So what seems to mark out non-agential forms of affective irrationality is that individuals who manifest these forms of irrationality are constitutionally unresponsive to reason. An individual who manifests affective irrationality of the agential sort, by contrast, is constitutionally responsive to reason even though he in fact fails to be thus guided (hence manifesting the irrationality that he does).

What, then, does being responsive to reason in our affective lives consist in? For Aristotle, the answer seems to be this: our affective responsiveness to reason is to a large extent anchored in our conception of the good and our pursuit of it. It is our ability to conceive of the good that enables us to apprehend or to perceive values through our emotional experiences, and to perceive

[32] *EN*, Book VII, chapters 4–5.

them accurately or inaccurately. Crucially, it is in virtue of being the means by which we apprehend situations in this way—evaluatively accurately or inaccurately—that our emotional experiences become apt for rational assessment, construed agentially. Thus, to say of an individual's emotional dispositions that they are wholly blind to reason owing to disease or brutishness is to say that such an individual's affective life is not in any way guided—nor is it even guidable—by any conception of the good. It is shaped exhaustively by the desire for the pleasant and the aversion to the unpleasant. This denudes a subject's emotional experiences of their value-disclosing function—and *a fortiori* their aptness for rational assessment. (I will return to the relationship between our capacity to rationally desire the good and our capacity for evaluative perception towards the end of this section, in 4.5.2.2 below.) By contrast, the individual who fails merely contingently to be guided by reason in his affective response(s) is able to conceive of the good and is therein able to grasp the significance of a rational critique of this response—and this is the basis upon which, it seems, Aristotle holds the individual to be agentially irrational.[33]

Given that the distinction is drawn in this way between agential and non-agential forms of affective irrationality, the scope of agential affective irrationality looks to be far wider, and the scope of affective irrationality stemming from disorder (disease/injury) correspondingly far narrower, for Aristotle than it is for those who tend nowadays to equate the notion of emotional disorder with distress. For the criterion that Aristotle thinks marks the absence or presence of a conception of the good in an individual's life—and thus of the individual's responsiveness or unresponsiveness to reason—is that of 'unnaturalness'.[34] This needs to be briefly explained. Although Aristotle invokes this criterion in the context of a discussion about the imperfections of human nature that are rooted in 'deplorable dispositions' generally, he notably includes in this discussion the example of anger (he also mentions fear/cowardice at *EN* 1149a7–8). So it seems reasonable to suppose that included in this broad category of 'deplorable dispositions' are affective ones. But because the focus

[33] This seeing, i.e. of the import of rational critique of one's affective responses, is succinctly illustrated in the following personal account of a 25-year-old man, Peter, who reports that psychotherapeutic engagement helped him become less angry and controlling in his relationship with his 23-year-old partner, Amy. Looking back at how volatile he used to be with his partner and acknowledging that he used to bombard her with texts and phone calls demanding to know her movements, Peter says he would get 'quite angry' if Amy didn't keep up with his communications. He now reports feeling ashamed of his behaviour. Recounting his perspective then, Peter states: 'I just thought it was normal . . . . I was thinking, "there's no problems [sic] here . . . . This is how I've grown up, [this is how] I've seen relationships handled. It's no one else's business." ' Regular counselling sessions, he says, helped him understand the impact of witnessing . . . coercive control during his own childhood. He continues: "Upon some deeper reflection, I noticed the impact I was having on her, I realised things need to change" (adapted from https://www.bbc.co.uk/news/uk-68942471).

[34] Whereas distress can be ubiquitous, unnaturalness is uncommon.

of Aristotle's discussion here is on 'deplorable dispositions' generally, the criterion he offers us for distinguishing agential and non-agential forms of irrationality does not make specific reference to emotion (nor is it illustrated by reference to emotional dispositions). Instead, the demarcating criterion appeals to a notion of 'unnaturalness', i.e. those deplorable dispositions that violate certain fundamental taboos of the human species are ones that fall outwith the bounds of reason, with cannibalism being offered as an example (cf. Natali 2009: 108). Aristotle couches the presentation of this criterion in terms of what an individual finds attractive or aversive in possessing a 'deplorable disposition', and the idea here seems to be that an individual's finding pleasant something that reasonable people would find deeply aversive, e.g. cannibalism, is indicative of there being no conception of the good that guides this individual's affective dispositions. Reason appears to be altogether absent from this individual's life (due to disorder or the individual's brutish nature), so that it is not possible to say that of such an individual that their affective dispositions are responsive to reason. In such a life, responsiveness to reason is not present even as a potentiality.[35]

This is not the case for ordinary affective dispositions that tend towards excessiveness/paucity. These Aristotle views through the lens of continence and incontinence, meaning to say that he holds the view that a person who tends to get angrier than the circumstances warrant is someone who, because he is capable of conceiving of the good and could become less angry through effort and through habituation, is responsive to reason (even though he currently fails to experience anger in a reasonable way, i.e. a way that is appropriate to the eliciting circumstances). This disposition towards excessive anger yields poor judgement in particular situations: its subject will inaccurately perceive slights where there are none, or perceive slights to a greater degree than is actually present, which in turn leads to further miscalculations (often with adverse consequences for both the agent and the targets of his or her anger).[36] For this irrational affective disposition and the poor judgements it yields, the

[35] Whilst acknowledging that the category of 'Anti-Social Personality Disorder' is notoriously vague (see Hare and Harpur 1991 for discussion), and therefore restricting the following point to the narrower category of psychopathy, it is helpful for illustrative purposes to note here that the debate continues as to whether true psychopathy is shaped by anatomical and/or affective anomalies that disable its subject from experiencing a range of morally significant affects. *If* psychopathic subjects are thus disabled (and this is of course a big 'if'), then they would count as 'brutish' in this Aristotelian sense—meaning that they are correctly deemed to be pathologically irrational (and not agentially irrational) on account of being constitutionally unresponsive to reason. For a construal of psychopathy that is consonant with this view, see Duff 1977.

[36] While both vice and wickedness are thus forms of bad character, *vice*, as contrasted with virtue, is distinct from *wickedness*, which is what the deplorable dispositions rooted in disorder or a brutish nature manifest; cf. Natali 2009: 113.

agent can be held accountable, on the Aristotelian view, and he can be held accountable because such an agent does in fact have a conception of the good. This enables him to reflectively appreciate that his disposition towards excessive anger results in poor judgement.[37]

The idea that we can be held personally accountable for our affective dispositions is a familiar one in everyday life, especially in situations that are practically significant. Alluding to our ordinary practice of praising and blaming agents for their affective dispositions, Jonathan Glover writes:

> We are responsible for the harm we do unless some excuse shows either that we did not understand or could not control what we did. We do not escape responsibility by saying that our selfishness, laziness, hatred, or love of cruelty caused us to do the harm. We are held responsible for shaping ourselves so that these features either fade away or at least are kept under control. That they persist in uncontrolled forms reflects badly on our past decisions not to confront them. (Glover 2015: 315)[38]

To make vivid this point, i.e. that we view people as admirable or as warranting opprobrium on account of their affective dispositions, I'd like to turn to an example which centres upon the virtue of courage. Before turning to the example, a few remarks about courage. In Book III of the *Nicomachean Ethics*, Aristotle identifies two constituent feelings in the affective disposition we refer to as courage: feelings of fear and confidence. Courage, said to be the mean between these two feelings, is identified by Aristotle as a virtue of character.[39] Aristotle defines the virtue thus:

> The man . . . who faces and who fears the right things and from the right motive, in the right way and at the right time, and who feels confidence under the

[37] The notion of 'good judgement' contains several ideas: that of accuracy (which is what I will focus on in this section), but also that of being commendable in a value-substantive sense. We speak of good aesthetic judgement or good moral judgement, for instance. I will discuss 'good judgement' with a focus on this latter idea in Section 4.5.3 below where I look at the role of emotions in shaping our judgement.

[38] Glover draws our attention to this practice within the context of arguing against the widespread tendency to assume that if we do have freedom of choice, this freedom has to be total or complete. Against this, Glover puts forward a more nuanced picture that recognizes simultaneously that we can be held responsible for our affective dispositions and that we do not create ourselves out of nothing (Glover again: 'If we go back far enough, we will reach decisions largely or entirely caused by factors outside our control. These causes may be our genes, or not being loved as a child, or some horrible event that happened to our mother in pregnancy . . . . Our degree of responsibility for who we are is limited by the things we have to take as given. For good or bad, how we turn out is *partly* . . . a matter of "moral luck" ' (Glover 2015: 315–16, italics added).

[39] *EN* 1107b1–5: 'With regard to feelings of fear and confidence courage is the mean; of the people who exceed, he who exceeds in fearlessness has no name (many of the states have no name), while the man who exceeds in confidence is rash, and he who exceeds in fear and falls short in confidence is a coward.'

> corresponding conditions, is brave; for the brave man feels and acts according to the merits of the case and in whatever way reason directs. (*EN* 1115b17–20)

The way the two feelings of fear and confidence interact to constitute the virtue of courage is helpfully summarized by the Aristotelian commentator, Daniel Putnam, in these terms:

> Fear and confidence are . . . deeply intertwined, but the two emotions rely on distinct perceptions: the danger of the situation [fear], [confidence in] the worthiness of the cause, and [confidence in] the perception of one's ability. Courage lies in the interface where the limit of our confidence meets the reality of a feared situation. (Putnam 2001: 469)

So as well as feeling fear to the appropriate degree, having an accurate grasp of our own skills and abilities—and having confidence in these—is also a core constituent of the virtue of courage. This combination of fear and confidence is manifest in the example of courage that I now want to turn to. On 26 September 1983, whilst on duty, a Soviet lieutenant colonel named Stanislav Petrov decided to act in breach of military instructions by not reporting computer readings on the country's early-warning systems that appeared to 'detect' an incoming nuclear missile strike from the United States. Petrov (correctly) judged this to be a false alarm, and in doing so, helped to pre-empt the world from coming closer to being plunged into a nuclear holocaust. Petrov's own account of his judgement on this fateful day is revealing:

> The siren howled, but I just sat there for a few seconds, staring at the big, backlit, red screen with the word 'launch' on it. [The system was telling him that the level of reliability of that alert was 'highest'. There could be no doubt. America had launched a missile]. A minute later the siren went off again. The second missile was launched. Then the third, and the fourth, and the fifth. Computers changed their alerts from 'launch' to 'missile strike'. There was no rule about how long we were allowed to think before we reported a strike. But we knew that every second of procrastination took away valuable time; that the Soviet Union's military and political leadership needed to be informed without delay. All I had to do was to reach for the phone; to raise the direct line to our top commanders—but I couldn't move. I felt like I was sitting on a hot frying pan . . . . [Although the nature of the alert seemed to be abundantly clear, what made Mr Petrov suspicious of the apparent alert was just how strong and clear it was.] There were 28 or 29 security levels. After the target was identified, it had to pass all of those 'checkpoints'. I was not quite sure it was

> possible [for a target to pass through all these checkpoints], under those circumstances. [In the end, Mr Petrov called the duty officer in the Soviet army's headquarters and reported a system malfunction.] (BBC 26 September 2013)[40]

Petrov's response to the putative nuclear alert was (very naturally) affectively charged: shock and fear nearly paralysed him. This is hardly a hysterical response given the scale of the putative threat (as well as the very real risk of incurring a severe personal penalty for disobeying protocol). But this fear was constrained by Petrov's confidence in his own judgement that (a) the threat of a US nuclear missile strike was unlikely, and that (b) it was far likelier that the computer readings and the howling sirens were a result of a system malfunction. We look back at this nauseating near-miss and admire Petrov not just for deciding not to raise the alarm, but crucially, also for his courageous disposition, i.e. a disposition that yielded an accurate perception of the threat and which guided his decision to report a system malfunction (rather than to raise the alarm).[41]

What holds of Petrov's courageous disposition also holds of all our agential emotional dispositions: they are praiseworthy or blameworthy, and we can be held accountable for our affective dispositions in virtue of the fact that in responding emotionally to particular situations as we do, we are perceiving value properties—accurately or inaccurately—in these situations, therein exercising our rational agency well or poorly. While it is true that we do not always overtly hold to be blameworthy the subjects of mundane irrational affective dispositions, it would be a mistake to infer from this that we don't do this because such dispositions are disordered, i.e. because such dispositions aren't situated on the continuum of emotional dispositions for which we can be held accountable. Given Aristotle's criterion for demarcating agential from non-agential irrational affective dispositions (namely, the presence/absence for the subject of a conception of the good), most mundane irrational dispositions would arguably be situated on this continuum: so long as the subject of an affective disposition has a conception of the good, she is able to reflectively appreciate that her irrational disposition merits censure (amongst other things, on account of its yielding poor judgement). It is the subject's receptivity to the

[40] https://www.bbc.co.uk/news/world-europe-24280831.

[41] That Petrov manifests a courageous disposition, and that this disposition was a significant determinant in his decision (rather than the decision being the product of cool deliberation), is suggested by the fact that Petrov acted in an emergency, and by his own description of his experience of the emergency. In this vein, Aristotle remarks: 'It is thought the mark of a braver man to be fearless and undisturbed in sudden alarms than to be so in those that are foreseen; for it must have proceeded more from a state of character, because less from preparation; acts that are foreseen may be chosen by calculation and reason, but sudden actions must be in accordance with one's state of character' (*EN* 1117a17–22).

good that is foregrounded when we view her affective disposition through the lens of continence/incontinence. Bearing this in mind, if we do not overtly hold accountable a subject for manifesting an irrational affective disposition, this is likely because the irrationality manifest is simply not sufficiently significant or exceptional to elicit an overt judgement of (say) blame. Nonetheless, these affective dispositions (exceptional or mundane) are agential, at least by Aristotelian lights.

For the Aristotelian, the principal reason why it is important not to lose sight of the fact that in responding emotionally to situations, we are exercising our rational agency (well or badly)—and thus that we can be held accountable for our affective dispositions—is that this is simply the nature of human emotional dispositions: they share in reason (*EN* 1102a34–1103a4). This fact about our affective dispositions is especially visible in the zones of apparent indeterminacy in our evaluative perception. These are zones in which it isn't obvious which emotional response (or a range thereof) is the most accurate one, or where it is possible for more than one kind of response to be accurate or appropriate. We navigate these zones by engaging in rational critique, i.e. justifying the response(s) we judge to be the right one by (for instance) appealing to features of the eliciting circumstances that may have gone unnoticed by others, and in turn revising our own emotional response when others persuade us that there is good reason to do so. This is an important means by which we develop our emotional literacy and our capacity for evaluative discernment, and therein to become better at exercising our epistemic, moral, and deliberative agency.[42] That we find it natural to subject our emotional responses to rational critique is indicative of our taking them to be informed (and inform*able*) by reason.

One qualm about the claim that we are accountable for our (agential) emotional dispositions is that it isn't a psychologically realistic one. In Section 4.6 below, I will explore several of the mechanisms by which our rational agency can be said to be exercised through our emotional responses; a clearer grasp of these mechanisms should dispel this qualm.

[42] Developing our capacity for evaluative discernment matters to our flourishing since our ability to realize the good in our lives turns upon how well we identify circumstances that are evaluatively germane and how well we discern value herein. I will explore the topic of our affective development in Section 4.6.1 below.

4.5.2.2 Desiring the Good and Perceiving Value: Two Connections

One point that has emerged from the discussion so far is that our faculty for evaluative perception is deeply intertwined with our conception of the good, in at least this way: our emotional dispositions constitute our practical orientation towards the good (i.e. to whatever degree that we are thus oriented), and it is through this conception of the good that we become able to evaluatively perceive as good or otherwise (and not merely as pleasant or unpleasant)[43] the particulars we encounter and engage with daily and over time. For instance, a person may respond not just *aggressively* to a physical threat, but in virtue of being able to conceive of the good, also to respond *angrily* to a perceived slight to his sense of dignity, say, or to something that appears to attack what is just (as he conceives of this). It is possible to respond angrily to threats towards dignity or justice only in light of a conception of the good. Creatures with no conception of the good may respond aggressively to threatening stimuli, but they are not capable of responding angrily to a stimulus that is perceived to be an attack on values (such as dignity or justice). We have seen that it is in virtue of this relationship between our conception of the good and our faculty for evaluative perception that our emotional responses can be said to be reasons responsive, i.e. to be informable by reason.

But our ability to conceive of and pursue the good and our ability to perceive value are intertwined in one further way that has so far been dealt with only implicitly: excellently cultivated emotional dispositions—ones that are thoroughly and effectively informed by a conception of the good—enable an agent to exercise her *deliberative* agency in a fine manner. Such an agent is successful in her practical reasoning in the following sense. Unlike continent and incontinent agents, the excellent agent's emotional dispositions are integrated and aligned with what she perceives to be good, and what she perceives to be good is (since she is an excellent rational agent) in fact truly good. Continent and incontinent agents have to wrestle with recalcitrant emotional dispositions that may lead to distortions in their practical reasoning. For instance, being tempted by the (brutely) pleasant may lead such agents to overvalue or altogether misvalue things, i.e. being drawn affectively towards something that should in fact be perceived as bad for them in light of their own conception

[43] The concept of *goodness* here is being used in the broadest sense, representing all the evaluative perceivings it is possible to experience in the personal, moral, prudential, aesthetic domains, as well as all the other domains in which we perceive values in particular situations.

of the good. An example here might be of someone who rationally wishes[44] to be a faithful and loving partner to their spouse, but who finds him- or herself strongly sexually attracted to a neighbour or to a work colleague. The excellent or virtuous agent has no such recalcitrant emotional disposition to contend with.[45] This agent is properly affected by her circumstances, and in this way pursues the good harmoniously in her practical reasoning and her actions, and therein flourishes on account of exercising her deliberative agency well.

It is perhaps worth stating here that the excellent or virtuous agent is something of an ideal. It is certainly an achievement—a mark of psychological maturity, even wisdom—to have our conception of the good be so firmly rooted in the soil of our appetitive and perceptive functioning. But having a clear view of what this ideal is, is helpful for understanding exactly how it is that our emotional dispositions—through their orientating role and their disclosure of values—contribute to the exercise of our rational agency and therefore to our flourishing, as this is conceived of within the Aristotelian framework.

4.5.3 Emotions and Good Judgement

It might be contended that all that has been said so far about the orientating and evaluative-perceptive roles of emotional dispositions in the exercise of rational agency applies to the wise and to the misguided alike. Criminal dictators may be wholeheartedly (i.e. affectively) oriented towards what *they* take to be the good, and whose emotional dispositions accordingly colour in their evaluative perceptions.[46] Yet, we judge such individuals as living (or as having lived) hideously distorted, vicious lives. This possibility brings into view a third principal way in which our emotional dispositions essentially determine

[44] i.e. someone whose *boulesis* is to be a faithful and loving partner to their spouse.

[45] In saying that the virtuous agent has no such recalcitrant emotional dispositions to contend with, I do not mean to portray such an agent as somehow being purified from the messy contingencies of actual human life (though rigidity or reclusiveness, say). So it is important to clarify the sense in which the virtuous agent is said not to be beset by recalcitrant emotions that he has to wrestle with. Here, I find David Carr's characterization of the virtuous agent helpful: 'The virtuous may feel attraction to the partners of their best friends but have the wisdom and self-understanding to put this in its proper human place. It is a brute fact of life that one is humanly prone to such feelings: however, the morally mature recognise that they cannot have everything they desire and that the satisfactions of acting on untoward attraction could not possibly outweigh the personal and interpersonal damage (to integrity, relationships and so on) that such conduct is likely to incur. It may be that some merely self-controlled agents are incapable of coming to mature terms with the prospect of such inevitably ungratified desire' (Carr 2009: 45).

[46] This possibility reflects the formal, rather than substantive, character of the Aristotelian claim that rational agents desire the good (and not just the pleasant). On this point, Rosalind Hursthouse states: 'Aristotle maintains . . . that desire is for the seeming (or apparent) good . . . ; when what seems good in the faculty of desire is truly good, then the desire is right, and the true judgment of reason about what is good will be in agreement with it' (Hursthouse 1991: 66).

our flourishing or languishing as rational agents. According to the Aristotelian conception of flourishing, it is possible to value the wrong things, and to have one's life become badly distorted as a result; conversely, getting it right about what matters most is the soil from which a human life grows into full fruition. Good judgement is essential for flourishing, and for Aristotle, our emotional dispositions play a decisive role in shaping our judgement. They do this in two ways. First, our emotional dispositions—and our affective states more broadly—furnish us with our starting points in practical reasoning about what to do. Second, the emotional dispositions that excellent others display provide us with a living, embodied model for living well ourselves; we learn from them how to cultivate the excellences of character and intellect that are essential for human flourishing. I will consider each of these points in turn.

4.5.3.1 Judgement and Our Starting Points in Practical Reasoning

Aristotle conspicuously never attempts to justify the analytical connection that he takes to hold between flourishing and the excellences of character and of intellect. Nowhere does he attempt to persuade us that we have strong reason for adopting a morally sound outlook on life. Rather, he proceeds on the assumption that he is addressing those who want to know how to become good, being already persuaded that this is a valuable and worthwhile undertaking (*EN* 1095a2–12, 1095b3–12, 1103b26–31, 1151a18–19). This appears not to be an omission on Aristotle's part, but to stem from an underlying view about how people come to be good or bad, virtuous or vicious. According to this view, our moral vision is rooted ultimately in our pre-cultivated (or 'natural') emotional dispositions and affective states. As we grow and mature, these emotional dispositions are transformed through moral education and training so that they become cultivated emotional dispositions (virtues in the 'strict' sense). Our mature characters are constituted by these cultivated virtues.[47] Our cultivated virtues and practical wisdom are closely connected to each other: cultivated virtues are informed and guided by practical wisdom (this is what makes them character virtues or excellences in the strict sense, as distinguished from the virtues or emotional dispositions we are naturally endowed with as human beings), and they in turn provide the starting points for our practical reasoning (*EN* 1144a7).

To elaborate on this view: we do not come to inhabit a morally sound outlook on life by being persuaded through argument and reason. We can't argue our way into caring about living a morally good life. Our natural emotional

[47] *EN* 1144b3.

dispositions—towards compassion, towards generosity, towards fairness, etc.—are the ground from which moral growth and development springs. Equipped with these dispositions by natural endowment, our elders can then seek to inculcate in us good habits from early childhood, so that we grow to inhabit a life in which temperateness, generosity, courage, fairness, compassion, etc. come easily to us.[48] This means not only that we find it easy to feel and act in these ways, but crucially, also that, as adults, we find these qualities desirable. We are drawn to them in others, and we are ourselves attracted to a morally sound vision of life. While we do also have to cultivate the virtues of intellect in order to be able to employ our mature emotional dispositions in a practically wise or virtuous manner (and therein to flourish, i.e. to be an excellent rational agent), inhabiting a morally sound outlook through well-cultivated emotional dispositions is what enables us to exercise good judgement (indeed, good judgement just *is* the ability to discriminate correctly through feeling and through action (Broadie 1991: 77)). It is by possessing good judgement that we are able to identify substantive values correctly, to recognize what is objectively worthwhile, to distinguish between what merely appears to be good and what is genuinely good. That is to say, it is by possessing good judgement—made manifest through our cultivated emotional dispositions—that we find our footing in a world of moral significance at all. Being kind and being inclined towards generosity of spirit: it is by being affectively disposed in these morally commendable ways that we are capable of reasoning practically in a manner that is sensitive to moral constraints, for it is in light of our emotional dispositions that we find intelligible the starting points in practical reasoning that we do. A selfish or a mean-spirited person will see no point (except through the promise of personal gain) in helping others, and this is contrary to how a generous person views the matter. The general principle, 'All things being equal, it is good to help those in need wherever we are able to,' will be self-evident to the generous person only, providing her with an altogether different starting point (from that of the selfish person's) in reasoning about what to do on an occasion when she is presented with a request for help.

Aristotle, being no relativist about values, makes it categorically clear that he thinks that there is a fact of the matter which starting points are the right ones, stating 'Wickedness perverts us and causes us to be deceived about the starting-points of action' (*EN* 1144a35–6). While our practical reasoning skills may grow to be very sophisticated, and we may be able to construct

[48] Burnyeat 1980.

impressive rationalizations for why we did or did not act in the way we did, the entire edifice of our practical reasoning bottoms out ultimately in our character, that is, our mature emotional dispositions.[49] So an adequate answer to the questions that can be asked of any morally significant situation, 'Did the moral agents in that situation see and judge rightly? Were they *able to* see and to judge rightly?', must, according to the Aristotelian, take into account the agents' respective emotional dispositions. It is because our emotional dispositions furnish us with our starting points in our practical reasoning that Aristotle remarks, 'It makes no small difference . . . whether we form habits of one kind or another from our very youth; it makes a very great difference, or rather *all* the difference' (*EN* 1103b25–6).

4.5.3.2 Learning to Cultivate Good Emotional Dispositions: The Social Role of Emotions

Aristotle holds that, in the first instance, it is through the moral education of our emotional dispositions rather than through argument and persuasion that we come to inhabit a morally sound outlook on life, an outlook that is made manifest in good judgement. Aristotle also famously holds that the moral education of our emotional dispositions consists in habituation (and here we might add *enculturation* more broadly), rather than in memorizing a list of general rules and principles of conduct and seeking cognitively and volitionally to adhere to these principles in deciding and in acting. The transformation of our natural emotional dispositions into well-cultivated virtues thus depends crucially, in early life, on the elders who nurture us and bring us up, and as we grow older, our moral learning is enriched further by our intelligent observation of the people we consider to be good. This places human relationships at the centre of moral development throughout our lives and accords an important social function to our emotional dispositions in enabling us to flourish as human beings.

In fact, right at the heart of Aristotle's ethical inquiry we find a character who plays a pivotal role in Aristotle's characterization of his conception of human flourishing: the man (or the person) of practical wisdom (Broadie 1991: 76; cf. *EN* 1141b14–16). Sarah Broadie writes of this person that s/he 'is the personification of that ungeneralisable *orthos logos* [right reason] that is the form of an appropriate response' (ibid.: 76). The practically wise or virtuous person tends to manifest right reason in her responses to the eliciting circumstances, which is to say that her response—an emotional-cognitive-volitional complex

[49] 'Knowledge of what to do . . . depends on goodness of character' (Broadie 1991: 250).

issuing in action—tends to be appropriate to the circumstances. There are two immediate points to be made about this. First, the reason why Aristotle appeals to a *personification* has to do with his particularism. Because he eschews the notion that it is possible to codify standards of moral conduct in terms of a set of principles, one alternative way for characterizing the standards of moral exemplariness in an informative way is by pointing at instances of moral excellence. We deepen and refine our own sense of what right reason—embodied in appropriate responses—consists in by observing and learning from people who are disposed to respond appropriately. So the personification of the practically wise person occupies a central position in Aristotle's characterization of the normative standard of responding appropriately. Aristotle writes, '[Character] excellence, then, is a state of character concerned with choice, lying in a mean, i.e. the mean relative to us, this being determined by reason, and by that reason by which the man of practical wisdom would determine it' (*EN* 1106b36–1107a2). Second, although practical wisdom requires the exercise of intellectual excellences in drawing relevant distinctions and apportioning the features of a given situation their due weight, we have seen that the very fact that the practically wise agent is positioned so as to be capable of exercising right reason and responding appropriately is due to her having well-cultivated emotional dispositions, i.e. her excellences of character. It is good judgement—the ability to discriminate correctly through right feeling—that renders possible and effective right reason (Broadie 1991: 80).

If we are to continue to develop morally, then, part of what this involves is looking to person(s) of practical wisdom for guidance and modelling, closely and intelligently observing how they feel, how they judge, and how they act. Doing this well entails being able to accurately identify such persons and being disposed to finding them attractive role models. How excellent character is manifest thus becomes a significant point of moral interest, and as has been noted in the foregoing discussion, feeling—and lack thereof—manifests a person's character as much as her actions do.[50] In order to determine that a person is an excellent epistemic, moral, and deliberative agent, then, we shall have to look at her emotional dispositions since it is through these dispositions (as well as her actions) that her practical wisdom, her right reason, is

[50] Aryeh Kosman observes: 'Aristotle's moral theory must be seen as a theory not only of how to *act* well but also of how to *feel* well; for the moral virtues are states of character [emotional dispositions] that enable a person to exhibit the right kinds of emotions as well as the right kinds of actions. The art of proper living, we should say, includes the art of feeling well as the correlative discipline to the art of acting well' (Kosman 1980: 105).

made manifest.[51] It is not least because of this that Aristotle takes a serious interest in human feelings in his ethical inquiry.

The discussion in this section (4.5.3) has sought to articulate a third important way in which our emotional dispositions shape the exercise of our epistemic, moral, and deliberative agency (therein shaping our flourishing or languishing): good judgement is anchored in well-cultivated morally sensitive emotional dispositions. This returns us to Winthrop Cohen, the US war veteran whose life I briefly sketched out above. Cohen's life would have been easier in the sense of being significantly less (subjectively) distressing and unpleasant, had he not suffered guilt to the depth and extent that he did. We are now in a position to see, however, that it is in his experience of this guilt that Cohen manifests good judgement. For it is in and through the guilt experienced with searing and penetrating depth that Cohen sees clearly the true moral significance of his action of killing an enemy doctor who was in the midst of working to save lives. He exercises his epistemic and moral agency more fully through this clear-sightedness and so, in his being haunted by guilt, comes closer to flourishing in the Aristotelian sense. This conception of what it means to flourish as a human being has written into it an agent's capacity to emotionally grasp the true moral significance of his actions, and by bringing it to bear on our understanding of Cohen's predicament, we see so much more than merely a man suffering from 'depressive symptoms'. We see, importantly, a man who exercises good judgement—or at the very least, a man who exercises better judgement than he would have done had he been happily oblivious to his own past action.[52]

I say 'importantly', but whether we do find it as important as I've just made it out to be will depend on how heavily we weigh the human ability to grasp moral truths. How important do we consider this ability to be to our understanding of what it means to live a distinctly human life—and thus (at least for the Aristotelian) to our flourishing? Indeed, how important do we think it is to live a life that is distinctly human at all? A disinterestedness in this last question may lean us towards focusing on Cohen's 'depressive symptoms' and

[51] Broadie again: 'What characterises a prohairetic state (as distinct from a skill, say) is that it shows itself in the agent's acting and failing to act *because of his feelings*' (Broadie 1991: 81, emphasis in original).

[52] In a world of moral complexity such as ours is, paying serious attention to one's circumstances (in circumstances that warrant this seriousness) can essentially involve experiencing emotional distress. Just as it is for Winthrop Cohen, so too it was for Dietrich Bonhoeffer: his recognition of the inevitable conflict between his commitment to pacificism and his participation in the plot to assassinate Hitler was gravely distressing, and essentially so. Bonhoeffer's distress embodies his wisdom exactly because we see, in his distress, not an avoidance or a denial of the moral starkness of his awful predicament, but rather a clear-sighted grasp of the gravity of the truth of this predicament. Bonhoeffer's distress is explored in depth in Schliesser 2008.

the distress that this is causing him instead, this focus in turn motivating us to expend our creativity and energy on working out how best to alleviate this suffering—as quickly and as effectively as possible—without so much as wondering about the rather curious fact that here is an organism who is capable of being morally haunted by his own past actions.

4.6 The Mechanisms through Which Our Emotional Experiences Form an Integral Part of the Exercise of Our Rational Agency

I have so far been filling out the Aristotelian claim that our emotional experiences partake in human reason by outlining four ways in which they figure in the exercise of our rational agency: in orienting us towards the good, in our evaluative perception, in furnishing us with our starting points in practical reasoning, and in our cultivation of good judgement by learning from (virtuous) others. The term 'figure in' is worth clarifying in its own right, however, because our emotional experiences are often states we feel *subject to*. When we experience intense rage or grief, it can feel as though we are *struck by* these emotions. Their phenomenology is suggestive of passivity, not agency. This can make the main claim put forward in the previous section—namely, that our emotional experiences figure essentially in the exercise of our rational agency—seem paradoxical. In this section, I urge caution about taking matters at face value because the appearance of passivity that is suggested by the phenomenology of our emotional experiences is not inconsistent with the claim that they figure essentially in the exercise of our rational agency. While much of the previous section (Section 4.5) has been spent identifying the roles that our emotional experiences play in the exercise of embodied reason (as this is conceived of within the Aristotelian framework), it is worth spending a little time reflecting on some of the mechanisms by which they discharge these roles, for this will help us to appreciate that the phenomenology of passivity does not preclude the possibility that the human exercise of rational agency is constituted in part through our emotional responses and the tendencies that shape them.

It is helpful to begin this inquiry into mechanisms by noticing a pervasive habit of thought that encourages the view that we are passive subjects of our emotional experiences, namely that of privileging the performance of intentional actions when we picture an agent exercising his or her rational agency. In acting intentionally, an agent exercises direct, wilful control over his body

and the objects that are the target of his actions. This habit of thought construes the exercise of human rational agency in strictly present-tense volitional terms. By equating the exercise of rational agency with the exercise of direct wilful control in this way, i.e. volitionally in the present moment, we are led from the phenomenology of passivity in our emotional experiences to the view that our emotional experiences can't constitute an exercise our rational agency since this phenomenology implies exactly the opposite.

I want to suggest, first, that although construing the exercise of our rational agency in control-based volitional terms tends to yield a narrow view of reason's extent and scope in human life, there is an important way in which our emotional experiences can be said, on a sufficiently broad conception of volition, to constitute an exercise of our rational agency even in this control-based sense. I discuss this in Section 4.6.1 below. Having noted this, however, I then note that the scope of our rational agency arguably extends far wider than just the sphere of what we exercise control over. Some of our emotions, while experienced seemingly passively, simultaneously seem to form an essential part of our sense of who we are, therein grounding the very possibility of exercising our rational agency in a manner that is self-directing and autonomous. Recognizing this dimension of our rational agency will require us to move beyond a control-based understanding of what having and exercising this agency consists in. I will explore this point further in Section 4.6.3 below.

4.6.1 Exercising Control over Our Emotional Responses

Emotions cannot usually be experienced or relinquished at will, and it is especially with respect to the valence quality of our emotional experiences that we can feel passive. When we experience intense rage or grief, for instance, it is the intensity of the *felt sense* or valence quality of these emotions that can lead us to experience ourselves as being *struck by* these emotions. Yet, although occurrent emotional states typically have a valence quality, it is important to recall that they can be motivating and that they are also content-bearing states that can be assessed as rational/irrational. I explored this point in Section 4.5.2 above, noting that emotional dispositions are states of character that are amenable to being critiqued and for which we can be held accountable. Indeed, it is in virtue of this multifaceted structure that our emotional experiences are apt for discharging the four roles that have been identified in Section 4.5 (and in these ways can be said to 'share in' our reason).

To make intelligible the idea that our emotional states are amenable to being rationally assessed, in spite of the appearance of our passivity with respect to

them, I now want to turn to consider one set of mechanisms in virtue of which our emotional states come to constitute an exercise of our rational agency in their own right. This set of mechanisms I shall refer to as *enriched habituation*, and I take the concept of enriched habituation to be a development of Aristotle's view of character transformation that famously emphasizes behavioural habituation specifically. A typical passage from the *Nicomachean Ethics* on the idea of habituation reads:

> We are adapted by nature to receive [character virtues], and are made perfect by habit . . . . By doing the acts that we do in our transactions with other men we become just or unjust, and by doing the acts that we do in the presence of danger, and by being habituated to feel fear or confidence, we become brave or cowardly. The same is true of appetites and feelings of anger; some men become temperate and good-tempered, others self-indulgent and irascible, by behaving in one way or the other in the appropriate circumstances. Thus, in a word, states of character arise out of like activities . . . . The virtues we get first by exercising them. (*EN* 1103a25–b26)

Julia Annas emphasizes that the cultivation of better habits is 'neither easy nor mechanical. Thought and reflection are needed,' both in order to embark on the project of cultivating better emotional dispositions through habit, and then in intelligently monitoring the way a person's attitudes, feelings, and behaviour evolve as he continues to work on improving his emotional habits (Annas 1993: 57–8).[53] Annas' characterization of this process of critical engagement is situated within the picture of habituation that Aristotle himself explicitly discusses, where the focus is on behaviour specifically.

We can add to this picture, however. Given that our understanding of human psychology has evolved quite considerably since Aristotle wrote 2,500 years ago, notably in the development of depth psychology and our understanding now of the pervasive influence that early attachment patterns can exert over our lives, it seems reasonable to allow that character transformation can involve other dimensions besides behavioural habituation. Importantly, enriching our concept of habituation by taking into account our current understanding of human psychology will be crucial to addressing the charge noted in Section 4.5.2 above, namely that the Aristotelian claim of emotional accountability presumes a psychologically unrealistic view of the relationship in which we stand to our emotions. I'll develop the concept of enriched habituation by addressing this charge.

[53] Cf. Burnyeat 1980.

The charge that the Aristotelian picture of our emotional lives—and the conception of flourishing which it informs—is psychologically unrealistic is motivated by at least one important concern. Aristotle's account of practical rationality can appear to presume what John Cottingham calls 'rational hegemonism', a view according to which our emotions are entirely governable by reason. Cottingham describes this view as follows:

> *Rational hegemonism*. A view of the good life as one . . . where the development of a full range of human capacities and dispositions, emotional as well as ratiocinative, enters into the blueprint for fulfilment—subject nonetheless to their being appropriately channelled and guided by reason. (Cottingham 1998: 36)

Contrary to this view of ourselves, the depth psychological framework of the twentieth century suggests that there is much about our emotional lives that is rooted in the unconscious workings of our psyche—thus lying as a matter of principle beyond our rational control. It is futile simply by dint of conscious, deliberate effort to try to corral our emotional tendencies into the enclosure of reason: the tenacity with which these tendencies cling to us in the face of such effort is only too familiar. As they seem to defy our rational control, how then can we be held accountable for our emotional tendencies—particularly those that subvert or undermine our pursuit of the good? The Aristotelian conception of flourishing, insofar as it presupposes rational hegemonism, looks to be naïve and psychologically unrealistic.

It seems to me that this charge rests upon an unduly restrictive picture both of the concept of Aristotelian habituation and of the human predicament as portrayed by the depth psychological framework, by attachment theory, and other paradigmatic approaches found in contemporary psychology. We can loosen these restrictions by drawing together and integrating the two perspectives. This integration promises to enlarge and enrich the notion of Aristotelian habituation, but it needn't stretch the notion to the point that would require us to overthrow the presumption of rational hegemonism altogether (although this presumption will itself be cast in a new light as a result of our widening the notion of Aristotelian habituation). So while it seems true that without being informed by contemporary psychological theory and psychotherapeutic approaches, Aristotelian rational hegemonism does indeed threaten to become a vacuous ideal, uninhabitable by human beings—and this would go against the grain of Aristotle's naturalism (discussed in Section 4.4.1 above)–I shall argue that there is a specific sense in which the hegemonism of rationality is maintained even on an enlarged notion of habituation. It is maintained in virtue of the fact that the Aristotelian conception

of flourishing, as it has been presented so far in this chapter, can provide a distinct overarching aim for psychotherapeutic engagement, and it does so in a way that makes it meaningful for us (still) to hold ourselves and each other accountable for our emotional tendencies.

To see this, we can begin with the following observation. Our turn towards depth psychology and other modern psychotherapeutic models, in search of a greater understanding of ourselves, is seldom if ever left just at that, i.e. we do not take these models to yield merely a passive observational understanding of the unconscious workings of our psyche, workings that we accept as being forever out of our conscious, rational reach and about which we can do nothing. Rather, the understanding we seek generally seems much more active, even participatory: at the very least, we seek out this understanding in the hope of bringing about a change in our emotional tendencies that will make our lives better in some way. Now, 'better' may be understood solely in terms of reducing or removing distressing emotional experiences. In this case, and if we deem distressing emotions to be disordered on account of their distress, then the following view of psychotherapeutic engagement is the one that would most obviously suggest itself to us: the entire point of turning towards depth psychology and to engage in psychotherapeutic work would be to attempt to get rid of distressing emotional tendencies (as Charles Taylor remarks), tendencies that are assumed to be disordered on account of the distress they cause, and which in turn are taken to be devoid of meaning in virtue of being deemed to be disordered.

However, the Aristotelian conception of flourishing provides us with a distinct take on the turn towards depth psychology in the quest to deepen our self-knowledge. Equipped with this conception, it now becomes possible to construe the practical understanding of ourselves that we seek out as being aimed at the cultivation of emotional tendencies in ways that tend towards the finer exercise of our rational agency. We seek, that is, to develop the capacity to realize our conception of the good (a conception that can itself evolve and change) more effectively and seamlessly. This quest might take the shape of excavating a depressive$_{\text{DRIC}}$ experience to bring to consciousness its subject's particular conception of the good which lies latent in its depths (as described in Chapter 3). It might take the shape of uncovering and understanding how our early life experiences have shaped our emotional tendencies in ways that seem to subvert or to thwart our pursuit of the good, often on account of the irrationality they seem to manifest. The inaccurate perception of value yielded herein poses a systematic obstacle to our being able to apportion things their due weight, which directly sabotages our pursuit of the good (we'd like to be

able to indulge less in food; we'd like to be less prone to feeling fearful or bitter or angry; but our emotional tendencies always seem to get in the way). In engaging psychotherapeutically, we plumb the depths of the (largely mysterious) workings of our psyche in order to understand ourselves better, in the hope of resolving the psychological knots and binds that power up and continue to give fuel to our irrational emotional dispositions, as part of an effort to develop well-cultivated emotional dispositions that enable us to exercise our deliberative agency in finer ways, i.e. ways that are more aligned with our conception of the good.

At the heart of this understanding of psychotherapeutic engagement lies the recognition of the possibility of character development through the transformation of our emotional dispositions. This is a far richer vision of the overarching aim of psychotherapeutic engagement than that which is aimed squarely at the treatment of disorder. In contrast to the latter view, according to which problematic emotional dispositions are judged as disordered and are thus taken to be devoid of meaning, our emotional dispositions are viewed on this alternative as subjectable to rational critique, in the ways just outlined in the previous paragraph.

At this point, though, the following concern may be pressed: even if we accept the Aristotelian claim that our emotional dispositions are thus subjectable, is this claim sufficient for asserting the further claim that we can rightly be held accountable for these dispositions? All that we seem to have so far is a contingent relation between emotional tendencies that are seen by us as obscuring or thwarting our pursuit of the good, on the one hand, and on the other, our judgement that they are problematic for this reason. We may decide that it is apt to claim that such emotional tendencies are subjectable to rational critique in light of this contingent relation, but surely we need more than this for accountability. Don't we in addition need to be responsible for the formation of the tendencies deemed problematic, and isn't the key insight of the depth psychological framework that the processes of emotional disposition formation (and maintenance) lie outwith our conscious control—and thus that we can't be responsible for *this*?

I want to suggest here that it is the possibility of emotional evolution through self-reflective psychotherapeutic engagement that dissipates this apparent dichotomy between an emotional disposition's lying outwith our conscious control and its being something we can be held responsible for. The point of crucial significance here is this. The maintenance of an emotional disposition outside our conscious perspective *at a given time* does not entail that we can have no part in reshaping or revising it *over time*, and enriched habituation is what makes the latter possible. To explain: we are changed in all sorts

of ways by our life experiences, but the kind of change that puts our affective tendencies in greater alignment with our conception of the good—as well as the kind of change that gives rise to a worthwhile conception of the good—seldom if ever happens by chance alone. It is brought about by a person, endowed with the capacity for reason, who reflects upon his emotional tendencies and engages psychotherapeutically to understand himself better and to reshape these tendencies. Take Alec, who has struggled for some years now with a disposition towards anxiety which he himself judges (cognitively) to be excessive. This struggle leads him to enlist psychotherapeutic support. In thus engaging, Alec needn't be told (merely) to fake courage till he makes it (as the narrower notion of behavioural habituation might prescribe). In today's context, the work of bringing about a change in character can (and arguably should) be informed by an understanding of the human predicament as it is laid out by the various modern psychological models of the human psyche. Thus, depending on his personal circumstances, this work could very well involve Alec's undertaking depth psychological self-examination; it might also directly address the various facets of his experience of anxiety itself. Consider for instance the appetitive dimension of anxiety: effectively transforming an anxious disposition perhaps requires taking into account our natural desire for the pleasant in the form of feeling safe. The chronically anxious person does not enjoy a feeling of basic safety—as manifest in his perpetual sense that the world is a threatening place—and so working to increase this individual's feeling of safety could form part of an enriched habituation process aimed at helping him to feel fear appropriately, i.e. perceiving threats only where such actually exist, and to the degree that they actually do; in a nutshell, judging threats accurately. Importantly, this process of enriched habituation necessarily involves the Alec's intelligent and reflective active participation throughout: it involves his preparedness to engage creatively with his feelings of fear, to examine the origins of these feelings and to recalibrate them against the reality of his present context, to work on increasing his general feeling of safety in the world (perhaps by developing the feeling of safety in his own body), to monitor his evolving fear responses, and suchlike.[54]

For the sake of comprehensiveness, it is useful to bear in mind how very broad and eclectic are the range of practices that we now have at our disposal, and that the part we play in psychotherapeutic engagement now goes well beyond merely that of lying on a couch and free associating in the presence of an

[54] In fact, the safety feeling is underscored in a growing body of research and practice in psychotherapeutic approaches that take trauma as its central construct. For a sample of this research, see Porges (2004 and 2011); Payne, Levine, and Crane-Godreau (2015); Gibson (2019).

impassive observer-psychoanalyst. The latter point becomes especially clear when we embed our understanding of these practices within an Aristotelian framework. Thus, with respect to the role that our emotional experiences play in enabling us to perceive value, notice that the cultivation of our emotional sensibilities requires us to exercise our self-reflective ability to appraise as irrational those of our emotional responses that are irrational and to recalibrate and modify these responses over time by various means—for instance through reframing our understanding of situations, through depth psychological self-examination and personal integration, through psychophysical resourcing and meditative techniques, by modelling those whose emotional responses are reliably more accurate and practicing responding in emotionally more accurate ways ourselves, and if an irrational response tends to recur in interpersonal contexts, through dialogue and mutual support. All these form part of the rich tapestry of enriched habituation. In the orientating role that our emotional experiences play, it is worth noticing too that emotional cultivation can involve our attempts to widen and deepen the range of our emotional sensibilities—for instance by observing wiser others, and by critically examining intelligent character portrayals in literature and in film—thereby enlarging our affective perspective on life. This enlargement can result in a transformation of what we conceive of as good, of what we value, in addition to modifying the patterns of our emotional responses themselves. It is in these numerous ways that we manifest our ability not merely to reflectively observe our emotional tendencies as passive by-standers, nor even merely to manipulate them indirectly (by deliberately avoiding emotional triggers, say),[55] but to actively engage our emotional experiences *qua* content-bearing states directly through our use of reason.

It is true that the duration involved in the processes of emotional evolution just described will likely be much longer than the time it takes to act intentionally with respect to our own bodies and to objects in our vicinity. This will mean that the notion of direct, wilful control cannot apply to emotional change in the way it does to intentional action. All the same, since our efforts to cultivate wiser emotional dispositions through enriched habituation do have a target, just as intentional action does, it doesn't seem *too wide* a stretch to say that these efforts are intentionally structured. Justin Oakley, drawing on Michael Stocker's work, makes a useful distinction here between complete foresight and schematic foresight. While we generally have complete foresight of simple physical acts—such as raising our arm and holding our breath—and

[55] Which is to say: our emotions do not figure in the exercise of our rational agency merely in these shadowy ways.

are thus able to spell out the targets of these acts to a high degree of specificity, the targets of more complex physical acts, such as composing a musical piece, are not specifiable in this way. There is an open-endedness about the target that is realized (i.e. it is not altogether predictable what the finished musical piece will actually sound like). Yet the act of composing is very clearly intentionally structured: to engage in the activity of composing, the composer has to employ his musical skills and he does so intentionally, i.e. in order to compose a musical piece.

Our engagement in the various activities of enriched habituation is similarly intentionally structured: the Aristotelian *telos* of exercising our rational agency well enables us to see ourselves as engaging with schematic foresight in the various activities and practices of enriched habituation.[56] Widening our conception of the volitional sphere to include schematic foresight puts us in a position to appreciate that this sphere need not (and should not) be restricted to just our exercise of direct, wilful control over ourselves and our immediate environment. It is the habit of privileging of intentional actions when thinking about human volition that leads us to focus on relatively brief time slices (minutes or hours, at most several days), but the processes of enriched habituation underscore the point that the human volitional sphere extends far wider than this, and that it importantly encompasses more gradual, temporally extended forms of reason-informed character transformations that are intentionally structured. So while it may be true that, when seen at a particular time, many aspects of our emotional dispositions look to be formed and maintained outside our conscious perspective, it does not follow from this that we can play no active part in revising these dispositions over time. Modern psychological models of the human psyche suggest that it is possible to bring into consciousness these aspects of our emotional dispositions, and on this basis to seek to cultivate finer emotional tendencies that are rooted in an ever-deepening understanding of ourselves. Our active, reason-informed participation is woven into the fabric of this endeavour, and this is what warrants the claim that we can rightly be held accountable for our emotional dispositions. The hegemony of rationality is not overthrown, but now we come to see it in a new light. By integrating the Aristotelian and the depth psychological perspectives, the view that emerges is not of reason dominating or suppressing our emotional tendencies (such a view *would* indeed yield a vacuous ideal of human flourishing). Instead, equipped with the notion of enriched habituation, we come to see that reason can (as Cottingham puts it) ' "guide" or "lead the way"—not

[56] Oakley has carefully elaborated this point (Oakley 1992: 135–40).

so much [through] the suppression of the emotions as their channelling into fruitful paths' (Cottingham 1998: 44).

Summarizing, we may note that through the various mechanisms of enriched habituation, our rational agency is exercised twice with respect to our emotional experiences. Firstly, in virtue of the fact that we employ our reason in engaging the activities of enriched habituation, the evolution of our emotional lives that results from our engagement looks to be intentionally structured. Secondly, since at least some of our emotional experiences emerge constitutively (and not merely causally) from our engagement with the activities of enriched habituation, they can—in all the roles they play, as enumerated in Section 4.5 above—rightly be said to instantiate an exercise of reason in their own right (unlike brute or unhabituated emotional states).

The availability of the lens of enriched habituation has an important bearing upon the way we think about psychotherapeutic engagement, for it now becomes possible to construe psychotherapeutic work not principally in clinical terms, but instead as a form of psychoeducation that is aimed at supporting individuals to develop their character by improving their emotional and psychological literacy and through the cultivation of wiser emotional dispositions. Working therapeutically with problematic emotional dispositions no longer has to entail presuming that these dispositions are disordered—and it thus no longer entails presuming that what they call for is clinical treatment aimed at suppression, or better still, at eradication—since such dispositions when viewed through the lens of enriched habituation are now understood to be malleable states that are amenable to rational critique and revision, states for which we can be held accountable. This is not in fact a new idea. As long ago as the 1930s, Carl Jung wrote:

> What was formerly a method of medical treatment now becomes a method of self-education, and therewith the horizon of our modern psychology is immeasurably widened. The medical diploma is no longer the crucial thing, but human quality instead. This is a significant step. All the implements of psychotherapy developed in clinical practice . . . are now put at our service and can be used for our self-education and self-perfectioning. (Jung 1933: 54)[57]

[57] In his illuminating treatment of this topic, Cottingham notes that in breaking with Freud's clinical spin on psychoanalysis, Jung comes to adopt a 'humane Aristotelian model of carefully developed habits of feeling and action.' Cottingham elaborates: 'There is a kind of convergence between the aims of psychoanalysis, as interpreted by Jung, and those of traditional Aristotelian-style ethics. So far from being . . . an unended revisiting of the painful traumas of the past, the process of 'transformational' analysis aims to move forward to a phase of . . . growth. There is of course a major difference from the Aristotelian blueprint . . . . The notion of a systematic redemption, or reclamation of the past, so crucial for what gives the psychoanalytic movement its appeal, is simply missing in the ratiocentric world of the *Nicomachean Ethics*. But for all that, the two approaches come together insofar as the goal in each case is the realization of a harmonious

What the Aristotelian framework adds to this understanding of the overarching aim of psychotherapy is the specification of the notion of self-perfectioning in terms of the excellence of our epistemic, moral, and deliberative agency as an ideal—and along with this construal, an emphasis on the exercise of our rational agency in and through our emotional tendencies as they are and as they evolve. In its turn, the BED analysis of the concept of emotional disorder clarifies the foundation upon which this framework rests. Specifically, what the BED analysis makes lucid is that, to a large extent, it is matter of choice whether to view those emotional experiences that we deem problematic in light of the EF or the HF, and thus whether to view such experiences as disordered or otherwise. This is because whilst the claim of disorderedness is anchored principally in negative appraisals made in the dimensions of valence and agential efficacy, the BED analysis shows that one and the same emotional experience may simultaneously bear existential significance and be judged as rational or irrational (in agential terms). By conceiving of human flourishing primarily in terms of subjective good feeling, the HF encourages a prioritization of the former appraisals over the latter, therein anchoring the claim of disorderedness. By contrast, greater emphasis is placed on the latter appraisals when we regard the emotional experience as an exercise of our rational agency—particularly in its orientating and value discerning roles—and there is no value-neutral standpoint from which the claim of disorderedness can trump the latter Aristotelian (EF) framing of this experience, since (as the BED analysis reveals) we are faced with the task of negotiating competing appraisals of the emotional experience, and which order of prioritization we favour will depend largely on how we conceive of human flourishing. It is against the backdrop of this analysis of the concept of emotional disorder that we see how psychotherapeutic work can be construed as a distinctly pedagogical undertaking—that is, as a form of psychoeducation oriented towards character development and maturation, separate from the clinical aim of treating emotional disorders. (I shall return to the function(s) of psychotherapy in Section 5.3.)[58]

and fulfilled life . . . . One which allows our natural capacities to grow . . . . So that we can reach our full potential as human beings' (Cottingham 1998: 148).

[58] It might be argued that Cognitive-Behavioural Therapy (Beck and Alford 2009) aims at improving a person's exercise of their rational agency by encouraging them to revise 'erroneous beliefs'—which generate distressing (and disordered) emotional dispositions—and to form more accurate beliefs. If the overarching teleological framework within which this belief revision is embedded is that of fixing a disorder, however, this obscures the Aristotelian apical goal of becoming a more excellent rational agent, therein substantially altering the interpretative context within which the notion of rational agency is situated.

4.6.2 Recalcitrant Emotions

For all that has been said about our engaging the activities of enriched habituation, there is nonetheless the familiar phenomenon of recalcitrant emotions, i.e. emotional dispositions that seem to stick rigidly to us in spite of the revisions we might make to the beliefs that underlie them, or in spite of all the other ways we might try to loosen their grip on us. In virtue of not being amenable to being transformed through enriched habituation, we tend to judge recalcitrant emotions as lying outside the scope of our rational agency, and we are correlatedly more reluctant to hold persons accountable for emotional dispositions of this kind. But even here, it is interesting to note that our responses to recalcitrant emotions are finely poised between a recognition of their stickiness and our judgement of their normative significance. A phobia of spiders can be easily relegated to the so-called realm of law, located outwith the space of reasons (McDowell 1994; cf. Sellars 1956b). Hatred or suspicion towards a particular racial group, by contrast, is not so easily relegated, even if a person has with sincerity sought to transform this emotional tendency. The moral import of this tendency leads us to expect more of the person who is beset by it: we look for more evidence that her efforts at enriched habituation are genuine and sincere, that she has utilised the resources available to her as best as she can. When we have exhausted all these considerations, and if her efforts are sincere, we are led to the view that this person is alienated from this emotional tendency ('we' here includes the person whose tendency it is). It is a part of her psyche in an empirical sense, but it does not truly belong to her.

This move to cast the emotional tendency as something from which a person is alienated is, I think, a reflection of our ambivalence about relegating it to the realm of law completely, an ambivalence rooted in our recognition of its moral import. Relegating it wholly to the realm of law would put it on a par with processes like digestion and respiration: such processes lie categorically outside the scope of our rational agency. By contrast, speaking of the tendency towards racial suspicion as a tendency from which its subject is alienated places it not quite outside the scope of our rational agency, but not at its heart either. We attempt to strike this balance between its recalcitrance and its moral import by situating the disposition to racial suspicion at the liminal edge of a person's rational agency (this, on the proviso that we are convinced the person has sincerely sought to transform this tendency). Compare: we *could sp*eak of a normatively irrelevant emotional tendency such as arachnophobia as a tendency a person feels alienated from; but we could just as easily consign it completely to the realm of law (which effectively places

it outside the scope of its subject's rational agency altogether) at no great cost. This ambivalence we feel about where to situate in an agent's psyche recalcitrant emotions that bear normative significance puts into relief the fact that we ordinarily do view as integral to the exercise of our rational agency emotional tendencies that are amenable to change through enriched habituation (i.e. non-recalcitrant emotions).

I now want to turn to one other means by which at least some of our emotional experiences form an integral part of our rational agency, namely by constituting our embodied conception of the good. The suggestion here is that our emotions serve as an anchoring mechanism: certain emotions we experience anchor and complete the judgements we make, the commitments we undertake, the plans we formulate, the actions we perform, and the communicative responses we offer. Such emotional experiences, in sum, form an essential part of the very fabric that constitutes our rational agency.

4.6.3 The Fabric of Our Rational Agency: Endorsement, Disavowal, and the Configuration of Our Intrapsychic Landscape

I wish to begin exploring this idea by noting another way by which we can feel passively subject to our emotions, and this is through a familiar phenomenon that Harry Frankfurt terms 'volitional necessities'.[59] In contrast to emotional tendencies we judge to be 'recalcitrant'—tendencies that are, by definition, experienced as ego dystonic to its subject—volitional necessities lie at the very heart of our sense of who we are *qua* rational agents. They are, as Frankfurt puts it, what 'fixes our shape as active beings' (Frankfurt 2004: 50), and they do this by guiding and delimiting our rational agency, 'setting the boundaries of our practical life' (ibid.: 50). We have encountered this idea that emotions are an anchoring mechanism, implicitly, in Section 4.5.1 when we considered the role that our emotional dispositions play in orienting our entire organism towards the realization of the good in our lives. Vasily Grossman's passion for truthfulness (and the emotional tendencies that contribute towards constituting this passion), and Winthrop Cohen's deepening grasp of the value of human life (again, with all the emotional tendencies that constitute and that manifest this insight), orientate both men towards the good practically, therein anchoring their rational agency. I want to suggest here that Grossman's

[59] That is, a manner other than the way by which we feel related to the valence quality of our emotional experiences.

and Cohen's respective emotional tendencies serve as an anchoring mechanism in virtue of being Frankfurtian volitional necessities.

My reason for introducing Frankfurtian volitional necessities into the discussion at this point is that Frankfurt has written extensively about volitional necessities as a mechanism that anchors and constrains individual human agency, therein showing how our volitional necessities come to form an essential part of the agency that we exercise. Frankfurt's work thus promises to shed light on how our emotions can, when they take on the guise of volitional necessities, function as an anchoring mechanism—and in this way come to form an essential part of the means by which we exercise our rational agency. Before expanding on this claim further, it is perhaps worth briefly clarifying the assertion that emotional tendencies can be volitional necessities. Frankfurt is emphatic that volitional necessities are volitional, stating:

> As in other modes of caring, the heart of the matter [regarding love, which is a volitional necessity] is neither affective nor cognitive. It is volitional. Loving something has less to do with what a person believes, or with how he feels, than with a configuration of the will that consists in a practical concern for what is good for the beloved. This volitional configuration . . . [guides the lover] in the design and ordering of his relevant purposes and priorities. (Frankfurt 2004: 42–3)

In the first chapter, I noted that our emotions are composite states that contain cognitive, affective, and motivational dimensions. In denying that the affective lies at 'the heart of the matter' in a volitional necessity, I take it that Frankfurt isn't denying that volitional necessities can take the form of emotional tendencies. On the contrary, his example of love as a volitional necessity suggests otherwise. But it is in virtue of containing a motivational dimension, specifically, that our emotions can take on the form of volitional necessities. I read Frankfurt as saying in this passage that it is this specific dimension that figures most prominently in an emotional tendency *qua* volitional necessity.

Having clarified this point, I now want to turn to the mechanism of anchoring itself and to draw upon the concept of Frankfurtian volitional necessities to shed light on this mechanism. It is interesting to notice that, although Frankfurt's liberalism is radically un-Aristotelian, Frankfurt nevertheless does share with Aristotle the view that essential for exercising our rational agency well is our having a hierarchy of ends that is encompassed ultimately by a set of final ends (volitional necessities are what, for Frankfurt, constitute our final ends).[60] What is distinctive about this hierarchy of ends in

[60] Frankfurt 2004: 52–3.

agency that is specifically rational (in contrast, say, with agency that is merely biological) is that they are ends *for* their subjects. Thus, as rational agents, we can reflectively assess our motivations in light of these ends, as well as critically appraise the ends themselves. So a further observation we find in both Frankfurt's and Aristotle's accounts of human agency is that our rational agency is structured by the reflective attitudes we adopt towards our mental states (including our emotional experiences), most broadly through the attitudes of endorsement and disavowal. Aristotle draws on this observation to distinguish the different grades of human flourishing, i.e. from the finest exercise of human rational agency by the agent who is practically wise or virtuous, through to the thoroughly poor exercise of this agency by the vicious agent (whose conception of 'the good' is corrupt; he misconstrues what counts as good). What distinguishes the practically wise from the continent and the incontinent, for Aristotle, is the reflective stances they respectively adopt towards what motivates their actions. The practically wise person fully and wholeheartedly endorses these motivations—motivations that will importantly include emotional states—and the actions that ensue. She is psychically unified. The continent and incontinent by contrast are not: both find themselves with a set of motivations not all of which they fully endorse. The difference between them is that the continent agent is able, through effort, to make effective only those motivations that she fully endorses; whereas the incontinent agent is unable to do this. So the organizing principle that Aristotle uses to distinguish these three categories of rational agent turns out to be framed in terms of a broad pair of notions made familiar more latterly by Frankfurt, i.e. that of reflectively identifying with, or disavowing, our motivating states (and our mental states, more generally).

Although Aristotle presses this point in service of an account of human flourishing, and Frankfurt is concerned with human freedom and autonomy, what is germane to my present concern is that Aristotle and Frankfurt both discern in our human or rational agency an intrapsychic landscape that is configured by the reflective stances that agents adopt towards their (lower-level) mental states. Frankfurtian 'volitional necessities' are the most complete expression of the reflective stance of avowal or identification. They are also the most ultimate. Volitional necessities are our answer to the question, 'How should I live?' In spite of being choiceless in a Frankfurtian sense, it is through these necessities that we manifest our responsiveness to this question and therein we manifest our nature as agents who are capable of self-reflectively valuing our own lives and who seek to steer ourselves accordingly. To restate the point, our volitional necessities constitute our *embodied*, motivationally

potent, conception of the good, a conception without which our theoretical ideas concerning the good would fail to latch on to our practical agency—and thus, arguably, without which it would not be so much as possible for us to exercise our practical agency in a way that is apt for normative assessment (see Section 4.4 above).[61] Our volitional necessities thus anchor our rational agency, thoroughly informing and permeating the multitude of ways we exercise this agency. In serving as anchors, they come to form part of the very means by which we exercise our rational agency.

A corollary of this point is that these volitional necessities, and a whole range of emotional experiences besides, constitute an exercise of reason in themselves since they are experienced not brutely but rather *as* endorsed or *as* disavowed. Since we engage our reason in adopting the reflective stances that we do—these stances being shaped to a significant degree by our understanding of what is morally good, what is prudential, what is epistemically sound, etc.—our emotions-as-endorsed and as-disavowed directly express our engagement with reason. To illustrate this point, I want to return to Winthrop Cohen. In his experience of guilt, Cohen suffers considerable distress—distress that is significant enough to detrimentally affect his physical functioning. What is remarkable in the face of this distress is not merely the fact that Cohen does not disavow the guilt he feels, nor does he wish it away.[62] What is striking is that Cohen seems to accept that his guilt—through which is manifest his growing awareness of the intrinsic and irreplaceable value of each individual human life—has a perspective-shaping role, coming to define his sense of who he is *qua* moral, human agent. And this forms an integral part of Cohen's experience of this guilt. It is *his* guilt, i.e. this possessive pronoun indicates not merely ownership as a matter of empirical fact, but rather points at a truth about the contours of Cohen's individual moral agency. It is through this guilt-as-endorsed—and the deepening compassion for human life that it manifests—that Cohen perceives and makes sense of his intersubjective world. It is here, in its anchoring of Cohen's rational agency—and through this anchoring, the emergence of a determinate relation in which he comes to stand to his guilt—that we see how Cohen's guilt constitutes an exercise of his

[61] To be clear: it is our *conception* of the good that volitional necessities embody, and this conception is fallible. We could fail to be responsive to what is genuinely good. When that is the case, then according to Aristotle we fail to exercise our rational agency well. The main point here is that the phrase 'our conception of the good' is value-neutral; this is what permits us to identify the common ground between Aristotle and Frankfurt being discussed here.

[62] The anti-depressant medication which he accepts taking, he seems to accept mainly because it takes the edge off the physical difficulties with sleeping, eating, and sexual functioning. Nowhere in Kleinman's account is Cohen reported as saying he wishes the guilt itself away.

rational agency. This is so in spite of the fact that the guilt is something Cohen is very obviously struck by, i.e. in spite of its being experienced passively.

Frankfurt's work is helpful because it points at this essential characteristic of our rational agency, i.e. that it is configured by our reflective stances of endorsement and disavowal (the most ultimate stance of endorsement being found in our volitional necessities; and conversely, the most categorical stance of disavowal is to be found in those states we feel wholly estranged or alienated from). In having this characteristic pointed out, we are then in a position to see *that* our emotional experiences are situated within this intrapsychic landscape, and that it is where an emotional experience is thus situated (more than our sense of being passively subject to it) that matters in determining how the experience shapes the exercise of our rational agency. Those emotional experiences that are reflectively endorsed wholeheartedly and in an ultimate way—the volitional necessities—lie at the centre of the fabric of each individual's rational agency, defining his or her sense of self *qua* moral, deliberative, and epistemic agent and thereby anchoring the exercise of this agency. But between the volitional necessities and those emotional dispositions from which we feel estranged lies a continuum of variation. Those agents who are practically wise and are able to exercise their rational agency perfectly well (Aristotle's *phronimos*) will manifest an agential structure that is perhaps exhaustively characterized by motivational (and other mental) states that are reflectively endorsed. In rational agents who are less perfectly wise, practically speaking, we'll find—as we scan this fabric—emotional experiences about which an ambivalence is felt, as well as emotional experiences that are reflectively disavowed. Altogether, through their respective positions in this intrapsychic landscape configured by our reflective stances, our emotional experiences help to give our individual rational agency its distinctive shape. Emotional anchoring, manifest most paradigmatically through our volitional necessities, is thus an important mechanism in virtue of which our emotional experiences can be said to constitute an integral part of the exercise of our rational agency.

4.6.4 Feeling Passive and Exercising our Rational Agency

Pulling together the discussion in this section (Section 4.6), we can note the following point. Focusing just on the distressing quality of an emotional experience, in the here and now, is what tends to happen in the clinical encounter. While this narrow focus makes intelligible the psychiatric framing of much

of our emotional distress, i.e. as states we are passively subject to, it threatens to obscure the true nature of our emotional predicament. As the discussion in the first, third, and current chapters show, our emotional experiences have many more facets to them than is usually initially apparent in the clinical encounter. Crucially, in coming to see that there are these further facets, we begin to see that, although we may feel passively subject to our emotional experiences, these experiences can in fact constitute an exercise of our rational agency. They do so firstly in the sense that at least some of our emotional dispositions can be sculpted by the processes of enriched habituation, making them states that we can exercise a degree of control over (albeit gradually) and through which we express our rational perspective (as this is made manifest in our evaluative perception and in our practical deliberation). This point applies to good and poor exercises of our reason alike. The second sense in which our emotional experiences can constitute an exercise of our rational agency is, as we've just seen, by anchoring our agential perspective. This comes about by our emotions being experienced as endorsed or as disavowed. In this way, our emotional lives are intricately interwoven with our sense of who we are *qua* rational agents. Perhaps the deepest problem with applying the psychiatric lens indiscriminately to our emotions is that it tends to obscure this understanding of the ways in which our emotional experiences can play an essential role in constituting our rational agency and, therein, in helping define our sense of self *qua* rational agent.

By identifying two specific mechanisms through which our emotional experiences form an integral part of the exercise of our rational agency—enriched habituation and anchoring—I have sought to show in this section that our exercise of this agency is not restricted in scope to just those states in which we feel active. It can, and does, encompass certain states in which we feel passive—most notably, our emotional experiences. Feeling as though we are passively struck by our emotions, in sum, does not rule out the possibility that our emotional responses can constitute an exercise of our rational agency.

4.7 Conclusion: Conceptions of Flourishing and Negotiating the BED Framework

What has emerged in the foregoing discussion are two radically different ways of thinking about what it means to live well and to flourish. Now that we have both the health-based and the excellence-based conceptions in view, it

is possible to offer the following reply to the defence of the pathologization of distressing emotions presented at the end of the previous chapter. There, we saw the advocate of pathologization arguing that it matters little, ultimately, whether we pathologize a depressive$_{\text{DRIC}}$ experience and respond to it in an interventionist manner, or whether we resist pathologizing it and engage with it hermeneutically, since our ultimate goal *either way* is the one that guides the psychiatric enterprise from the outset: to alleviate the suffering wrought by negatively valenced emotions and to increase the preponderance of positively valenced emotional experiences in our lives. We can say now that this is merely *one* interpretation of the significance of the choice between pathologizing and not pathologizing the depressive$_{\text{DRIC}}$ experience, and that a very different interpretation becomes available in light of a distinct ultimate goal. If we take our ultimate goal not to consist in maximizing happiness as good feeling, but rather to consist in exercising well our epistemic, moral, and deliberative agency, then the positive appraisal of a depressive$_{\text{DRIC}}$ experience along the dimension of existential significance will be prioritized over the negative appraisal in the dimension of valence on the grounds that the depressive$_{\text{DRIC}}$ experience is understood to be integral to fostering the good exercise of our rational agency in virtue of its bearing existential significance. It is valued inherently for this reason, not because it is seen as a means to promoting good feeling. Specifically, in engaging a depressive$_{\text{DRIC}}$ experience hermeneutically, an agent works to uncover her conception of the good, learning to identify it in a more explicit way and thereby becoming more conscious of the misalignment between her decisions and her current daily circumstances, on the one hand, and on the other, that which she most deeply values. This hermeneutic engagement forms an essential part of the agent's work to redirect her life so that she is better able to pursue and to realize her conception of the good, and in so doing, to exercise her distinctly human agency more fully. On this alternative picture, the exercise of our rational agency in a fuller, more complete way is taken to be the ultimate good.

More broadly, once we have this picture of flourishing in view, we are able to grasp the possibility that what we might ultimately want in life is to accurately grasp the true measure of things, to commit ourselves to concerns and projects that are worthy of our love and our energy, to be well disposed towards our fellow human beings—in spite of the fact that living such a life almost invariably involves experiencing distressing emotions alongside pleasant ones. Indeed, on this view, the valence quality of our emotional experiences is much less of a focal point, and certainly no benchmark, of flourishing. Crucially,

when viewed from the Aristotelian perspective on human flourishing, to respond to our negatively valenced emotional experiences principally by seeking to remove them in an interventionist manner is potentially to deny ourselves the opportunity for personal growth tending towards the exercise of our rational agency in finer and more excellent ways.

5
Emotion and the Concepts of Health and Disorder

> It's a very appropriate week to be talking about mental health or, if you prefer an older-fashioned term, 'happiness'.
>
> **Andrew Marr, BBC Radio 4**[1]

5.1 Introduction

We can now return to the question with which this inquiry began—that is, how should the boundary between normal and disordered emotion be delineated?—and to the dispute between Kendler and Horwitz and Wakefield over how to do this. In the Introduction and the first chapter, we've seen how Horwitz and Wakefield seek to constrain the concept of emotional disorder in evolution-theoretic terms. In the next section of this chapter, I will attempt to show that this approach is unviable because it is limited in two important ways, and I suggest that the BED framework, conjoined with the Aristotelian excellence-based conception of flourishing, offers us a more comprehensive—and therein a more promising—approach to delineating the boundary of the concept.

Having done this, I will situate the task of delineating the concept of emotional disorder within the broad cultural current that has been termed 'the therapeutic turn', in which the moral outlook on human life is said to be gradually being replaced by a medical one. Assembling together and drawing upon the conceptual resources made available by the BED analysis (in conjunction with the recognition of the top-down influence that conceptions of flourishing exert over our negotiation of the basic evaluative space), I will suggest

[1] This was part of BBC broadcaster and journalist Andrew Marr's opening remarks to an episode of *Start the Week*, first broadcast on 5 April 2021.

The Concept of Emotional Disorder. Gloria Sibson Ayob, Oxford University Press. © Oxford University Press 2024.
DOI: 10.1093/9780198909620.003.0006

that nothing in this 'therapeutic turn' provides us with sound independent reasons for delineating the concept of emotional disorder in such a way as to subsume under it an ever-greater range of emotional distress.

5.2 Delimiting the Boundary of the Concept of Emotional Disorder

In response to Horwitz and Wakefield's contention that psychiatry over-pathologizes emotional distress, recall that Kendler's response was to say that it is Horwitz and Wakefield who have drawn the boundary in the wrong place, thereby risking 'over-romanticizing' this distress. Having set out the basic evaluative dimension (BED) framework in the first chapter, I suggested in Section 1.8 that this dispute between Kendler and Horwitz and Wakefield is best understood by viewing each party as privileging a different set of basic evaluative dimensions to anchor their application of the (meta-evaluative) concept of emotional disorder. We have seen that Horwitz and Wakefield articulate their position in terms of the concept of biological (dys)function, and that this concept, when applied to our emotional experiences, tracks the dimensions of representational accuracy (a component of emotional rationality) and low-level volitional efficacy (the drive to meet basic needs). Horwitz and Wakefield's resistance to the pathologization of distressing emotional experiences derives from their privileging these two dimensions. Kendler's defence of pathologization, by contrast, turns upon his privileging the dimensions of emotional valence and higher-level agential efficacy (our pursuit of intentional concerns and projects).

Why is Kendler seemingly indifferent to Horwitz and Wakefield's attempt to delimit the scope of the concept of emotional disorder by appealing to the idea of emotional disorder as biological dysfunction—and is Kendler's stance an unreasonable one? I want to suggest that Kendler's indifference implicitly poses the following question, which seems a fair one to ask. It may be true that humans have evolved to experience sadness because it was at some point in our ancestors' history biologically adaptive to be able to feel distressing emotions under certain conditions. But why should we care about that historical fact in the present moment, when there now exists the technological means to immediately alleviate our distress at no great survival/reproductive cost to ourselves? My ancestors may have once benefited greatly from social support in the face of significant personal loss, but what they got through social support I can now get through a selective serotonin reuptake inhibitor. Or again, while my ancestors may not have had the cognitive capacity to *deliberately* act

submissively in the face of a loss of status suffered, thereby needing sadness to protect them from further aggression, I am now able to work this out in a conscious and deliberate way. Why continue to suffer from sadness as a result of lost status, then, instead of alleviating this sadness through pharmacological means, since in any case (i.e. even without experiencing sadness) I know better than to challenge those who have contributed to my loss of status and whom (I can plainly see) I am in no position to wrestle with successfully?[2]

Horwitz and Wakefield reply to this by asserting a normative claim, namely that medicine should only be treating biological dysfunctions, specified in evolution-theoretic terms. Whilst in general terms this sounds like a plausible claim to make—since we want to be able to recognize values other than medical ones (and to do this we need to delimit the scope of the concept of disorder)—in the context of the question just raised, the claim will only be as convincing as the strength of the biological dysfunction account when applied to our emotional experiences. Now, Horwitz and Wakefield's normative claim is relatively straightforward to understand and to apply in the context of physical medicine because the evolution-theoretic framework enables us more or less satisfactorily to explain the structure and the workings of physical organs and mechanisms. But the application of the evolution-theoretic framework to psychopathology is a far more contentious matter—and nowhere is this more evident than in connection with our emotional lives. I will argue that the overriding difficulty that besets the evolution-theoretic framework here is that the framework provides too partial and limited an understanding of our emotional experiences. As a result, relying on this framework to draw the boundary between normal and disordered emotion will mean drawing the boundary in the wrong place, as revealed by the fact that this boundary doesn't allow us to make distinctions between the concept of disorder and other normative concepts that we consider valuable (or at least, the evolution-theoretic framework can only do this at great cost to its own extent and reach). If Horwitz and Wakefield's normative claim cannot be convincingly applied to the domain of human emotions, this will leave Kendler's question unaddressed. I conclude that the BED analysis, combined with the excellence-based conception of flourishing, offers us a more promising way of delineating the concept of emotional disorder and of replying to Kendler's question on this basis.

[2] I take it that it is (in part) this question about the practical value of enduring such suffering that is being pressed in Kendler's charge that Horwitz and Wakefield 'over-romanticize' human suffering.

5.2.1 The First Limitation of the Evolution-Theoretic Framework

We've already seen in Section 1.9.1 one way in which the evolution-theoretic framework is limited: it serves to delineate the boundary of emotional disorder in the way that Horwitz and Wakefield wish it to only when it is restricted to understanding emotional kinds exclusively in terms of the basic survival and reproductive needs that they are hypothesized to serve. Kendler pushes against this by focusing our attention on the fact that we negotiate our emotional lives not just from a concern to get our basic needs met, but also with an eye on pursuing our intentional projects, concerns, and interests. We can broaden this worry out further by noticing more generally that humans realize possibilities that aren't explicable exclusively in evolution-theoretic terms, and that, to a significant extent, we understand and make sense of our emotional experiences through the realization of these distinctly human possibilities. The Aristotelian framework set out in the previous chapter foregrounds this point. Recall that at the foundation of this outlook lies the recognition that whilst humans desire what is pleasant, they also desire what is good. Whilst the desire for pleasantness (and aversion towards the unpleasant) may be capturable within the evolution-theoretic framework, this framework is ill-equipped to accommodate our desiring of the good, because good-directed activities and practices move us beyond evolutionary constraints: in being able to conceive of the good and to pursue it, we act and feel in ways that transcend these constraints, ways that can even put us at a significant biological disadvantage. Crucially, our emotional experiences are shaped by, and in turn find their expression in, these good-directed activities and practices. When we critically appraise our emotional experiences along the basic evaluative dimensions, we are always appraising them as they occur within the context of these practices.

A vivid example of this is the emotional disposition of courageousness in the pursuit of truthfulness or justice. From antiquity to the present day, human history is populated by individuals who display remarkable courage in standing up to harms that can be spelt out only by reference to our capacity for conceiving of (and desiring) the good: from Socrates who accepted his death sentence and drank hemlock in ancient Athens in an effort to defend the ideal of an open society that prizes rational inquiry, to the Russian journalist Anna Politkovskaya who knowingly risked her life (and who was eventually assassinated in 2006) pursuing the truth about human rights abuses

perpetrated during the Russian-Chechen war (1999–2009).[3] Their emotional dispositions—their character—place such individuals at a huge biological disadvantage. In the two examples just cited, the courageousness of both individuals resulted directly in their deaths. Should we therefore judge this courageousness to be dysfunctional?

This isn't an easy question to settle because, within the evolution-theoretic framework, biological functions are determined in evolutionary time, rather than on our observations about what is biologically advantageous or disadvantageous in the present day. So we have to make an imaginative leap into the world of our hunter-gatherer ancestors and picture an emotional trait that is taken to be the evolutionary prototype of Politkovskaya's and Socrates' courage, and to ask of this trait whether it is functional or dysfunctional. We can imagine a tribesperson living in the savannah 100,000 years ago who encounters a dangerous and powerful predator (a lion, say) and who, in spite of having seen the lion in plain sight, rashly stands up to it—seemingly oblivious to the actual danger posed by the lion. The individual in this case seems to display *excessive* boldness—that is, he fails to feel appropriate fear, thereby manifesting a failure accurately to register the threat posed by the predator—and his emotional insensitivity therefore appears to lack biological fitness. On these grounds, we might conclude that this boldness or rashness signifies a malfunctioning of the mechanism of courage. But there is another way to interpret this prototypical boldness in functional terms: the tribesman's standing rashly in the face of the imminent threat of being mauled buys the rest of his tribe valuable time to escape. It therefore promotes genome fitness. On this interpretation, far from signifying a malfunction in the mechanism of courage, this boldness in fact signifies the proper functioning of this mechanism. The difficulty that confronts the evolutionary theorist here is that there is no way of conclusively settling the matter and deciding between these two functional interpretations of courage. So long as the first interpretation cannot be ruled out, and on the assumption that the tribesman's boldness is the evolutionary antecedent of Politkovskaya's courage,[4] the emotional trait of remarkable (or excessive, depending on how you look at it) courage could signify a dysfunction.

[3] Politkovskaya entitled what was to be her final collection of dispatches, 'Is Journalism Worth Dying For?' (2011). The collection contains an article about torture she was working on—which she did not get to complete writing—that she discussed publicly on Radio Free Europe/Radio Liberty on 5 October 2006, two days before she was murdered. This is mentioned by Viv Young, who reviews Politkovskaya's posthumous collection in the New York Journal of Books.

[4] An assumption I find problematic. I will say more about this analogue in n. 10 below.

It is important to understand the full significance of this claim, and so I will spend the remainder of this section unpacking it further. I want to begin by noting the following commonplace: we recognize in everyday life myriad ways of normatively appraising an emotional response as appropriate or inappropriate. Some of these ways crucially turn upon *not* viewing the response as pathological. That is to say, when we judge an emotional response that is inappropriate as an unreasonable or an irrational one, we might indeed be judging it to be disordered, or we might instead be critically appraising it as a poor exercise of epistemic, deliberative, and/or moral agency (on the part of the agent whose response it is). Thus an irrational emotional response might be criticized for the poor judgement it manifests on the agent's part—indeed, this criticism might be viewed as a moral failing if it is rooted in a broader assessment of the agent's character—and in criticizing an irrational emotional response in these ways, we are holding its agent accountable for the poor exercise of his or her epistemic, deliberative, and/or moral agency. These judgements are diametrically opposed to judging an irrational emotional response as disordered insofar as we do not hold the agent of the latter response accountable for the irrationality she manifests.[5] Generalizing this point (and recalling the discussion in Section 1.3), it seems that there are two distinct kinds of judgement of emotional irrationality that we make:

- $\text{Judgement}_{\text{pathology}}$ (Jp). An emotional response is appraised as irrational, and its irrationality is explained by reference to the concept of disorder (naturalists such as Horwitz and Wakefield will flesh this out by saying that the agent responds emotionally irrationally to her circumstances due to her suffering from some (hypothesized) emotional dysfunction), and
- $\text{Judgement}_{\text{agential}}$ (Ja). An emotional response is appraised as rational or irrational, and this is seen as an instance of the agent's exercising her rational agency (respectively) well or poorly.[6]

[5] The suggestion that we can be held accountable for our emotional responses merits clarification. Theorists who hold that our emotional responses are suited to being rationally critiqued—and are revisable in light of such critique—accept that this critique (implicitly) makes reference to a temporally extended, multifaceted process that involves far more than a mere appeal to the relevant evidence that was available to the subject at the given time. Habituation, the cultivation of good human relationships (including relationships with good role models), and the development of enlightened (non-deceptive and morally praiseworthy) self-understanding are amongst some of the complex and lengthy processes involved in the revision of our emotional dispositions. This makes the rational critique of our emotional responses quite unlike the rational critique and revision of our beliefs. But like the latter, this view of our emotional responses, i.e. as being apt for rational criticism, does indeed take our emotional experiences to be partially constitutive of the exercise of our epistemic, moral, and deliberative agency. In the previous chapter, we looked at some length at one way in which this view has been articulated, i.e. by Aristotle and neo-Aristotelian commentators.

[6] These two appraisals correspond, respectively, to the pathological construal and agential construal of affective irrationality, first introduced in Chapter 1. The Aristotelian account of human emotion presented

Using the terminology introduced in the first chapter, in the first type of appraisal, we construe instances of emotional irrationality in *pathological* terms; in appraisals of the second type, we construe instances of emotional irrationality in *agential* terms. For brevity, I shall henceforth refer to the former kind of appraisal as (Jp)-type rationality appraisals, and to the latter kind as (Ja)-type appraisals. Now, by locating the source of (Jp)-type irrationality in an emotional dysfunction specifically, the evolution-theoretic framework appears to promise us a way of marking off (Jp)-type rationality judgements from (Ja)-type ones, as follows: only those irrational emotional responses that are caused by an emotional dysfunction instantiate (Jp)-type irrationality; all other irrational emotional responses instantiate (Ja)-type irrationality. Notice that this promise will be fulfilled only if the notion of emotional dysfunction can itself be delimited in such a way as to enable us to sustain a meaningful contrast between (Jp)-type and (Ja)-type irrationality judgements—and this is where the evolutionary theorist is confronted by a problematic dilemma, as I shall now argue.

Turning first to the first horn of this dilemma, suppose that the trait of remarkable courage (displayed by Politkovskaya and by Socrates) turns out to be an emotional dysfunction by evolution-theoretic lights, as per the first of the two functional interpretations considered earlier. Then, if the concept of emotional disorder is analysed in terms of the notion of emotional dysfunction (as per the evolution-theoretic framework), the evolutionary theorist is

in Chapter 4 helps us understand more fully the agential construal (although it should be noted that this construal isn't by any means exclusive to the Aristotelian account). Viewing an inappropriate emotional response according to one or other of these two construals has far-reaching practical implications. To illustrate: take someone who is chronically fearful, and who may thus be diagnosed with an anxiety disorder within a clinical setting. Viewing this excessively fearful disposition as constituting a poor exercise of one's epistemic and moral agency, rather than as pathological, opens one to otherwise unconsidered possibilities: (i) taking it on trust (from a better judge) that one's fear responses tend to be disproportionate to the circumstances—or altogether unmerited by them—and becoming curious about why and how a better judge sees things so differently; and (ii) trying to re-habituate one's fear responses based on this calibration. Besides the focus on evaluative perception or discernment here, this re-habituation may also involve engaging in practices that work to increase one's feeling of safety in the world. It involves also (iii) becoming curious about the fearful disposition itself, i.e. its origins and its ramifications in one's life. Importantly, underlying the realization of these possibilities is the recognition of the further possibility of reason-informed revision, of growing to be someone who see things in a less fearsome light. One exercises one's agency both in the cultivation of a more rational fear disposition and in the evaluative perceptions that are transformed by this—from inaccurately seeing the world as (pervasively) threatening previously, one is now able to discern actual threats more accurately. Aristotle describes the disposition towards excessive anger and fear as *incontinent* anger and incontinent fear (not as *pathological* anger or fear), capturing precisely this point. Judging the original fearful disposition to be an anxiety disorder masks these possibilities, or at least interprets them through an interventionist filter, thereby potentially missing the distinctively agential character of this revision in a person's fear-related dispositions.

led to conclude that Socratic courage is disordered.[7] But given the distinct grammar of the concept of disorder (spelt out most sharply in the strict construal of the concept of pathology, discussed in Chapter 3), this threatens to obliterate the (Ja)-type judgement that we do in fact make of such courage, namely that it is deeply admirable and commendable.[8] Quite simply, this is because the judgement of disorderedness has a monopolizing impact on the all-out assessment we make of Socratic courage: whenever this judgement is made, it places the emotional experience being appraised outside the bounds of what we can be held accountable for, thereby obliterating all other normative appraisals that presuppose agential accountability which we might ordinarily make of the very emotional experience being so judged (including the judgement that the courage demonstrated by Politkovskaya or by Socrates is admirable and praiseworthy). Whenever a judgement of disorderedness is made, the grammar of the concept of disorder entails that this judgement becomes the only lens through which we comprehensively appraise our emotional experiences, and this results in an extremely narrow and impoverished view of our emotional lives. This is the first horn of the dilemma that besets the evolution-theoretic analysis of the concept of emotional disorder.

To avoid this problematic conclusion, and to dissipate the broader worry that the evolution-theoretic concept of emotional disorder threatens to obliterate the other normative appraisals we make of our emotional states, the evolutionary theorist might seek to isolate a purely biological component within our emotional states, and to insist that the concept of emotional (dys)function applies to just this component, thereby leaving the remainder of an emotional state open to being normatively appraised in all the non-health-related ways that we ordinarily assess it by. That is, the hope is that by isolating a purely biological component in our emotional states, the evolutionary theorist will be in a position to delimit the notion of emotional dysfunction in such a way that makes it suitable for marking off (Jp)-type rationality judgements from (Ja)-type ones, even *within* the assessment of a single token emotional response. In effect, this delivers a composite picture according to which an emotional response may be judged (narrowly, in biological terms) to be dysfunctional, whilst simultaneously being judged (broadly, in moral terms) to

[7] In the case of Socratic courage, in which the subject is put in grave danger for being courageous, the emotional response would also be harmful to the subject. Thus this emotional response would satisfy both the harm condition and the biological dysfunction condition of Wakefield's analysis of the concept of disorder, thereby counting not only as dysfunctional but also as disordered.

[8] The fact that we commend such courage in the pursuit of what we value is reflected, for instance, in the numerous awards conferred on Politkovskaya for her brave and valuable reporting; these include awards that have been established in her name—for instance, the Anna Politkovskaya-Arman Soldin Prize for Courage in Journalism, awarded by the French government annually since 2013.

be praiseworthy. The evolutionary theorist's success in carving out a purely biological component in our emotional states will thus be decisive in ensuring that the evolution-theoretic conception of emotional disorder does not result in an impoverished and unduly narrow view of our emotional lives. So the overall plausibility of the evolutionary analysis of the concept of emotional disorder will depend on whether or not it is possible to carve out a purely biological component in our emotional states.

How then might this putative biological component be isolated? I will now sketch out what I think is the most obvious way of doing this (and this seems to echo much of what Horwitz and Wakefield say), but to anticipate, I think that this move is ultimately unviable. In order to prevent the collapse of (Ja)-type rationality judgements into (Jp)-type ones, the evolutionary theorist could try to isolate a purely biological component in emotional states by drawing a distinction between two levels at which human nature is manifest—namely, first nature and second nature—and to identify the biological with just first nature. On this picture, our second nature is constituted and defined by our capacity to conceive of and to pursue the good. The exercise of our epistemic, moral, and deliberative agency (including through our emotional responses) is apt for normative appraisal by reference to rational and ethical standards, and both the exercise of our rational agency and our normative appraisals thereof are taken to be constitutive of our second nature. All else is part of our first nature. That is, the fabric of our first nature is taken to consist in the instincts, drives, desires, proto-beliefs, and primitive aspects of human emotional traits that function to support and enhance human biological fitness in various ecological niches over evolutionary timescales. The concept of emotional (dys) function, according to this proposal, would be characterized exclusively by reference to this first nature alone.

Given this characterization of the concept of emotional dysfunction, the evolutionary theorist is now able to say that the concept of emotional disorder simply isn't applicable to (the entirety of) the emotional trait of courageousness as manifest by Politkovskaya since, thus manifest, it is an expression of Politkovskaya's second nature, being rooted in her pursuit of truthfulness and justice (these are values that become available to an agent only by her being able to conceive of the good). This leaves it open to us to normatively appraise Politkovskaya's courage as admirable and praiseworthy, even if this courage is deemed to be dysfunctional, narrowly speaking (i.e. in strict biological terms).

It is perhaps worth pointing out here that the two sets of judgements needn't always be in conflict with each other. To see this, consider Winthrop Cohen's guilt. Cohen's guilt may be biologically adaptive in virtue of guilt's being an

enabling condition for the particular ways by which humans socially cooperate with each other. This is an observation about human first nature, and thus it is germane to the concept of emotional (dys)function. But Cohen's guilt is *also* a clear-sighted seeing of a grievous moral harm that he committed. This would be an observation about human second nature: the guilt Cohen experiences is constitutive of his grasping this moral truth, and on these grounds, we positively appraise Cohen's guilt along the basic evaluative dimensions of emotional rationality and interpersonal functioning (he demonstrates good judgement), as well as along the dimension of existential significance (his guilt manifests a deep and serious compassion for human life as an element of his mature conception of the good). Cohen's guilt, construed as a manifestation of clear-sighted moral discernment, is an expression of his second nature and as such the concept of emotional (dys)function would be inapplicable to it. This leaves it open for us to positively appraise Cohen's guilt in terms of (Ja)-type judgements—in this case along the evaluative dimensions of emotional rationality, interpersonal functioning, and existential significance—alongside the appraisal of this guilt in terms of the (Jp)-type judgement, namely that it signifies the proper functioning of his pro-social emotional mechanisms.

In this way, this strategy for isolating a purely biological component in our emotional experiences promises to rescue the prima facie plausibility of the evolution-theoretic analysis of the concept of emotional disorder in the face of the threat of its collapsing all (Ja)-type rationality judgements into (Jp)-type ones (in instances where the two sets of judgements come into conflict with each other). But this strategy comes at a heavy price for, by pegging the notion of emotional (dys)function to our first nature exclusively, the resulting evolution-theoretic analysis of the concept of emotional disorder appears to sever the promised explanatory connection between emotional dysfunction and (Jp)-type irrationality judgements, i.e. a connection that presumes that all and only those irrationality judgements rooted in an ascription of emotional dysfunction are (Jp)-type judgements. This is because it is possible on this view for an emotional response to be judged (widely) as rational, for instance as admirable and praiseworthy, in spite of this very same token response being deemed (narrowly) to signify an emotional dysfunction.

Just as problematically, this composite view—which restricts the domain of emotional dysfunction (and thus of disorder) to just our first nature—threatens to exclude from view those aspects of our emotional states that are expressive of our second nature that *intuitively* bear upon the judgement of emotional disorder. Depressive$_{\text{DRIC}}$ experiences furnish us with an example of this. As it was spelt out in Chapter 3, it is possible for a depressive$_{\text{DRIC}}$ experience to be the expression of a serious misalignment between an agent's

outward circumstances and her most deeply held values. To put the point in Aristotelian terms (as we did in the previous chapter), the experience is an expression of an agent's frustrated attempt to pursue the good as she conceives of it (however inchoately), or a failure to pursue it altogether. Depressive$_{DRIC}$ experiences, thus construed, are thoroughly embedded in our second nature. Yet there is at least one normative appraisal of such experiences that bears directly on the judgement of disorderedness, i.e. they may be positively appraised in the basic evaluative dimension of existential significance (as we've seen when discussing James Hollis' interpretation of depressive$_{DRIC}$ experiences). This positive appraisal gives us grounds for resisting the (all-out) pathologization of depressive$_{DRIC}$ experiences. But appraising any emotional experience in terms of its existential significance is, of course, a manifestation *par excellence* of our second nature. So in restricting the scope of the concept of emotional dysfunction to just those emotional states and traits that manifest our first nature, the evolution-theoretic framework will fail to register this limit on the concept of emotional disorder, let alone to provide an account of it. Indeed, for all that Horwitz and Wakefield do say explicitly about the function of sadness, a depressive$_{DRIC}$ experience will end up counting as disordered because there is no obvious or apparent loss to which the depressive$_{DRIC}$ experience would be seen as an appropriate response (and sadness in response to loss is the only form of non-disordered or normal sadness that Horwitz and Wakefield discuss explicitly).

This shortcoming is also evident in the dispute between Horwitz and Wakefield and Kendler that we looked at in the Introduction and the first chapter. Against Horwitz and Wakefield's insistence that sadness is not disordered if it is a response to loss, Kendler draws our attention to his patients' intentional level concerns, concerns that look to have at least *prima facie* relevance to the concept of emotional disorder. Specifically, Kendler insists that the immobilizing effect of intense sadness on his patients' lives gives us grounds for judging the sadness to be disordered (in saying this, we have seen that Kendler prioritizes the negative appraisals made along the basic evaluative dimensions of valence and higher-level agential efficacy). In light of the basic evaluative dimension (BED) analytical framework set out in the first chapter, we were in a position to see that Horwitz and Wakefield's evolution-theoretic position does not give us conclusive grounds for dismissing Kendler's considerations. All we have are two parties to the dispute, each of whom prioritizes appraisals in a different set of basic evaluative dimensions, and neither of whom offer us reasons for favouring their prioritization over the other party's prioritization (this was the upshot of the discussion in Section 1.9.1). But

now we can add, further, that the evolution-theoretic concept of emotional dysfunction—if restricted to just the first nature components of our emotional traits—seems in fact to be altogether blind to the considerations that Kendler brings into the fold, i.e. considerations that refer to emotional states as expressed in our second nature.

The best that the evolutionary theorist can hope to do—in order to encompass our emotional traits as they are lived out through our second nature, whilst restricting the concept of emotional dysfunction to just our first nature—is to place the second nature expression of such traits on the evolution-theoretic radar by construing them as spandrels, that is, as accidental by-products of the selection of the first nature component of these emotional traits. To do this, however, would still fall far short of giving us an understanding of what it means to appraise a depressive$_{\text{DRIC}}$ experience as existentially significant, or to appraise remarkable courage in the face of a grave threat to a person's life as admirable and morally praiseworthy. In order to grasp the content of these appraisals, we have to bring into view our desire for the good. It is only by recognizing that humans are capable of conceiving of and desiring the good that the positive appraisal of death-defying courage for the sake of upholding that which we value most becomes intelligible as such. And since these latter appraisals directly constrain and limit our intuitive application of the concept of emotional disorder, any adequate characterization of this concept must surely equip us to grasp the content of these appraisals. To conclude, because our grasp of these appraisals draws upon conceptual resources that require us to go beyond the modified version of the notion of emotional dysfunction currently being considered (according to which emotional (dys)function is pegged exclusively to our first nature), the evolution-theoretic analysis of emotional disorder as emotional dysfunction looks to be too narrow to yield an adequate understanding of this concept as it is situated in the nexus of evaluative concepts we ordinarily use to make sense of and to appraise our emotional experiences.

To be clear, the problem here is not so much that the judgement of emotional dysfunction, made of the first nature component of an emotional response, is cancelled out by the judgement that the response is praiseworthy (there is no cancelling out since the picture of our emotional responses, according to the strategy being considered now, is a composite one). The problem rather is that the resulting evolution-theoretic account does not have the means to account for the content of an appraisal of the emotional response as existentially significant, or as morally praiseworthy. What is problematic about this is that such appraisals do bear upon our *intuitive* sense that the emotional response is not disordered. This latter datum will simply have to be left unexplained by the

evolutionary theorist. Seeking to isolate a purely biological component in our emotional states by pegging it exclusively to human first nature thus entails the following consequence: we are offered a composite picture according to which it is possible for an emotional response to be judged (narrowly, in biological terms) as dysfunctional, whilst simultaneously being judged (broadly, in moral terms) as praiseworthy, but there is no acknowledgement within this picture of the fact that the latter judgement would *ordinarily* be taken as a reason to resist the pathologization of the emotional response.[9] This leaves a large hole in the analysis presented by the evolutionary theorist of the concept of emotional disorder. This hole, moreover, is a puzzling one: narrowly construed, the emotional response in question is dysfunctional—and because it is biologically harmful to its subject (at least in the case of Socratic courage), it counts as disordered. But construed broadly, its existential significance or its moral praiseworthiness seems to give us good reason not to pathologize the emotional response. Now, since the evolution-theoretic approach is silent with respect to the latter judgement, we're left staring at an apparent aporia, one there seems to be no way of resolving within the constraints of the evolutionary framework.

The evolution-theoretic concept of emotional disorder is thus beset by a problematic (I think, fatal) dilemma: on the one hand, within the evolution-theoretic framework, (Ja)-type rationality judgements threaten to collapse into (Jp)-type rationality judgements whenever an emotional response or disposition is judged to be disordered by evolution-theoretic lights. That is, when this judgement is made, it threatens to obliterate all the other normative appraisals we make of that very same emotional experience. The resulting view of our emotional lives will be, at the very least, an extremely narrow and impoverished one. If all we can say of Socratic courage is that it is emotionally dysfunctional, then we have failed spectacularly to grasp the possibilities that human life affords. In order to avoid this difficulty, the evolutionary theorist might try to isolate a purely biological component in our emotional experiences and to restrict the application of the concept of emotional dysfunction—in terms of which the concept of emotional disorder is analysed—to just this component. This would leave the rest of a given emotional experience open to being normatively appraised in non-health-related ways. I have suggested that a way of doing this would be to identify the biological component in our

[9] Or conversely, in the case of Kendler's example, a composite picture according to which it is possible for an emotional response to be judged (narrowly, in biological terms) as functioning properly, whilst simultaneously being judged (broadly, in intentional terms) negatively along the dimension of agential efficacy, but there is no acknowledgement within this picture of the fact that the latter judgement would *ordinarily* be taken as a reason to pathologize the emotional response.

emotional experiences with just those ways by which our emotional experiences are manifest in our first nature. But in restricting the concept of emotional dysfunction thus, the evolutionary theorist thereby excludes from view the myriad ways by which our emotions find expression through our second nature, many of which directly bear upon and constrain the judgements of emotional disorder that we ordinarily make. So once again the resulting concept of emotional disorder threatens to be too narrow, this time in the sense that it fails to offer us a sufficiently comprehensive understanding of the relationship in which this concept stands to its relevant complements.[10]

Faced with this dilemma, the evolutionary theorist might seek to challenge the dichotomy between (Jp)-type and (Ja)-type irrationality judgements. According to this line of reply, the suggestion would be that it is possible for an irrational emotional response to be both disordered *and* one for which it is appropriate to hold its subject accountable. However, so long as we find it important to preserve the following possibility, this reply merely postpones the evolutionary theorist's problems: sometimes, in saying of an irrational emotional response that it is disordered, we mean precisely to say that its subject cannot (coherently) be held accountable for it. This is one way that we do in fact appraise the rationality of our emotional experiences. So what would distinguish this kind of judgement of irrationality, from a judgement of irrationality that purports to view the emotional response as both disordered and one for which we can coherently hold its subject accountable? At this point, we begin to lose our grip on what we're speaking about because when a judgement of irrationality is simultaneously a judgement of disorderedness (thus, an (Jp)-type judgement), it is typically a judgement to the effect that the agent's irrational emotional response signifies or is caused by a pathological process, *rather than* signifying a poor exercise of his epistemic and/or moral agency for which an agent can be held accountable. Being thus caused, the appropriate way to seek to overcome this irrationality is by treating the pathology, rather than by the agent's working to cultivate a more rational emotional disposition. It is precisely this contrast that the concept of disorder is typically used to signal.

[10] The fact that Politkovskaya's courage is anchored in her pursuit of the good makes it quite unlike our ancestral tribesman who feels insufficient fear and is thus insensitive to the imminent threat of being mauled. Of the latter, we want to say either that he has miscalculated the degree of danger he is in or his own strength (or both). Neither of these remarks presupposes reference to an agent's conception of the good. This difference problematizes the assumption that the tribesman's boldness is the evolutionary antecedent of Politkovskaya's courage, and so the evolution-theoretic framework may have nothing at all to say about the latter. If so, then the first horn of the dilemma being described in this section will not apply to the evolution-theoretic concept of emotional dysfunction, leaving the concept firmly impaled on the second horn.

In other words, once we remove this particular way of contrasting the two kinds of judgement of irrationality—so that an irrational emotional response judged as disordered can nonetheless also be a response for which we can hold its subject accountable—we pull the rug from under us altogether. There is now no longer a meaningful distinction to be drawn here at all between the two kinds of judgement of irrationality. This would mean either that (i) all judgements of emotional irrationality are taken to signify a poor exercise of his epistemic and/or moral agency for which an agent can be held accountable, in effect rendering the concept of emotional disorder vacuous, and therefore redundant (so we no longer have any use for the evolution-theoretic concept of emotional dysfunction), or else (ii) the judgement of disorderedness has to be understood as swallowing up all judgements of irrationality, so that no judgement of emotional irrationality can be construed as a critique of its subject's exercise of her epistemic, deliberative, and/or moral agency, something for which she can be held accountable. At this point, we come face to face once again with the threat of all other normative appraisals of our emotional experiences being obliterated by the judgement of disorderedness, whenever this judgement is made. Challenging the dichotomy between (Jp)-type and (Ja)-type irrationality judgements thus doesn't seem like a promising way out of the dilemma.

It is worth noticing that this dilemma does not arise for the evolution-theoretic analysis of the concept of physical disorder because we generally tend not appraise the workings of our physical organs (the heart, say) in non-health-related ways; and we certainly do not appraise their functioning in terms of the normative standards that refer to the concepts of rationality or ethical goodness. To restate the point, there is no 'second nature' displayed by human lungs or the human heart, i.e. the behaviour of these organs is indifferent to our conception(s) of the good. This contrast between the concept of physical disorder and that of emotional disorder helps to underscore the problematic nature of the concept of emotional dysfunction (construed in evolution-theoretic terms). The problem, in a nutshell, can be stated as follows: the development of human second nature sees the emergence of the normative appraisals of our lives in terms of rationality, ethical standards, and questions of existential significance (our receptivity to these normative standards are constitutive of our second nature), and our emotional lives find expression through our second nature, so making them subject to these appraisals. Now, it is unclear to what extent, if at all, our second nature as it is presently formed was already present in evolutionary time—but it is important to recognize the momentous character of the shift from a form of

life exhaustively shaped by first nature to one that manifests a sophisticated second nature. The Aristotelian framework set out in the previous chapter helps us to see that our emotional lives today have been thoroughly transformed by this shift.[11] It is this transformation in our emotional lives that explains why the evolution-theoretic biological dysfunction conception of disorder or pathology cannot be extended without great difficulty from the physical domain to the domain of human emotion. The evolution-theoretic framework looks most plausible when it is used just to explain the workings organs and mechanisms that haven't undergone a transformation from first nature to second nature (i.e. our physical organs). The further away we move from this starting-point, the shakier the framework becomes.[12]

Summarizing: given that our task is to decide in the present day whether or not the concept of emotional disorder can be rightly applied to this or that emotional response, the lens that the evolution-theoretic framework imposes upon us, through which we view the appraisals we make of our emotional experiences, looks to be far too narrow to yield a viable conception of emotional disorder. The narrowness of the evolutionary-theoretic framework—as applied to our emotional experiences with the aim of delineating the concept of emotional disorder—seriously problematizes Horwitz and Wakefield's normative claim that psychiatry should only be treating emotional dysfunction (where 'dysfunction' is construed in evolution-theoretic terms). It is hard to stand by the evolution-theoretic concept of emotional disorder, anchored in the notion of emotional dysfunction, when the understanding it proffers of the appraisals we make of our emotions (appraisals that bear directly on the judgement of emotional disorder) is so partial. This, I hold, undermines the attempt to delimit the scope of application of the concept of emotional disorder by appealing specifically to the evolution-theoretic idea of emotional dysfunction. While the claim that medicine should only be treating biological dysfunctions may therefore be an important and useful one in the domain of physical health, when applied to the domain of human emotion, its usefulness evaporates. The appearance of usefulness is sustained by the chimera of a purely biological component of emotional experience that is isolatable from all other facets of our emotional lives.[13]

[11] For a neo-Aristotelian articulation of this distinction between first nature and second nature in contemporary philosophy, see McDowell 1998, esp. pp. 184–92.

[12] This is incidentally why I am less sanguine than the evolutionary theorist, Randolph Nesse, about the prospects of rapid advancement in the study of human emotions by rooting this study in evolutionary thinking, merely on the basis that it is this theoretical perspective that 'has rapidly advanced the study of animal behavior' (Nesse 2007: 160). An explanatory schema that has worked well for organisms who don't manifest a second nature will not necessarily work as well for organisms who do.

[13] A more localized formulation of this problem—using the example of intense, immobilizing sadness (IIS) construed as a biological function—was presented in Section 1.9.2. It is perhaps worth remarking here

Where does this leave us with regard to Kendler's (covert) question, 'Why should we care in the present moment about the historical fact that emotional distress under certain conditions was biologically adaptive for our ancestors, when there now exists the technological means to immediately alleviate our distress at no great survival/reproductive cost to ourselves?' Horwitz and Wakefield's answer, 'Because medicine should only be treating emotional dysfunctions', is germane and persuasive only if the concept of emotional dysfunction on which it rests is itself sound, that is, if it doesn't distort or altogether obscure the myriad normative appraisals we make of our emotional experiences, appraisals that directly shape our judgement of normal/disordered emotion. If it isn't sound, then what? Then Kendler's preferred liberal conception of emotional disorder remains untouched by Horwitz and Wakefield's appeal to the evolution-theoretic concept of emotional dysfunction, and we shall need to look elsewhere to rein in the scope of emotional disorder and thereby to continue to recognize the possibility that distressing or incapacitating emotional experiences aren't necessarily pathological. In virtue of its greater comprehensiveness, I suggest that the BED analysis combined with the Aristotelian framework (as set out in the previous chapter) offers us a more promising way of doing this.

5.2.2 The Second Limitation of the Evolution-Theoretic Framework

There is a second, more general, way in which Horwitz and Wakefield's evolutionary-theoretic approach is limited, which renders it inert to the question of why it matters that we seek to delimit the scope of the concept of emotional pathology at all. And there is correspondingly a way in which the Aristotelian framework constitutes a more comprehensive outlook, a comprehensiveness that therein furnishes us with an answer to the question of why it matters. This makes the latter approach a more attractive way of delineating the boundary of the concept of emotional disorder.

Let me begin here by noting that it is not clear what anchors Horwitz and Wakefield's insistence that we delimit the concept of emotional disorder in

that the systems-theoretic conception of emotional dysfunction will likely be beset by similar difficulties, because any framework that accords an explanatory primacy to biological functioning will restrict its own explanatory resources in such a way that leaves it unable to deal with second nature traits and dispositions. In developing a second nature, humans become capable of engaging in complex practices that transcend biological imperatives (survival and reproduction), and crucially for our present inquiry, much of our emotional life is understood and appraised within the context of these practices.

the way they propose, i.e. that psychiatric medicine should only be treating biological dysfunctions, construed in evolution-theoretic terms. We might agree with Horwitz and Wakefield that it is a good idea to delimit the concept of emotional disorder, sharing with them the background thought that it is only in this way that we could hope to hold on to value pluralism and prevent our lives from being consumed entirely by medical values. But Horwitz and Wakefield do not offer us a positive account of what this value pluralism yields; they observe simply that distressing emotions (when functioning normally) are an inherent and valuable part of the human condition. Interestingly, Horwitz and Wakefield do hint at the basic evaluative dimension of existential significance in offering this observation. They write:

> One argument against medicating normal sadness is that it treats as pathological what is actually an inherent and valuable part of the human condition. For thousands of years, people have used religion, spirituality, and philosophy to understand how their unhappiness is tied into larger questions about life. Such questioning allows people to comprehend how their emotions are related to basic aspects of human existence and to gain a deeper appreciation of their feelings than palliation through medication can provide . . . . Using pills represents an escape from truly confronting life's problems. (Horwitz and Wakefield 2007: 190)

Without an account of why it matters that we use 'religion, spirituality, and philosophy to understand how . . . unhappiness is tied into larger questions about life,' however, it isn't clear why we should choose this option (and limit the scope of the concept of emotional disorder accordingly), rather than go down the route Kendler proposes. Kendler's emphasis on emotional valence and higher-level agential efficacy, moreover, has the obvious advantage of being practically expedient, addressing people's immediate concerns about their emotional distress when they enter the consulting room. If both these options are set against the health-based conception of flourishing, then the question raised at the start of Chapter 4 becomes germane once again: if human flourishing consists in enjoying a preponderance of positively valenced emotional experiences (and a relative dearth of negatively valenced ones), then why does it matter, ultimately, whether we achieve this through medication or through questioning and reflective self-examination? The normative claim that medicine should only be treating biological dysfunctions does not stretch as far as to give us an answer to this question, and that is because it says nothing about the nature of human flourishing (nor is it intended to).

By contrast, the Aristotelian framework is first and foremost an account of human flourishing, so we'd expect to find an answer to this question written

into the very fabric of the framework—and indeed, the framework offers us precisely this, as the discussion in Chapter 4 has established. Summarizing this answer in terms of the BED analysis, we can say that the Aristotelian framework provides us with a reason for privileging the BEDs of emotional rationality, existential significance, interpersonal functioning, and agential efficacy *over* the BED of emotional valence (as psychiatry does). The reason is that human flourishing on the Aristotelian view consists in the excellent exercise of our epistemic, deliberative, and moral agency (rather than in enjoying a preponderance of good feeling), and that an essential part of exercising this agency well is the experiencing, under relevant circumstances, of negatively valenced emotions that are positively appraised along the former basic evaluative dimensions. This offers us a reason to care about delimiting the scope of the concept of emotional pathology in such a way that it isn't expanded to include all negatively valenced emotional experiences.[14]

[14] I wish to pause here to address what is perhaps the most basic motive for committing to naturalism, namely the idea (attractive to naturalists) that it is *intrinsically* worth demarcating the boundary of pathology using a value-neutral criterion. If we take this idea as our starting point, it might appear that far from being a shortcoming, it is in fact a strength of the naturalistic approach to understanding disorder that it does not engage our concerns about flourishing.

This idea, it should be noted, is not a free-standing one: its coherence turns upon the tenability of a distinction that is central to the naturalist's outlook, namely the distinction between the theoretical study of disease or pathology and the practical uses of the findings of this study in the field of medicine. While the practical pursuit of medical care may be shaped by considerations of human values, the theoretical study of pathology should, according to this perspective, be a value-independent scientific undertaking (for a programmatic assertion of this claim, see Boorse 1975). This distinction between the theoretical and the practical concepts of health, if workable, would be what enables the naturalist to delineate the boundary of pathology using the concept of biological function, even in the case of human emotions, unhindered by any consideration about flourishing (or other human values).

Now, our grasp of this distinction turns upon our understanding of the way the two concepts herein are related to each other, since the naturalist contends that our practical uses of these concepts must be informed by their theoretical analogues. In the case of physical health, we do have available a reasonably straightforward picture of this relationship, which runs something like this: our biological survival is a causal precondition of our flourishing, and thus our physical health is instrumentally valuable for our flourishing. Importantly, the relationship looks to be exclusively an instrumental one: that our physical health causally enables us to flourish seems to be the principal reason why we value being physically healthy. This instrumental relationship makes it possible for us to insulate the task of demarcating the boundary of physical pathology from our concerns about flourishing, and this is what sustains the postulated separation between the theoretical and the practical concepts of health. How so? Broadly, with regard to the theoretical study of pathological entities, this insulation allows us to take as the apical goal in terms of which physical mechanisms are individuated and identified the goal of biological survival (and reproduction). If flourishing is the apical goal in terms of which we make sense of our lives globally, then this apical goal is extraneous to the task of identifying biological functions by reference to the apical goal of biological survival since the functions thus identified are understood to be causal enabling conditions of our flourishing, rather than figuring constitutively as elements in our flourishing.

Here, then, we have two distinct reference points in terms of which to articulate the separateness of the theoretical concept of health and the practical concept: whilst the theoretical concept has as its subject of investigation biological functions as identified by reference to the apical goal of survival, the practical concept is shaped by our concern to flourish—and these points of reference are external to each other, i.e. the one (good biological functioning) picks out a set of causal conditions that enables the other to be realized (human flourishing). This picture is paradigmatic of our understanding of how the two concepts of theoretical pathology and practical pathology are related to each other.

5.2.3 Whither to, Emotional Disorder?

Without recourse to the notion of emotional dysfunction, Horwitz and Wakefield may press the question: how according to the BED analysis—coupled with the Aristotelian excellence-based conception of flourishing—is the concept of emotional disorder to be characterized and its scope delimited? I want to begin by summarizing what is being ruled out. The foregoing discussion has sought to show that it is not possible to isolate a purely biological component in our emotional experiences and to apply the judgement of normalcy/disorderedness to that component alone. What the BED analysis, combined with the excellence-based conception of flourishing, reveals is that non-biological appraisals of emotional experiences—that is, appraisals not pertaining to an organism's survival or reproductive viability—also directly shape the determination of an emotional experience as disordered or otherwise.

This brings us to the heart of the matter. The concept of emotional disorder is an inherently meta-evaluative one, which renders it opaque (cf. Chapters 1–2). In this way, the concept is very unlike more transparent concepts of disorder such as the concept of physical disorder, and possibly certain other concepts of mental disorder (e.g. psychosis). The judgement that an emotional experience, response, or disposition is disordered is the upshot of a complex negotiation between at least five basic evaluative dimensions, a negotiation that is often influenced from the top down by our conceptions of human flourishing. Presuming the health-based conception of flourishing will give us grounds for weighting more heavily the negative appraisal of an emotional experience in the BED of affective valence, whilst adopting the excellence-based

By contrast, as we've seen in the previous section, i.e. Section 5.4.1 (and of course in the first chapter), we make both biological appraisals of our emotional experiences (i.e. appraisals pertaining to the survival/reproduction advantage conferred by these experiences) and non-biological appraisals (i.e. appraisals that do not pertain to survival/reproduction), all of which constrain our judgement of emotional disorder. This is reflective of the broader point that considerations of human flourishing—and not just reference to biological survival and reproduction—bear directly upon the delineation of the boundary of emotional pathology, which entails that the apical goal in terms of which the boundary between pathology and non-pathology is delineated is not exclusively that of survival/reproduction. Indeed, we've seen in some detail, in Chapters 3 and 4, how considerations of flourishing directly inform our delineation of this boundary. Thus with regard to the domain of human emotions, we look not to be able to sustain two distinct reference points in terms of which to construe the separation of, and relationship between, the theoretical and practical concepts of pathology. Without being able to effect this separation, we lose the context that makes intelligible the task of individuating and identifying biological mechanisms (here, of emotional functioning). So while the *idea* of the intrinsic worthiness of delineating in value-neutral terms the boundary of emotional disorder may seem an attractive one, it is an idea whose apparent coherence looks to dissolve under closer inspection, resting as it does upon a presumed distinction concerning concepts of pathology that ultimately cannot be sustained with respect to our emotional experiences.

conception of flourishing will tend to give us reason to foreground and prioritize the appraisals made along the other BEDs. It is out of this complex negotiation of values that the judgement of emotional disorder comes to be made.

So what, if anything, can be said by way of a positive characterization of disordered emotional states? I suggested in Chapter 1 that, to a first approximation, a negative appraisal of an emotional response in all relevant BEDs—and no positive appraisal of this response in any dimension—will tend us towards the judgement of emotional disorder. But is there something more that can be said about the nature of an emotional response that is thus judged, something in virtue of which it can be said to be disordered? Otherwise, why apply this meta-evaluative concept at all? These are important questions that merit further investigation (I will say something further about this in Section 6.3). Here, I wish to offer one cautionary thought. It is a fact about our current practices that we have an interest in extending the concept of disorder into the domain of human emotions. That is, we have an interest in viewing at least some of our emotional states as disordered. However, in being attracted to the idea that it is possible for us to suffer from emotional pathology—that is, in being drawn to the idea that certain kinds of emotional experience may be a form of illness or disorder—we should be wary of the risk of oversimplifying the concept of emotional disorder. The aim of this inquiry has been to show just how pressing and real this risk is given the multitude of ways it is possible, in our characterization of the concept, to miss out and overlook non-health-related appraisals of our emotional experiences that embody aspects of human life we value deeply, appraisals that turn out to bear upon the determination of a particular emotional response as disordered or otherwise. This risk of blindness to whole swathes of axiologically significant facets of human life, facets lived out through our emotional experiences, is I believe the root source of the threat of over-pathologization of our emotional lives.

The result of appreciating the true complexity involved in judging an emotional response to be disordered is the recognition that it may not be possible, ever, to reify the state of emotional disorder—either in terms of the notion of biological dysfunction, or of a unique phenomenological or biochemical profile, or simply of distress. We may come to the view, on this basis, that the task of judging an emotional response to be disordered is an inherently dynamic process that involves negotiating the five basic evaluative dimensions in perceptive, imaginative and highly context-sensitive ways. And this might lead us to suspect that diagnosing a disordered emotional state is nothing like diagnosing a heart condition, and perhaps not even much like diagnosing psychosis.

5.3 Mental Health, Moral Virtue, and the Therapeutic Turn

In the previous chapter (Section 4.6.1), I suggested that psychotherapeutic work can be understood to play a pedagogical role that aims at helping us exercise our rational agency in better ways through the cultivation of more reasonable emotional dispositions and the development of greater self-understanding, where this includes making more explicit to ourselves our conception of the good (as embodied in our most deeply held values). This is not the most widely held view of psychotherapy today, however. With the dawning of the so-called 'Therapeutic Turn'[15], what has become more familiar is the idea that psychotherapy aims at helping us to recover from illness (illnesses that are more often than not emotional in nature). Charles Taylor, whose characterization of two approaches to depressive$_{DRIC}$ experiences we looked at in Section 3.4, articulates this latter conception of psychotherapy in the following passage:

> Psychoanalysis may seem, and partly is, an intermediate phenomenon. Unlike behavioural therapies, or those relying mainly on drugs, it involves a hermeneutic, an attempt to understand the meaning of our unease. *But its goal is the same*; the hermeneutic delves into the unavoidable, deep psychic conflicts in our make-up. But *these have no moral lesson for us*; the guilt or remorse points to no real wrong. *We strive to understand them in order to reduce their force, to become able to live with them*. On the crucial issue, what we have morally or spiritually *to learn from* our suffering, it is firmly on the therapeutic side: the answer is 'nothing'. (Taylor 2007: 621, italics added)

Of the two approaches to the suffering engendered by 'deep psychic conflicts' Taylor identifies in this passage, the therapeutic/spiritual and the pathological, it is interesting to note that Taylor too sees the line that divides pathologizing and non-pathologizing approaches as centring upon the notion of a pedagogical path that leads to greater psychological and moral maturity—with pathologizing approaches eschewing this notion altogether.[16] This

[15] The phrase 'the therapeutic turn' is used by Charles Taylor to refer to the shift in our culture from 'a hermeneutic of sin to one of sickness' (Taylor 2007: 620). This shift has been commented upon by numerous authors, including Taylor. For other discussions of this shift, see for example Rieff 1966 (in his monograph, *The Triumph of the Therapeutic*), and more recently, Madsen 2014.

[16] See Section 3.4 for clarification of the perspectives here referred to as 'therapeutic/spiritual' and 'pathological' respectively.

encourages anyone who adopts a pathologizing approach (and this includes psychotherapists, according to Taylor) to respond to emotional distress by aiming exclusively to remove (or at least to reduce) this distress, doing this in an interventionist manner.

While Taylor is describing a widely held understanding of psychotherapy, it is important that we do not lose sight of the very different conception of psychotherapy sketched out in Section 4.6.1. This matters not least because psychotherapy is a key domain in which the moral outlook on human life is widely suspected of being gradually replaced by the concepts of mental health/illness. How we reconcile these two views of psychotherapy—psychotherapy as aiming respectively at character development and at treating illness—will therefore have potentially far-reaching implications for the way we construe the nature of distressing emotional experiences. So how should we think about the practice of psychotherapy, given that we are presented with two views that are, on the face of it, distinct from each other? The least revisionary stance we can adopt here would be to accommodate both views of psychotherapy simultaneously, i.e. to say that psychotherapy is engaged in both moral and medical concerns. This is indeed how I think we ought to understand the nature of the psychotherapeutic endeavour as it is practised today. Importantly, as we shall see below, the conceptual resources afforded by the BED/EF approach make available a satisfactory way of achieving this accommodation. There is, however, a less optimal way of doing so—less optimal because it involves obfuscating the boundary between the medical and the moral. Let me begin with the latter.

5.3.1 Mental Health and Moral Virtue: The Assimilationist Stance

We might seek accommodation by holding that the boundary between medical and moral discourse is not a hard and fast one, and in this vein, viewing the psychotherapeutic endeavour—construed as a singular process—as simultaneously facilitating psychological growth *and* as promoting recovery from illness, doing so in virtue of the supposed fact that both of these really come to the same thing. According to advocates of this assimilationist stance, medical and moral discourse are best integrated with each other to produce a single, fused 'moral-medical perspective', so that the two modes of discourse are understood as making the same kind of normative assessment. This stance has been adopted, amongst others, by Martin (2006) and Pickard (2009) and

below I will look at some of the reasons that authors have given for adopting this stance. To anticipate my response: I will draw upon the BED/EF framework to disentangle the main threads that lead to the adoption of this stance, and to show that the two views of psychotherapy can be accommodated without having to collapse the boundary between the medical and the moral in the way that the assimilationist stance does.

I want to begin by offering a reason for being wary of the idea of a single 'moral-medical perspective'. The stance has an air of paradox about it, arising from the key nodes in the structure of our concepts of moral agency and pathology respectively. These nodes hold these concepts in place and make them recognizable to us as such, and revising them in order to dissipate the paradox that would otherwise beset the idea of a fused 'moral-medical' perspective comes at a significant conceptual cost. Specifically, to square both claims—that psychotherapy facilitates psychological growth and that it seeks to treat our maladies—whilst construing psychotherapy as a singular process requires us to substantially water down the concept of illness. Recall that according to the strict construal of this concept (cf. Section 3.2), the having of a pathological emotional experience is a meaning-free occurrence, the appropriate response to which is its removal by interventionist means (that is, by treating it). Inherent to the notion of psychological growth, by contrast, is a process of change the characterization of which makes essential reference to how the rational agent herself sees and understands the world, and how she relates to others and to herself. To illustrate this difference, let's return to the example of the depressive$_{\text{DRIC}}$ experience. Even if psychological growth and the treatment of illness both result in behavioural change, behavioural change that is expressive of psychological growth is fundamentally a change that is rooted firmly within the rational agent's own perspective. Her perspective changes, but not by external manipulation, for instance by having a depressed outlook replaced through efficacious interventionist means with a happier and more hopeful outlook. The agent comes to understand, rather, something essential about herself in a depressive$_{\text{DRIC}}$ experience, and this understanding is what then provides her with a *reason* to seek to bring her outward circumstances into greater alignment with her conception of the good. Throughout, in the work of self-examination and in coming to recognize that she has reason for seeking to bring about this realignment, the agent exercises her rational agency. In this, she has precisely *not* been the passive recipient of an intervention that shifts her from a state of despair to a state of hopefulness. But the concept of being treated for a disorder, as it stands, connotes the latter view of the agent's behavioural change. So the only way of assimilating the idea of recovering from illness with that of psychological growth is by significantly

loosening the strict construal of the pathology concept so that it is no longer necessary to assume that the having of a pathological emotional experience is an occurrence that is devoid of meaning. This loosening is what allows us to speak of treating or curing an emotional malady even when what we are doing is exercising our rational agency and coming to understand ourselves better by grappling with an existentially significant depressive$_{DRIC}$ experience. There is a twofold hazard of accommodating the two views of psychotherapy in this manner: firstly, the concept of pathology threatens to end up being loosened to the point of vacuity; secondly and relatedly, the distinctly developmental or maturational character of what we are doing threatens to be occluded by the medical overlay that is now made possible by this loosening.

If it is difficult to keep this distinction in focus between psychological development and being treated for illness, this is perhaps because depressive$_{DRIC}$ experiences can look a lot like depressive experiences deemed to be devoid of any existential significance. This can lead to a confusion about the nature of the experiential state that is presented in the psychotherapeutic encounter, i.e. their superficial resemblance to each other can lead to our mistaking an experiential state imbued with moral import for a pathological state. Both kinds of experience can manifest phenomenologically similar qualities of experience (as discussed in Chapter 2), and both can have a similarly devastating impact on the subject's daily functioning and on her relationship to herself and to life (e.g. to her sense of self-worth and her commitment to living well). Given this surface resemblance, it can seem reasonable to lump together the two kinds of depressive experience—effectively taking them to constitute a single category—and to adopt a laissez-faire approach that eventually erodes a distinction that had initially seemed important. Psychotherapeutic work that engages this (supposedly) unitary type of experience may sometimes look to be a medical intervention and sometimes look to facilitate psychological growth, but both processes are situated within the same type of practice (psychotherapy), and both processes are taken to be aimed at altering a single, unitary type of experience (an experience of a pathological sort, namely depressive illness). With this merging of experiential types and of our responses to it, the distinction between psychotherapy-as-pedagogy and psychotherapy-as-medical-treatment can seem to diminish in importance, giving way to the emergence of the notion of a fused 'moral-medical' perspective. ('Adopt a moral or a medical response, whichever works! In the end, it comes to the same thing.')

By registering the dimension of existential significance, however, the basic evaluative dimension (BED) framework enables us to recognize the

possibility that there might be emotional experiences that look a lot like pathological depression but which aren't in fact pathological precisely because they bear existential significance (depressive$_{DRIC}$ experiences being a paradigmatic example of this). This undermines the rationale just presented for endorsing the notion of a fused moral-medical perspective, in the following way. The BED framework underscores the fact that judgements of emotional disorder are meta-evaluative judgements, which means that the pathology status of an emotional experience is derived from appraisals made in more basic evaluative dimensions, including that of existential significance. A positive appraisal of an emotional experience in any dimension gives us *prima facie* reason to resist the pathologization of the experience (this must then be negotiated alongside appraisals of the experience in other dimensions). Conversely, a negative appraisal gives us *prima facie* reason to pathologize it. Applying this general point to the dimension of existential significance, and with the strict construal of the pathology concept in place, we can say the following: we positively appraise an emotional experience in the dimension of existential significance on the basis of deeming it to bear such significance, and this provides us with reason not to pathologize the experience. Being positively appraised within the dimension of existential significance, a depressive$_{DRIC}$ experience is viewed as being inherently meaningful, and the having of such an experience is deemed also to be meaningful. Conversely, we have reason to pathologize the experience if we deem it to be devoid of existential significance. Furthermore, once we bring to the fore the two conceptions of flourishing discussed in the previous chapter, we see that the decision not to pathologize an existentially significant depressive experience can matter in an ultimate way: it is not necessarily the case that whether or not the depressive$_{DRIC}$ experience is pathologized, the subject of the experience is ultimately led through psychotherapeutic engagement to the same place, namely a life in which negatively valenced emotional experiences are greatly reduced. Rather, the decision not to pathologize a depressive$_{DRIC}$ experience could in fact be partially constitutive of an agent's attempt to live a more enlightened life, one in which she engages her rational agency more fully, where this engagement is valued for its own sake.

This is in effect to assert the claim that some kinds of depressive experience are morally or existentially replete, while other kinds are simply conditions of pathology. Upholding this distinction between kinds of depressive experience in turn allows us to (a) see that the judgement of pathology legitimately takes hold *exclusively* in the latter case, and to (b) appreciate why (for instance) the Hollisian psychoanalytic undertaking is a distinctly moral project (rather than a supposedly fused moral-medical one). It addresses a moral predicament and

seeks to facilitate a subject's psychological growth by engaging her rational agency (in the manner described in Section 4.6). Thus, although the two kinds of depressive experience may be phenomenologically similar, and even if the impact on a person's functioning of depressive experiences is similar across both types of experience, the BED framework enables us to see that these similarities should not be invested with a boundary-delineating significance. Correctly delineating the boundary between disordered and non-disordered depressive experiences involves, rather, registering the existential significance (or otherwise) of such experiences. Registering the appraisal made along this basic dimension of evaluation helps to clear up the confusion about the nature of depressive experiences just described, thereby also divesting us of the paradoxical notion of a fused 'moral-medical' perspective since the latter notion is sustained largely by this confusion. Being divested of this paradoxical notion in turn means that there is no longer the need to loosen the strict construal of the concept of pathology.

5.3.2 Mental Health and Moral Virtue: The Pluralistic Stance

We are now in a position to appreciate that the assimilationist stance is a hasty and costly way of accommodating the view that psychotherapy is (legitimately) concerned with matters of both moral and medical import. We obtain this accommodation at the needless cost of collapsing a distinction that we would otherwise regard as significant and useful. The BED framework reveals to us that we have the conceptual resources to accommodate the moral and the medical strands of psychotherapy in a different, and better, way; better, because this alternative preserves the substance of the concept of pathology. Thus, by enabling us to grasp that depressive$_{\text{DRIC}}$ experiences are existentially significant experiences that can facilitate psychological growth, and by allowing us to continue to hold on to the strict construal of pathology, the BED framework gives us grounds for placing the psychoanalytical response to depressive$_{\text{DRIC}}$ experiences articulated by James Hollis squarely within the domain of moral discourse (as stated in the previous two paragraphs), and for recognizing that psychotherapy is not a singular process. Instead, it involves heterogeneous practices—some moral, some medical—that respectively address non-disordered and disordered emotions that we deem problematic but which are differentially appraised within the basic evaluative space. In construing the nature of psychotherapy this way, we adopt what I shall call the pluralistic stance.

Recognizing the pluralistic stance enables us to respect the distinction between moral and medical discourses in a comprehensive characterization of the psychotherapeutic enterprise. It is worth noting here that medical interventions do not just have to involve pharmacological interventions; they could include 'talking therapies' that aim squarely at alleviating suffering and therein at the removal of symptoms of illness in an interventionist manner. Examples include encouraging the depressed subject to keep a gratitude journal, to notice 'maladaptive' thoughts and to correct them, or to devote energy and time to noticing sensations of ease when these arise. What all medical interventions have in common—the point in virtue of which these interventions count as 'medical'—is that they do not engage the depressive$_{DRIC}$ experience itself as a meaningful experience. Instead, they attempt to alleviate suffering by intervening on other variables that are deemed to be causally efficacious in producing positively valenced emotional experiences.[17] If relevant variables are correctly identified, the result is that the subject may feel better—and on this basis, he may be deemed to have recovered from the depressive episode. This is achieved not through engaging hermeneutically with an intrapsychic conflict that lies at the heart of depressive experience (when indeed the experience is a depressive$_{DRIC}$ experience), but rather by bypassing it. By contrast, when the intrapsychic conflict is placed at the heart of the inquiry and an attempt is made to meaningfully engage with it, the psychoanalytic undertaking is most appropriately understood as a moral project, in which the aim of the undertaking is understood not to be a recovery from illness, but rather to facilitate the process of psychological maturation through the development of the agent's capacities involved in evaluative discernment, in living more authentically and in exercising greater personal autonomy. Amongst other things, this maturity manifests in the agent's living a life in which her emotional tendencies and her choices become gradually more aligned with her conception of the good.

[17] The practice of keeping a gratitude journal is an interesting one. Besides the interventionist construal just sketched out, there is a different construal of this practice, i.e. as a practice aimed at developing our ability to take measure of our life circumstances in a more balanced and/or realistic manner. Understood in these terms, the practice isn't employed in an interventionistic manner (aimed exhaustively at increasing the episodes of positively valenced emotional experiences), but forms part of our learning to better exercise our rational agency.

5.3.3 Further Considerations in Favour of the Assimilationist Stance: A Critique

I want to briefly consider two further ideas that incline us towards the adoption of the assimilationist stance (rather than the pluralistic stance). My aim in doing this is to suggest that once we have in view the BED analytical framework—reinforced by the Aristotelian (EF) conception of flourishing—these considerations turn out not to be sufficiently compelling to motivate a loosening of the strict construal of pathology (as the assimilationist stance prescribes). The first guiding idea that motivates the assimilationist stance is that it alone promises to yield a scientifically informed understanding of the human condition, and that this explanatory power derives specifically from the medical strand of the stance. On this view, morality in itself—when unassimilated to the medical perspective—is fancied as being somewhat akin to superstition, that is, a not-very-good way of explaining a phenomenon of interest (in this case, human misery as it shows up in emotional distress and emotionally driven bad behaviour). As the brain and behavioural sciences reveal to us the cogs and wheels that make up our psychological machinery, we see (so the thought goes) that there is less to condemn (morally), and more to treat and to fix (medically).[18]

This guiding idea presumes that a scientific understanding of human psychology is the exclusive province of medicine, and thus that a scientifically informed understanding of our emotional lives necessarily introduces the register of health and disorder.[19] But there is no reason for supposing that it must. Aristotelian naturalism underscores the point that a scientific understanding of human psychology is just as relevant to our moral outlook and to our understanding of ourselves as moral and deliberative agents, as it is to identifying and treating the psychopathologies we are taken to suffer from, since our moral agency—along with everything else human—is ultimately rooted in our psychological make-up and our biology. As we saw in Section 4.4.1,

[18] This idea goes back at least to John Dewey, who called for a rejection of the distinction between moral goods (virtue) and natural goods (health), but doing so by privileging the latter concept. (Dewey 2004, originally published 1920).

[19] We see this assumption at work in Mike Martin's claim, in defence of the assimilationist stance, that sound morality is healthy. According to Martin, it is only through the introduction of the concept of health that a moral perspective 'will be psychologically realistic in its demands and take account of ongoing scientific inquiry' into a given domain of mental functioning (in this case: emotional functioning), and in this way be sound. But we will feel the need to introduce the concept of health to develop an empirically informed understanding of human agency only if we think that scientific research into the mind is not relevant to our understanding of our moral development and our ethical conduct (in its own right).

Aristotle himself sought to articulate a conception of human flourishing that is psychologically realistic and liveable, which is why he took our emotions and appetites, no less than reason, to play a decisive role in our flourishing. Viewed from this angle, it seems strange to embed an empirically informed understanding of our psychological make-up and our biology exclusively within a non-moral paradigm (i.e. a medical paradigm). Or at least: it would look to us strange to do this, unless one were already in the grip of the guiding idea—one that is antithetical to Aristotelian naturalism about human rational agency—that a scientifically informed understanding of the human condition can be obtained only by viewing our condition through a medical lens.

A variant of this guiding idea appears in an argument presented by Hanna Pickard in favour of construing Cluster-B personality disorders—specifically those marked by dispositional impulsivity and a tendency towards excessive anger—as medical conditions (Pickard 2009: 96). In this argument, Pickard identifies the moral outlook as such with a particular brand of morality, one that is unappealing because it is punitive and looks to be psychologically improbable: moral conduct consists, according to this view, in adhering to a set of moral imperatives that the excessively angry person has to be persuaded to endorse, and psychotherapy-as-a-moral-endeavour consists in 'converting' this person through persuasion to adopt the moral outlook and by this means he will (hopefully!) begin to conduct himself in morally more commendable ways. (I'll refer this to the 'moral conversion' view for brevity.) Having characterized the moral outlook in this way, Pickard then asserts the following claim: noting that the Aristotelian account of character development through enriched habituation is informed by empirical research—and so looks to be more psychologically realistic and thus promises to be more efficacious—Pickard draws the conclusion that enriched habituation is best understood as a *treatment* for B-cluster personality disorders (and not just as a moral undertaking). These personality disorders are accordingly best construed as medical conditions.

Pickard's conclusion that the processes of Aristotelian (enriched) habituation must be construed as medical interventions aimed at treating a disorder, even if they *also* happen to be describable in moral terms as processes that facilitate character development, follows only if we arbitrarily restrict the scope of morality (as such) to just the moral conversion view and insist that we must overlay our conception of morality with the medical perspective in order to form a fused moral-medical perspective that enables us to encompass processes of enriched habituation within our medical response to individuals suffering from the excessive anger and/or dispositional impulsivity (character

traits that are prevalent in Cluster-B personality disorders, especially so-called borderline or emotionally unstable personality disorder).

But why must we do this? Independently of Pickard's argument, and in particular its restriction of our conception of morality to the moral conversion view, enriched habituation offers us a perfectly familiar and deeply compelling view of what growing to inhabit an ethical outlook (as such) involves, at least for human beings. It seems a gross distortion of the processes of enriched habituation to fail to recognize that it is first and foremost ethical in character. Furthermore, the pluralistic stance towards psychotherapy—as opposed to the assimilationist stance—makes available the option of not overlaying our conception of morality with the medical perspective. Avoiding this overlay enables us to see that enriched habituation constitutes a moral process *simpliciter* of character development (with no reference to the concept of medical treatment required). This view arguably captures more perspicuously the processes of enriched habituation that are engaged psychotherapeutically in the context of borderline personality disorder today: individuals are taught skills that help them to develop greater self-awareness (skills such as mentalization and mindfulness), to tolerate distressing emotional experiences and to exercise more effective control over emotional reactivity (this process is referred to as 'emotion regulation' in the psychotherapeutic literature, and includes techniques ranging from self-soothing (through deep breathing, for instance), to distraction, to 'radical acceptance').[20] These are skills we *all* acquire and become better at exercising as we mature psychologically, and they are integral to exercising our epistemic, moral, and deliberative agency well.[21]

Against the charge that overlaying an accurate understanding of enriched habituation as a moral process with the concept of treatment threatens to distort this understanding, a proponent of the assimilationist approach might contend that the concept of medical treatment has as a matter of fact been expanded to address problems of moral character (and we see this most clearly

[20] For an overview of approaches currently used psychotherapeutically with individuals diagnosed with borderline personality disorder, see Dixon-Gordon et al. 2017. For an introduction to the notion of emotion regulation more broadly, and for the suggestion that emotional regulation is an 'essential feature' of mental health, see Gross and Munoz 1995. See also Jazaieri, Urry, and Gross 2013. Although James Gross and his colleagues write about 'emotional regulation' in the register of health, it seems to me that much of what is said can be formulated just as adequately, if not more so, in overtly ethical terms. Doing this helps to enrich our understanding of the development of skills and capacities that are integral to the competent exercise of rational agency.

[21] This is not to overlook the fact that our moral education is augmented by a further range of sources—for instance, by closely observing morally exemplary figures (cf. Section 4.5.3), by studying rich portrayals of human life in fictional and historical writing and in film/theatre, and by scrutinizing intelligent political commentary. These sources widen our emotional range, they enrich our imaginative capacity and they improve our emotional literacy, therein enabling us (amongst other things) to evaluatively perceive more accurately, and to make finer-grained discriminations in our moral judgements.

precisely in the context of the personality disorders). It is no longer restricted to just techniques of getting rid of distressing emotional experiences and dispositions (deemed to be disordered on account of the distress they cause). In this way, the boundary is softened between moral virtue and mental health (and conversely between moral incontinence and vice, and mental disorder), and this in turn paves the way for us to adopt the assimilationist stance.

But the fact that the concept of treatment has expanded this way is itself the result of a confusion that I suspect ultimately stems from the predominance of a single conception of flourishing that controls our understanding of what the intended outcome(s) are of any process of enriched habituation that is engaged in the psychotherapeutic context, i.e. the health-based conception. As we saw in the previous chapter, according to this conception, the intended outcome of any psychotherapeutic response to emotional distress—this includes enriched habituation—is understood principally to be the reduction of negatively valenced emotional experiences (to make the condition 'more liveable', as Charles Taylor puts it). If this is the only outcome it is possible to envisage, then the characterization of enriched habituation as a form of medical treatment begins to look more innocuous.

Once we recognize that enriched habituation may serve an altogether distinct purpose, however, i.e. that of learning to better exercise our rational agency, we see that there are more apt concepts than that of *treatment* by which to describe the processes of enriched habituation: to wit, *education* and *cultivation*. That these latter concepts are perfectly good ones to hold on to can be seen if we flip the picture on its head and start speaking of *treating* children when we teach them to develop greater self-awareness and emotional literacy, to cultivate good emotional dispositions, to modulate their emotional responses so that they are able to evaluatively discern matters accurately. The use of the term 'treatment' in this context strikes us as absurd because age-appropriate psychological immaturity is not a condition of pathology that calls to be treated. What the Aristotelian conception of flourishing suggests is that, to lesser or greater degrees, psychological immaturity may persist long into biological adulthood (psychological growth being shaped by the many contingencies of life, including adverse early relational experiences for some), and that this fact about us is as relevant to the psychotherapeutic undertaking as the fact that we suffer from distressing emotions. Indeed, if psychological maturity is understood in terms of agential excellence (construed in Aristotelian terms), then the task of developing greater psychological maturity is probably a life-long project for most of us. Viewing the matter in this light, we are able to appreciate that the notion of cultivation is just as central and just as primary to the psychotherapeutic endeavour as is the notion of treatment, and that the

concept of cultivation better describes the processes of enriched habituation undertaken by those who suffer from excessive anger and/or dispositional impulsivity because it is this concept (rather than that of treatment) that captures the distinctly ethical character of this process.[22]

As an aside, it is perhaps worth remarking that this casts in a new light trauma-informed understandings of emotional distress, which are growing in influence in critical response to the pathologization of this distress (Sweeney et al. 2018). Whilst proponents of the trauma-informed approach are generally keen to emphasize that such distress is not pathological (often this is coupled with the further claim that emotional distress may even be adaptive (see for instance Filson 2013)), we might add that early relational trauma—and its impact on human development—is *morally* significant. This observation is less about pointing the finger of moral blame at the adults who inflict this trauma on children (they may themselves have experienced early relational trauma); it is, rather, about becoming clear about the nature of the predicament of a human being who grows into biological adulthood without the psychological resources that foster good human development vis-à-vis agential excellence. Having merely the lens of health/disorder before us leaves us trying to decide whether or not emotional distress rooted in early relational trauma signifies disorder, when in fact the more significant point seems to be that these adverse early experiences are ethically corrosive in that they pose a considerable obstacle to the realization of our full potential as rational agents. By the same token, a recognition of this point invites us to interpret anew the significance of wider social, economic, and political conditions that have long been noticed to bear a strong correlation with emotional distress and suffering.[23]

[22] The pervasiveness of privileging the idea of treatment in our therapeutic discourse today is succinctly illustrated in an article about anxiety and the intolerance of uncertainty in a major UK broadsheet, in which the author, Shayla Love, cites Brady Nelson, an associate professor of clinical psychology, as follows: 'Because IU [intolerance of uncertainty] is associated with many mental health conditions, *treating* it could help with any or all of them. Rather than attempting to *treat* multiple distinct mental health conditions, it might be more effective to focus *treatment* on increasing tolerance of uncertainty' (italics added). In light of the foregoing discussion in this section, it seems to me that none of the 'treatment' responses suggested in the article are unamenable to being conceived of in terms of the idea of cultivating a more reasonable emotional disposition (rather than that of treating a medical condition). Indeed, the final suggestion presented in this article makes simpler sense when interpreted in the latter terms: 'Life will always be full of unknowns. When facing the uncertain, you might feel discomfort, but there's also a chance to try to feel curious and open. 'Ask yourself, "What can I learn here?" ' [Naomi] Koerner [associate professor of psychology at Toronto Metropolitan University] said. When we can find a way to be excited or intrigued by the unknown, rather than afraid, uncertainty becomes much more palatable.' Fear of the unknown: are you more sensitive to uncertainty than others? | Well actually | *The Guardian* (https://www.theguardian.com/wellness/2024/apr/09/uncertainty-mental-health (last accessed 17 May 2024)).

[23] Recent work by Madsen (2018), Davies (2021), and others questions the move to cast these conditions as precipitants of mental ill-health and returns us to a consideration that Aristotle grasped vividly, namely that human flourishing is built upon the foundation of conducive childhoods, and that this foundation in turn rests upon a socio-political milieu that fosters good human growth (cf. Aristotle (transl. Stalley) 1995).

In sum, Pickard's argument for the assimilationist stance rests on the presumption of an impoverished conception of the ethical life as such (expressed in the moral conversion view), combined with the questionable assumption that only a medical view of traits like excessive anger and dispositional impulsivity can yield a psychologically realistic understanding of these traits. Neither assumption looks remotely attractive once we allow for the possibility that empirical research into the human mind can (and ought to) inform *both* our understanding of conditions of pathology *and* our understanding of human moral agency and its development—the latter possibility being registered especially through the concept of enriched habituation as it was introduced in Section 4.6.1. This concept helps us to see that there is no reason to suppose that psychological research into human behaviour necessarily has to be filtered through the medical/health perspective. Indeed, the addition of this perspective can be superfluous—it does not add anything to the moral perspective in terms of which character development is normally understood—if not altogether misleading.[24]

If there is a meaningful distinction to be drawn between our moral outlook and the psychiatric view of emotional distress, as I have been arguing there is, then it is important to continue to maintain our handle on this distinction even in domains of practice where the two concerns—the moral and the medical—are closely interrelated, as it is in psychotherapy. Indeed, once we recognize that psychotherapy can engage both distinct concerns—once, that is, we construe the functions of psychotherapy in pluralistic rather than in assimilationist terms—it becomes possible for us to see that psychotherapists may be engaged in the psychoeducational task of facilitating character development, alongside the distinct task of providing treatment. The line between the different tasks may be a very fine one in practice, but it is nonetheless a line

[24] Pickard in fact concludes that asking whether Cluster-B personality disorders are a medical condition (ibid.: 98) is an idle question. This claim overlooks the very real risk of pathologizing distressing emotional experiences by the back door, as it were, i.e. construing enriched habituation as a treatment process, we then argue on this basis that excessive anger (which is 'treated' through enriched habituation) must therefore be a pathological condition (since it is a condition we respond to by trying to 'treat'). A growing sense of disquiet about the perceived limitations of medicalization through biological reductionism, in psychiatry, has led Jussi Valtonen and Bradley Lewis to suggest that we ought to carve out the space for a discipline they call 'mental health humanities', which in our study of mental suffering is to be regarded as a 'coequal partner with natural sciences' (Valtonen and Lewis 2023). At the heart of this proposed discipline is the use of research methodologies that can help us deepen our understanding of the particularities of human experience. This echoes what Karl Jaspers advocated in his *General Psychopathology* (1913/1997). Whilst I am sympathetic to this approach, I'd like to note here that the BED/EF framework invites us to consider the underlying possibility that much of our emotional suffering may in fact be being *arbitrarily* cast as states of pathology, and that this possibility in turn invites us to wonder about what (much of) this suffering might look like to us if we bracket off the framing concepts of health and pathology in our interpretation of and responses to this suffering.

that makes all the difference to our self-understanding, either as (imperfect) rational agents or as subjects of disorder.

I mentioned at the start of this section that there are *two* ideas that incline us towards the adoption of the assimilationist stance (rather than the pluralistic stance). I now want to turn to the second of these two guiding ideas, which is the idea that to live well morally is to live healthily. It has been claimed that this is an ancient idea that goes back at least to Plato's writings. There is much debate about whether Plato was speaking metaphorically or literally, but for our purposes, we would do well to take a step back and ask whether by 'health' the ancient Greeks meant anything like what we mean by the term today, especially our concept of emotional health as it figures in the contemporary health-based conception of flourishing (outlined in Section 4.2). The ancient Greek concept of the health of the soul/psyche is probably better analysed in terms of the concept of moral virtue, with the latter being understood in a robust sense that accords a central position to the moral agent who can be held accountable for his or her actions and affective dispositions. One way to see this is by noting that Plato equates his notion of mental health with that of psychic unity. A person whose emotions and appetites are harmoniously unified with her reason (i.e. effectively structured by her conception of the good) is a virtuous or excellent human agent. That the structuring of our emotions and appetites by our conception of the good plays a defining role in Plato's characterization of mental health points to an essentially agential construal of Plato's notion of 'mental health'.

It is true that Aristotle adds that, in excelling through the achievement of psychic unity (that is, in the way of the wise rather than the wicked), such an agent flourishes as a human specimen—just as a well-fed plant is an excellent plant specimen, and so flourishes as a plant. We might be tempted to add on this basis that the excellent human agent and the excellent plant are respectively *healthy* specimens of the species to which they respectively belong (it is interesting to note here that Aristotle never in fact explicitly uses the adjective 'healthy' in this way). Tempted though we might be to add the predicate 'healthy' to the judgement of excellence, we shouldn't be misled by this into thinking that there is some common, substantive quality shared by both the 'healthy human' and the 'healthy plant'. They excel in altogether distinct and irreducible ways, moral excellence (in the sense described in the last chapter) being central only to human flourishing. Summarizing this point, I suggest that, given our own conceptual baggage around the concept of health, we are better able to understand the Aristotelian view of human excellence if we resist this temptation altogether. Still, if we do wish to insert the concept of

health into the Aristotelian picture of flourishing, we should be clear that the notion of health—as it figures in the idea that an excellent human specimen is a healthy specimen (though to repeat: this is not an idea that Aristotle himself expresses)—is an empty one, one that is derivative of but adds nothing to the judgement that a particular person flourishes in virtue of exercising her epistemic, moral, and deliberative agency excellently.[25]

That this is how we should handle the concept of health here is reinforced by the fact that we have good reason to hold that Aristotle eschewed an assimilationist view of the concepts of moral virtue and emotional health: to think otherwise (to suppose that Aristotle was an assimilationist) is to risk losing our grasp on a claim that is central to the Aristotelian conception of human flourishing, namely that emotional incontinence and vice are moral shortcomings, character failings, and not diseases of the mind (see discussion of the varieties of affective irrationality in Section 4.5.2.1; cf. Harcourt 2013: 50). Casting these shortcomings in terms of the framing concept of disorder, as the assimilationist stance encourages us to do, effectively masks our rational imperfection. This yields a seriously incomplete picture of our rational agency as such—an incompleteness that would make it considerably harder for us to appreciate practical wisdom and emotional continence for the achievements of our rational agency that they are.

The upshot of the discussion in this section is that the conceptual resources made available by the BED/EF analytical framework help to clarify the point that we have no reason for favouring the assimilationist stance towards psychotherapy over the pluralistic stance, and that so long as the pluralistic stance remains firmly in view, the therapeutic turn furnishes us with no independent reason to delineate the concept of emotional disorder in such a way as to subsume under it an ever-greater range of emotional distress.

[25] Using Terence Irwin's helpful terminology, the suggestion here is that we should 'inflate' the concept of health into that of moral virtue (Irwin 2013: 38–9).

6
Conclusion

6.1 The Argument Recapitulated

The psychiatrist meets her patient in a state of distress, and since it is a general aim of medical intervention to alleviate suffering, it is natural for this affective aspect of human emotional experience to be foregrounded in this encounter. Given also that this happens to be the aspect of our emotions over which we seem to have least control—it is that aspect with regard to which we seem most passive—an interventionist response appears to be both appropriate and adequate. So long as the focus of our attention remains squarely on the felt sense of distress, therefore, the current practice of identifying disorder with distress in the domain of our emotional lives looks to have *prima facie* plausibility.

What the analysis of the concept of emotional disorder presented in this book reveals, however, is that affectivity or the valence quality of an emotion is just one of at least five aspects of our emotional experiences that are significant to us. The main thesis that has emerged from this analysis is that, when we bring all five aspects into view, we are in a position to recover a distinct understanding of human emotion, according to which our emotional experiences are an important means through which we exercise our rational agency. This understanding is distinct from the pathology construal of distressing emotional experiences; the latter can even cancel out the former given the grammar of the concepts of pathology and of rational agency respectively.

Our emotional lives are poised in this way—between these two radically different views (agential vs pathological)—because the concept of emotional disorder is meta-evaluative. Disorderedness does not wear itself on the sleeves of our emotional experiences. This means, as we've seen in the first three chapters, that we have to negotiate disparate (and often conflicting) constraints placed upon us by the appraisals we make of our emotional experiences in the five basic evaluative dimensions (BEDs). It is only on the basis of this negotiation, whether undertaken explicitly or implicitly, that we arrive at the

The Concept of Emotional Disorder. Gloria Sibson Ayob, Oxford University Press. © Oxford University Press 2024.
DOI: 10.1093/9780198909620.003.0007

judgement of emotional disorder.[1] Crucially, it turns out that in negotiating these basic evaluative dimensions (BEDs), we are in fact deciding between two distinct conceptions of our emotional lives—one that places our emotions on the health-illness spectrum and the other that takes (large swathes of) our emotional experiences to signify an exercise of our rational agency. This deciding is foisted upon us because our negotiation of the basic evaluative space is informed from the top down by our global conceptions of flourishing, conceptions that accord a central position to our emotions but that construe them in radically different ways.

We saw at the end of the third chapter how, guided by the health-based conception of flourishing (HF) alone, we naturally tend to place our emotional experiences on the health-disorder spectrum and, on this basis, we are led to suppose that it ultimately doesn't matter which way we negotiate the appraisals we make of our emotional experiences in the basic evaluative dimensions. It is assumed that, whichever way we negotiate the BED space, what we are principally aiming at in our emotional lives is to reduce the occurrence of negatively valenced emotional experiences ('emotional pathology'), and to promote emotional experiences that are positively valenced (thereby achieving 'emotional well-being').[2] Placing our emotional experiences on the health-disorder spectrum in this way is a substantive and significant move, however, because it positions them orthogonally to the domain of rational agency. Guided in this way by the HF in negotiating the basic evaluative space, we might in turn be led to minimize the weight we assign to the positive appraisals made of a distressing emotional experience along the BEDs of emotional rationality and/or existential significance, failing to appreciate the full import of these latter appraisals and on this basis arriving at the judgement of emotional disorder.

It is only by bringing into view a radically distinct conception of flourishing, i.e. the Aristotelian excellence-based conception of flourishing (EF), that we recover a conception of human emotion that allows us to see that our emotional experiences can and do constitute an exercise of our rational agency. This correspondingly enables us to appreciate the value of appraisals made of our emotional experiences in the BEDs of rationality, existential significance, and interpersonal functioning, which in turn grounds the move to weight these appraisals more heavily than Kendler does. This shift in the way

[1] I have shown in Chapter 2 that attempts to circumvent or avoid this basic evaluative space are inconclusive at best, if not altogether misconceived.

[2] In contemporary empirical research into happiness, 'emotional well-being' is a term that is often used interchangeably with 'flourishing' as such.

we negotiate the basic evaluative space is underwritten by the fact that the EF construes our emotional experiences very differently, i.e. not as conduits of pleasure and pain, mainly, but as content-bearing states that occupy a central position in the rich nexus of mental states in terms of which human reason is defined. Once we take this fact about our emotions seriously, we are in a position to see that our emotional experiences can constitute an exercise of our rational agency in virtue of the roles they play in orientating us towards our conception of the good and shaping our practical deliberation, in enabling us to perceive value, and in anchoring our capacity for moral judgement (Section 4.5). The discussion in Section 4.6 was intended to make clear that the way we exercise our rational agency through our emotional lives is distinct from the way we are understood to exercise it through the intentional actions we perform. Recognizing this diversity in the way human reason is manifest is vital for appreciating that our emotional lives are essential to our flourishing as excellent epistemic, moral, and deliberative agents. Achieving this excellence is a distinct *telos*—distinct, that is, from the aim of seeking to reduce the occurrence of negatively valenced emotional experiences—and appreciating this plurality of possible ends helps us to see that it does in fact make an ultimate difference which way we negotiate the appraisals we make of our emotional experiences in the basic evaluative space.

Once we have EF in view, we are in a position to see that it would be a mistake—and a costly one too—to conclude on the basis of focusing exclusively on distressing affects that our emotions are something we are passive in the face of, i.e. neither something through which we exercise our rational agency nor something with respect to which we can exercise such agency. This emphasis on passivity encourages the utilization of the concept of pathology in our interpretation and response to emotional distress. But an indiscriminate concept of emotional disorder that largely (if not exclusively) tracks the valence quality of our emotional experiences threatens to obscure precisely this agential construal of our emotional lives, since the ascription of pathology status to a mental state is generally taken to displace the judgement that that mental state constitutes an exercise of our rational agency. Conversely, by apprehending our emotional lives in light of EF, we effectively place large regions of our emotional experiences outside the health-disorder spectrum altogether. While we saw that attempts have been made to expand the concept of health so that it includes a concept of 'mental health' that incorporates the good exercise of reason (recall the critical review of the 'therapeutic turn' in Section 5.3), nonetheless the fact remains that there are two distinct paradigms in terms of which we make sense of our emotional lives: that of health,

and that of rational agency. So long as this basic distinction is upheld—and I have argued in Section 5.3 that we have no good reason to eschew it—then by discerning in our emotional experiences the exercise of our rational agency, we are precisely *not* placing these emotional experiences on the spectrum of health and disorder.

But this leaves us with an obvious question: are there any emotional experiences that would count as disordered, according to the BED/Aristotelian approach? This is a question to which I shall return in Section 6.3 below. The immediate point to note is this. It is because the concept of emotional disorder is meta-evaluative (and thus not transparent) that the very same (token) emotional experience that is regarded as disordered on account of its negative valence quality may, when it is understood to bear existential significance or to be a rational response to our circumstances, be regarded as an exercise of our rational agency. The HF encourages us to adopt the former perspective on this (token) emotional experience, whilst the latter perspective is undergirded by the EF. The light in which the emotional experience is regarded therefore seems to shift in a kaleidoscopic way—one light casting us as subjects of disorder and another light as (imperfect) rational agents—depending on which of these background conceptions of flourishing we adopt. All this is to say that, when it comes to our emotional lives, it does not seem possible to carve out the domain of pathology in such a way that it can be assumed to be cleanly discrete from the domain of our rational agency, with different kinds of emotional experience falling neatly within the scope of each domain respectively. Instead, the image of crossroads seems to be a more helpful metaphor here. In our emotional lives we find ourselves standing at a multidirectional crossroad: often with regard to one and the same emotional experience (token or type), we are confronted with forking paths that lead us towards radically different ways of understanding ourselves and our situatedness—through this experience—in our intrapersonal, interpersonal, and physical worlds.

Diagram 6.1 summarizes the structure of the concept of emotional disorder that has been proposed and articulated in this book.

6.2 Revisiting the Horwitz-Wakefield and Kendler Exchange

I noted in the Introduction that resolving the dispute between Horwitz and Wakefield, on the one hand, and Kendler, on the other, may be less a matter of getting right the facts about emotional pathology, and more a matter of understanding what values are at stake in deciding whether or not an emotional

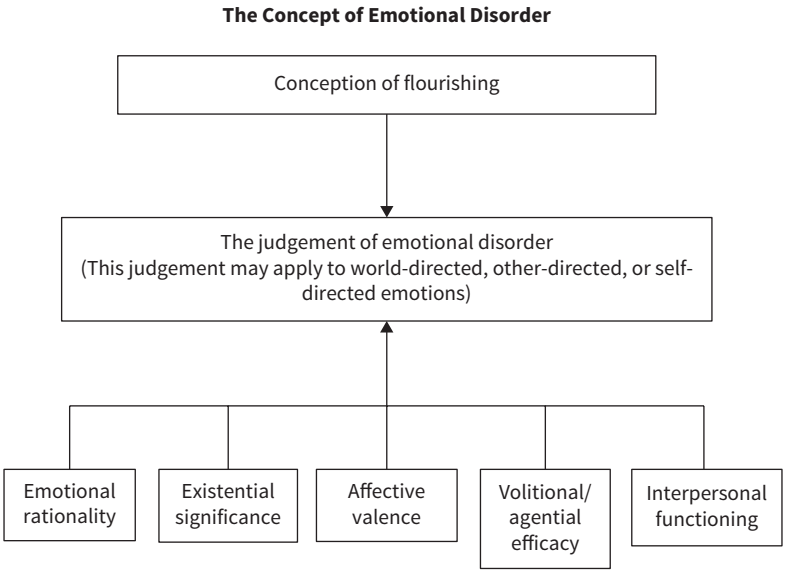

Diagram 6.1 The concept of emotional disorder

experience of a specific kind is pathological. The argument in this book has sought to elucidate this claim and to establish its truth. It is because the concept of emotional disorder is meta-evaluative that, in determining its scope, we find ourselves steering a path within a multidimensional value-laden space. We have now seen what some of these values are: those identified in the basic evaluative space, as interpreted in light of the two conceptions of flourishing, HF and EF, respectively.

Once we bring into view the EF conception of flourishing, however, something else in the exchange between Horwitz and Wakefield and Kendler becomes visible: in analysing the concept of emotional disorder, any departure point that assumptively places our emotional experiences on the health-disorder spectrum will struggle to fully register and to encompass their agential character. Horwitz and Wakefield's evolution-theoretic concept of emotional disorder and Kendler's simpler valence-based concept both share this departure point—and both therefore inherit this limitation. Kendler's concept barely registers the fact that we appraise our emotional experiences along the dimension of emotional rationality. By tracking the representational accuracy of an emotional response, Horwitz and Wakefield's evolution-theoretic approach does register this fact. However, in presuming that emotional experiences can be exhaustively characterised as biological

functions or mechanisms, the approach threatens to construe all judgements concerning the rationality (or otherwise) of our emotional responses in non-agential terms since the representational accuracy of a subject's emotional responses is taken squarely to signify the proper functioning of relevant biological mechanisms, rather than as signifying the exercise of a rational will that is partly constituted by well cultivated emotional dispositions.[3] Given that all our emotional experiences are taken by Horwitz and Wakefield to be expressions of either good biological functioning or biological dysfunction, this threatens to leave no room for us to recognize judgements of emotional rationality/irrationality, agentially construed. Both Kendler's and Horwitz and Wakefield's concepts of emotional disorder thus fail to do justice to the full range of appraisals we make of our emotional experiences, and it seems to me that the underlying cause of this limitation is that both Kendler and Horwitz and Wakefield assumptively place our emotional lives in a wholesale manner on the health-disorder spectrum.

By starting out with a focus on the *emotion* component of the concept of emotional disorder, instead of on the *disorder* component, the BED/Aristotelian approach is better placed from the outset to register the multi-faceted character of our emotional lives, principally because this approach doesn't assumptively place our emotional experiences on the health-disorder spectrum. This enables us (a) to recognize, firstly, the fact that we can and do manifest our rational agency through our emotional experiences, and, coupled with the further recognition of the fact that the concept of emotional disorder is meta-evaluative, (b) to appreciate that the task of demarcating the scope of emotional disorder will necessarily involve delineating the agential and pathological (thus non-agential) dimensions of our emotional experiences. This task of demarcation will be an essentially value-laden undertaking.

It is worth making explicit the point that the BED/Aristotelian approach seems to up-end the project of naturalising the concept of emotional disorder,[4] because the naturalistic programme presumes that all mental properties

[3] This point draws upon the distinction between capacities and virtues, which in turn alludes to a long-standing exploration of the distinction between skill and virtue (Aristotle discusses this in Book V of the *NE*, at 1140b22–5; the modern classic on this topic is Philippa Foot's 1978 paper, 'Virtues and Vices'). I have spent much time in Chapter 4 showing how we can be said to exercise our rational will through our emotional experiences (by looking both at the roles that our emotions play in the exercise of our rational agency and at some of the principal mechanisms by which they discharge these roles); but for a systematic study of the general contrast between the idea of properly functioning biological mechanisms and the idea of exercising our rational agency well, see Pestana 1998, esp. pp. 95–104. Whilst I do not endorse everything that is said in this discussion (this pertains especially to Pestana's proposed delineation of the boundary of the rational will), nonetheless Pestana's discussion helpfully maps out the logical geography of the concepts of biological mechanism and of rational agency.

[4] Boorse 1975; Wakefield 1992; Murphy 2017; Tsou 2021.

(which includes emotional properties) can be adequately characterized in terms of the concept of biological function. In order to be able simultaneously to do this whilst leaving us room to recognize judgements of emotional rationality/irrationality, agentially construed, the naturalist has to isolate a component in an emotional experience that is identified as its 'biological function', for it is only by doing so that we will be able to see this biological function as distinct from—i.e. as an enabling condition of—the emotional experience as an exercise of our rational agency. But we've seen that the naturalist runs into serious difficulties in trying to isolate a 'biological function' component in our emotional experiences (as discussed in Sections 1.9 and 5.2). These difficulties point exactly to the fact that disorderedness-as-dysfunction does not wear itself transparently on the sleeves of our emotional experiences (thus, that it is not possible to isolate it and to characterize it as a putative biological dysfunction). Rather, disorderedness is a feature it is possible adequately to ascribe to our emotional experiences only on the basis of having first negotiated the basic evaluative space. As far as the project of naturalizing the concept of mental disorder goes, this leaves the naturalist with two options:

(i) either to continue to assume that emotional disorder belongs to the same broad category of mental disorder under which all other kinds of mental disorder are also subsumed, and on this basis to hold that the concept of mental disorder cannot be naturalized, since the concept of emotional disorder is thoroughly evaluative;
(ii) or else, if relinquishing the entire programme of naturalism in psychopathology on the basis of the concept of emotional disorder alone seems hasty, the naturalist could opt to treat the concept of emotional disorder as *sui generis* and to assert that it is radically unlike the other kinds of conditions we identify as mental disorders.

The meta-evaluative nature of the concept of emotional disorder may offer us a principled reason for taking the latter option seriously, as I'll suggest in the next section.

6.3 Emotional Disorder: The Positive Story?

Approaching the compound concept of emotional disorder from the *emotion* side of this notion has enabled us to see just how complex and rich this concept really is. This complexity hints at a unique concept of disorder, which prompts

the question raised in Section 6.1 above: are there *any* emotional experiences that count as disordered, at least according to the BED/Aristotelian approach? Indeed, have we any reason to hold on to the concept of emotional disorder at all?

So far, the story has been told largely negatively, i.e. that an emotional experience that is not positively appraised in any basic evaluative dimension, and which is negatively appraised in one or more of these dimensions, is a candidate pathological state. But in virtue of what does such a state count as pathological? Given the difficulties that arise from an appeal to the notion of biological dysfunction to help answer this question, a reasonable next step would be to look at values-in analyses of the concept of disorder for an answer.[5] Consider then Bill Fulford's suggestion: pathology concepts are rooted in the notion of failures of ordinary doing (Fulford 1989). This notion, although more overtly value-laden than the appeal to biological dysfunction, still looks to be too narrow to use in the context of analysing the concept of emotional disorder, as it seems mainly to track negative appraisals made in the dimensions of agential efficacy and of emotional rationality. This limitation hints at a broader underlying difficulty, one that seems to beset both naturalistic and values-in analyses alike: they promise to offer a substantive account of disorder by identifying a transparent characteristic or feature of the conditions we judge to be disorders, a feature in virtue of which these conditions count as disordered. This approach therefore presumes that there is one single transparent feature (or a set of features) that can perform this explanatory task. But the BED analysis of the concept of emotional disorder reveals it to be opaque (on account of being meta-evaluative), not transparent. So it does not look to have a shape that will allow it to be fit into the familiar template of analysis of disorder concepts—which in turn raises the question, 'Is this so-called concept of emotional disorder a genuine concept of pathology at all?'

This last doubt arises if we presume that concepts of disorder have to be directly explanatorily substantive—that is to say, that they have to be directly explained in an informative way in terms of a set of individually necessary and jointly sufficient conditions that are taken to be definitive or prototypical of a general concept of disorder. Suppose instead that we are prepared to countenance the possibility that there are a variety of ways by which we connect particular psychological or mental conditions with the notion of genuine illness,

[5] In contrast to the naturalist outlook, proponents of 'values-in' analyses of concepts of pathology hold that these concepts are inherently value-laden and, on this basis, proceed to analyse them in overtly evaluative terms. A *locus classicus* of this approach is Fulford's 1989 book, *Moral Theory and Medical Practice*. For more recent values-in accounts of disorder, see for instance Thornton 2007, 2000; Bolton 2008, 2013; and Svenaeus 2011.

such that it isn't possible in principle to specify a universal template for subsuming particular disorder concepts under a generalised concept of disorder. This enables us to appreciate that the BED analysis may, after all, shed light on a cogent notion of emotional disorder, in spite of the fact that the concept it yields is not amenable to being directly characterized in explanatorily substantive terms.[6]

In the face of this, the following question might be pressed: why invoke the notion of disorder *at all* when it comes to our emotional experiences? Why, that is, should we connect *any* emotional experience with the notion of illness? The answer to this question lies, I believe, in the strict construal of pathology (discussed in Section 3.2), which exerts a powerful influence over our thinking. Consider, firstly, that an emotional state that is in no way positively appraised is an emotional state that is deemed not to have any rational, existential, interpersonal, or volitional significance. So it is a state that would formally satisfy the first three conditions of the strict construal of pathology: the having of such an experience would lie outwith the register of meaningfulness and in this way be external to its subject's sense of her core or real self. It would thereby also be an experience to which the subject is likely to feel passively related. If, secondly, this emotional experience is negatively appraised all around, then it would satisfy the remaining two conditions of this strict construal: being negatively valued, we would have reason to seek to alleviate or to remove it by interventionist means. A relatively clear example of such an experience might be a state of anxiety that is induced by an alcoholic hangover or by the overconsumption of caffeine (generally, however, we do not seek out medical intervention for this kind of state, given its usually temporary nature). Marked and rapid shifts from one type of mood state to another for no apparent reason at all (as seen for instance in bipolar disorder) might be another example. Depressed mood caused by hypothyroidism or by a neurodegenerative disorder would be further examples.

At this point in the inquiry, though, one might still wonder why it should be the concept of pathology specifically that we apply here. After all, isn't the strict construal of pathology itself explained by an underlying concept of disorder? Is it not a state's being biologically dysfunctional, or its being a failure of ordinary doing, that engenders the five characteristics of pathology, strictly construed? This is a distinct possibility that calls for further investigation.

[6] The claim that an emotional experience that is negatively appraised along all (relevant) evaluative dimensions merits the judgement that it is pathological is not in itself explanatorily substantive, I take it, since there is not yet any *prima facie* reason to believe that it is in virtue of some common feature—definitive of or prototypically instantiated by disorder—that the emotional experience comes to be appraised in this way.

The salient point herein would be that it seems likely that the only way we will succeed in articulating an adequate concept of emotional disorder is *indirectly*—that is, it is only by seeing that a given emotional experience is subsumable under the strict construal of the pathology concept (in the manner just sketched out) that we will be able to see how the emotional experience in question might satisfy the conditions of a general (substantive) concept of disorder. The opacity of the concept of emotional disorder makes it highly improbable that we will be able to subsume any given emotional experience directly under a substantive general concept of disorder, i.e. one that is delineated by an independent set of analytical conditions.

There is, however, another possible response to this question that cannot be ruled out. When we recall that the strict construal is a regulative assumption that makes intelligible the shape of our medical discourse(s), and not a substantive concept of disorder in its own right, it becomes easier to appreciate the possibility that there may simply be no way of subsuming our emotional experiences under a substantive general concept of disorder—not even indirectly. In light of the metaphorical crossroads at which we find ourselves standing in our emotional lives (as described in Section 6.1 above), if we are not to overlook any significant aspect of a given emotional experience, the process of arriving at the judgement of emotional disorder ideally has to involve the careful elimination of salient alternative possibilities along the various basic evaluative dimensions. If through this process of elimination an emotional experience looks to be devoid of rational (or any other) significance, and it is furthermore distressing and/or immobilising, it would be reasonable to seek to alleviate or to remove it through interventionist means. So long as we find this an intelligible move to make—so long, in order words, as it makes sense to us to view the emotional experience as pathological (strictly construed)—then perhaps this is all there is to be said about the experience as a state of disorder: it is meaningless, it is distressing and perhaps also immobilizing, and so it calls to be removed.[7]

Whichever of these two possibilities ultimately turns out to be the more viable, it seems to me that neither severs nor cuts off the concept of emotional disorder from the notion of genuine illness. For what is suggested by both possibilities is that the way in which the concept of emotional disorder is connected to the idea of being ill is *sui generis*—and so we shouldn't expect it to

[7] This possibility calls to mind Derek Bolton's (2008, 2013) suggestion that there may ultimately be no way of articulating a substantive analysis of the concept of mental disorder. The BED/EF framework presented in this book offers a principled account of why this is so, at least in the case of emotional states that are judged to be disordered.

look anything like the way concepts of physical disorder (or even concepts of mental disorder of other kinds) are thus connected. The key distinction here is between concepts of disorder that are transparent and those that are opaque on account of being meta-evaluative. Whereas it seems possible in the case of transparent concepts of disorder to directly say something further about what it is that renders a particular condition disordered, a meta-evaluative concept of disorder seems by its very nature to defy any direct substantive analysis.

I think we would lose something valuable if we eschewed altogether the framing of meta-evaluative concepts of disorder as pathological. They can after all, *ex hypothesis*, satisfy the conditions of the concept of pathology, strictly construed. It is this that connects such meta-evaluative concepts to the idea of having a (genuine) illness. But the absence of an immediate substantive explanatory substrate serves to emphasise the point that the task of judging an emotional response to be disordered is an inherently dynamic process that involves negotiating the five basic evaluative dimensions in perceptive, imaginative and highly context-sensitive ways. As I noted in Section 5.2.3, this leads naturally to the idea that diagnosing a disordered emotional state looks nothing like diagnosing a heart condition, and perhaps not even much like diagnosing psychosis. How could it, given the meta-evaluative structure of the concept of emotional disorder (the skilful application of which essentially involves navigating the innermost reaches of our subjective landscapes in subtle and discerning ways)?

This basic division between transparent and opaque/meta-evaluative concepts of disorder furnishes the naturalist about psychopathology with a principled reason for isolating the concept of emotional disorder and treating it as *sui generis*: the meta-evaluative structure of this concept renders it unlike other conditions we more familiarly identify as mental disorders, i.e. conditions that appear to be more transparently disordered. Noticing its intractability to being analysed in value-neutral terms, the naturalist may conclude, simply, that the concept of emotional disorder operates grammatically in a distinctive, perhaps even unique, way.[8] This in turn may lead the naturalist to question the legitimacy of the concept of emotional disorder itself, but this path towards scepticism is not germane to anyone who isn't a naturalist about the concept of disorder. A final thought about this dialogue

[8] How unique the concept of emotional disorder really is, will have to be determined by a more thorough and careful examination of other concepts of mental disorder. The point being made here is this: we have arrived at a position where we are able to say, in general terms, that the prospect of successfully naturalizing a given concept of mental disorder (any concept of mental disorder) will turn upon whether or not the concept is meta-evaluative. I have been exploring this idea with respect to our emotional lives, in this book; but it is an idea that may be useful to explore in other domains of psychopathology too.

with the naturalist: recalling the exchange between Kendler and Horwitz and Wakefield, it is perhaps something of an irony to learn that the most compelling way to police the boundary of the concept of disorder—at least when it comes to our emotional lives—is not by seeking out an analysis of the concept which isolates a component that eliminates all reference to human value, but rather by recognizing the fact that the concept of emotional disorder is thoroughly value saturated.

It has been observed that the construction of our psychiatric diagnostic categories with respect to human affectivity has proceeded largely without due and serious consideration of what constitutes normal human emotional experience.[9] By and large, it is theorists informed by the perspective of evolutionary biology who have stepped in to try and fill in this remarkable lacuna in our understanding of emotional pathology. While this perspective has yielded helpful insights, I have sought to argue that it does not provide us with a sufficiently comprehensive grasp of our emotional lives. A key aim of the present work has been to put forward a distinct perspective, one that seeks to reflect the fact that our appraisals of emotional experience—as normal or otherwise—seem to encompass and yet also to be broader than the biological imperatives of survival and reproduction. If the analytical framework set out in this book is even roughly right, then what it shows us is the importance of making room for our rational agency, construed non-reductively, in our worked out conception of normal human emotional experience.[10]

[9] That this gap in our understanding is problematic is expressed succinctly by Klaus Scherer: 'If we want to know why an emotional system does not function properly we should first understand *how* it functions' (Scherer and Mehu 2015: 202). We may generalize the point by saying that if we want to know why any emotional experience is deemed to be pathological, we should first understand what constitutes normal (non-pathological) emotional experience—however normality is construed.

[10] The inquiry in this book has brought to light the fact that there appears to be more than one substantive conception of normality that informs our sense of what counts as a normal emotional experience (the two conceptions we have been exploring at length are respectively the functional-theoretic conception and the excellence-theoretic conception). This plurality of conceptions may explain why, when it comes to the task of delineating the boundary of emotional pathology, it can seem so hard to achieve agreement about what the ground rules are—and hence why it can seem so hard simply to get started.

References

Abbott, E. (2010, orig. 1884). *Flatland: A Romance of Many Dimensions*, edition with notes and commentary by W. Lindgren and T. Banchoff. Cambridge: Cambridge University Press.

Aeschylus (2007). *Oresteia*, transl. I. Johnston. Arlington, VA: Richer Resources Publications.

Alvarez, A. (2002). *The Savage God: A Study of Suicide*. London: Bloomsbury.

American Psychiatric Association (APA) (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. Washington, DC: American Psychiatric Association (APA).

Anda, R. F., V. J. Felitti, J. D. Bremmer, J. D. Walker, Ch. Whitfield, B. D. Perry, Sh. R. Dube, & W. H. Giles. (2006). The enduring effects of abuse and related adverse experiences in childhood: a convergence of evidence from neurobiology and epidemiology. *European Archives of Psychiatry and Clinical Neuroscience*, 256, 174–86.

Annas, J. (1980). Aristotle on pleasure and goodness. In Oksenberg Rorty (ed.), 285–300.

Annas, J. (1993). *The Morality of Happiness*. Oxford: Oxford University Press.

Annas, J. (2004). Happiness as achievement. *Daedalus*, 133(2), 44–51.

Anscombe, E. (2000, orig. 1957). *Intention*. Cambridge, MA: Harvard University Press.

Aristotle (1980, reissued 2009). *Nicomachean Ethics*, transl. D. Ross. Oxford: Oxford University Press.

Aristotle (1991). *The Art of Rhetoric*, transl. H. Lawson-Tancred. London: Penguin Books.

Aristotle (1995). *Politics*, transl. R. F. Stalley. Oxford: Oxford University Press.

Aristotle (2013). *Eudemian Ethics*, transl. B. Inwood and R. Woolf. Cambridge: Cambridge University Press.

Aristotle (2018). *On the Soul, and Other Psychological Works*, transl. Fred D. Miller, Jr. Oxford: Oxford University Press.

Beck, A. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy*, 1(1), 5–37.

Beck, A. and B. Alford (2009). *Depression: Causes and Treatment*, 2nd edn. Philadelphia, PA: University of Pennsylvania Press.

Bentham, J. (1982, orig. 1781). *An Introduction to the Principles of Morals and Legislation*, J. H. Burns and H. L. A. Hart (eds.), London: Methuen.

Berryman, S. (2019). *Aristotle on the Sources of the Ethical Life*. Oxford: Oxford University Press.

Bolton, D. (2008). *What Is Mental Disorder? An Essay in Philosophy, Science, and Values*. Oxford: Oxford University Press.

Bolton, D. (2013). What is mental illness? In K. W. M. Fulford et al. (eds.), *Oxford Handbook of Philosophy and Psychiatry*. 434–50. Oxford: Oxford University Press.

Bonhoeffer, D. (2009, orig. 1949). *Ethics*, Dietrich Bonhoeffer Works, vol. 6, ed. Clifford J. Green, transl. R. Krauss, C. West, and D. Stott. Minneapolis, MN: Augsburg Fortress Press.

Boorse, C. (1975). On the Distinction between disease and illness. *Philosophy & Public Affairs*, 5(1), 49–68.

Broadie, S. (1991). *Ethics with Aristotle*. Oxford: Oxford University Press.

Burnyeat, M. (1980). Aristotle on learning to be good. In Oksenberg Rorty (ed.), 69–92.

Carr, D. (2009). Virtue, mixed emotions, and moral ambivalence. *Philosophy*, 84, 31–46.

Charland, L. (2004). Character: moral treatment and the personality disorders. In Radden (ed.), 64–77.

Charland, L. (2006). Moral nature of the DSM-IV cluster B personality disorders. *Journal of Personality Disorders*, 20(2) (special section: philosophy and personality disorders), 116–25.

Cooper, J. M. (1988). Some remarks on Aristotle's moral psychology. *Southern Journal of Philosophy*, 27 (supplement), 25–42.
Cooper, J. M. (1999). *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*. Princeton, NJ: Princeton University Press.
Cottingham, J. (1998). *Philosophy and the Good Life: Reason and the Passions in Greek, Cartesian and Psychoanalytic Ethics*. Cambridge: Cambridge University Press.
Cottingham, J. (2010). Integrity and fragmentation. *Journal of Applied Philosophy*, 27(1), 2–14.
Cottingham, J. (2022). *In Search of the Soul: A Philosophical Essay*. Princeton, NJ: Princeton University Press.
Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10(1), 59–70.
Cuthbert, B. and T. Insel (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine*, 11(1), 126.
Darwin, C. (1872/1999). *The Expression of the Emotions in Man and Animals*, 3rd edn with an introduction, afterword, and commentaries by P. Ekman. London: Fontana.
David, S., I. Boniwell, and A. C. Ayers (eds.) (2014). *Oxford Handbook of Happiness*. Oxford: Oxford University Press.
Davies, J. (2012). *The Importance of Suffering: The Value and Meaning of Emotional Discontent*. Hove, East Sussex: Routledge.
Davies, J. (2021). *Sedated: How Modern Capitalism Created Our Mental Health Crisis*. Bloomsbury: Atlantic Books.
Deigh, J. (1994). Cognitivism in the theory of emotions. *Ethics*, 104, 824–54.
Deigh, J. (2008). *Emotions, Values, and the Law*. New York, NY: Oxford University Press.
Deigh, J. (ed.) (2013). *On Emotions: Philosophical Essays*. Oxford: Oxford University Press.
Demazeus, S. and P. Singy (eds.) (2015). *The DSM-5 in Perspective: Philosophical Reflections on the Psychiatric Babel*. Dordrecht: Springer.
Dewey, J. (1957). *Human Nature and Conduct*. New York, NY: The Modern Library.
Dewey, J. (2004, orig. 1920). *Reconstruction in Philosophy*. Mineola, NY: Dover Publications Ltd.
Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542–75.
Dixon, T. (2012). "Emotion": the history of a keyword in crisis. *Emotion Review*, 4(4), 338–44.
Dixon-Gordon, K., J. Peters, E. Fertuck, and S. Yen. (2017). Emotional processes in borderline personality disorder: an update for clinical practice. *Journal of Psychotherapy Integration*, 27(4), 425–38.
Duff, A. (1977). Psychopathy and moral understanding. *American Philosophical Quarterly*, 14(3), 189–200.
Ekman, P. (1992a). An argument for basic emotions. *Cognition & Emotion*, 6(3–4), 169–200.
Ekman, P. (1992b). Are there basic emotions? *Psychological Review*, 99(3), 550–3.
Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4), 384–92.
Elliot, D. (2016). Defining the relationship between health and well-being in bioethics. *The New Bioethics*, 22(1), 4–17.
Elliott, C. (1996). *The Rules of Insanity: Moral Responsibility and the Mentally Ill Offender*. Albany, NY: State University of New York (SUNY) Press.
Engel, G. (1977). The need for a new medical model: a challenge for biomedicine. *Science*, 196 (4286), 129–36.
Feldman Barrett, L. and J. A. Russell (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, 74(4), 967–84.
Feldman Bettencourt, M. (2015). *Triumph of the Heart: Forgiveness in an Unforgiving World*. New York, NY: Penguin.
Filson, B. (2013). Self-injury: the attempt to cope, connect and communicate, webinar for MA Department of Public Health—Suicide Prevention Bureau and the National Association

of State Mental Health Program Directors (https://www.youtube.com/watch?v = Jfap TE0R6g0).
Foot, P. (1978). *Virtues and Vices: And Other Essays in Moral Philosophy*. Oxford: Basil Blackwell.
Foot, P. (2001). *Natural Goodness*. Oxford: Oxford University Press.
Fortenbaugh, W. W. (1964). *Transactions and Proceedings of the American Philological Association*, 95, 77–87.
Fortenbaugh, W. W. (2002). *Aristotle on Emotion*, new edn. London: Bloomsbury.
Fowers, B. J. (2016). Aristotle on eudaimonia: on the virtue of returning to the source. In J. Vitterso (ed.), *Handbook of Eudaimonic Well-Being*, 67–83. Switzerland: Springer.
Frances, A. (2010). Good grief. *New York Times*. Retrieved from http://www.nytimes.com/2010/08/15/opinion/15frances.html.
Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1), 5–20.
Frankfurt, H. (1998). *The Importance of What We Care About*. Cambridge: Cambridge University Press.
Frankfurt, H. (1999). *Necessity, Volition, and Love*. Cambridge: Cambridge University Press.
Frankfurt, H. (2004). *The Reasons of Love*. Princeton, NJ: Princeton University Press.
Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.
Frijda, N. H. (1993). The place of appraisal in emotion. *Cognition and Emotion*, 7(3–4), 357–87.
Fulford, K. (1989). *Moral Theory and Medical Practice*. Cambridge: Cambridge University Press.
Garson, J. (2019). *What Biological Functions Are and Why They Matter*. Cambridge: Cambridge University Press.
Garson, J. (2022). *Madness: A Philosophical Exploration*. Oxford: Oxford University Press.
Gibson, J. (2019). Mindfulness, interoception, and the body: a contemporary perspective. *Frontiers in Psychology*, 10, 1–18.
Glover, J. (2015). Psychiatry, folk psychology, and the impact of neuroscience—a response to Steven Hyman's Loebel Lectures. In J. Savulescu, R. Roache, and W. Davies (eds.), *Psychiatry Reborn: Biopsychosocial Psychiatry in Modern Medicine*, 304–319. Oxford: Oxford University Press.
Goldie, P. (2002). *The Emotions: A Philosophical Exploration*. Oxford: Clarendon Press.
Graham, G. (2010). *The Disordered Mind: Introduction to Philosophy of Mind and Mental Illness*. New York, NY: Routledge.
Greenspan, P. (1988). *Emotions and Reasons: An Inquiry into Emotional Justification*. New York, NY: Routledge.
Gross, J. J. and R. F. Muñoz (1995). Emotion regulation and mental health. *Clinical Psychology: Science and Practice*, 2, 151–64.
Grossman, V. (2011, orig. 1952). The Sistine Madonna. In *The Road: Short Fiction and Essays*, transl. Chandler and Chandler, 181–92. London: Quercus.
Haggard, P. (2019) . The neurocognitive bases of human volition. *Annual Review of Psychology*, 70, 9–28.
Harcourt, E. (2013). Aristotle, Plato, and the anti-psychiatrists: comment on Irwin. In K. W. M. Fulford et al. (eds.), *Oxford Handbook of Philosophy and Psychiatry*, 47–52. Oxford: Oxford University Press.
Hare, R. D. and T. J. Harpur (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, 100(3), 391–8.
Hari, J. (2018). *Lost Connections: Why You're Depressed and How to Find Hope*. London: Bloomsbury.
Herissone-Kelly, P. (2022). *Controlling Love: The Ethics and Desirability of Using 'Love Drugs'* (Cambridge Elements). Cambridge: Cambridge University Press.
Hollis, J. (2005). *Finding Meaning in the Second Half of Life*. New York, NY: Penguin Books.

Horwitz, A. and J. Wakefield (2007). *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. New York, NY: Oxford University Press.
Horwitz, A. and J. Wakefield (2012). *All We Have to Fear: Psychiatry's Transformation of Natural Anxieties into Mental Disorders*. New York, NY: Oxford University Press.
Hudson, S. D. (1981). Reason and motivation in Aristotle. *Canadian Journal of Philosophy*, 11(1), 111–35.
Hursthouse, R. (1991). Arational action. *Journal of Philosophy*, 88(2), 57–68.
Irwin, T. (1975). Aristotle on reason, desire, and virtue. *Journal of Philosophy*, 72(17), 567–78.
Irwin, T. (1992). Who discovered the will? *Philosophical Perspectives*, 6, 453–73.
Irwin, T. (2013). Mental health as moral virtue: some ancient arguments. In K. E. A. Fulford (ed.), *Oxford Handbook of Philosophy and Psychiatry*, 37–46. Oxford: Oxford University Press.
Jackson, M. and K. W. M. Fulford (1997). Spiritual experience and psychopathology. *Philosophy, Psychiatry, & Psychology*, 4(1), 41–65.
James, W. (1884). What is an emotion? *Mind*, 9(34), 188–205.
James, W. (1902). *The Varieties of Religious Experience*. New York, NY: Longmans, Green and Co.
Jaspers, K. (1913/1997). *General Psychopathology*, transl. Hoenig and Hamilton. Baltimore, MD: Johns Hopkins University Press.
Jazaieri, H., H. L. Urry, and J. J. Gross (2013). Affective disturbance and psychopathology: an emotion regulation perspective. *Journal of Experimental Psychopathology*, 4(5), 584–99.
Jung, K. (1933, republished 2001). *Modern Man in Search of a Soul*. Abingdon, Oxon: Routledge.
Kashdan, T., R. Biswas-Diener, and L. King (2008). Reconsidering happiness: the costs of distinguishing between hedonics and eudaimonia. *Journal of Positive Psychology*, 3(4), 219–33.
Kendler, K. (2005). Toward a philosophical structure for psychiatry. *American Journal of Psychiatry*, 162(3), 433–40.
Kendler, K. (2008). Book review of 'The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder' by Horwitz and Wakefield. *Psychological Medicine*, 38, 148–50.
Kendler, K. and P. Zachar (2019). *Toward a Philosophical Approach to Psychiatry*. Cambridge: Cambridge Scholars Publishing.
Kenny, A. (1963, republished 2003). *Action, Emotion, and Will*. London: Routledge.
Kierkegaard, S. (1941, orig. 1849). *Fear and Trembling* and *The Sickness unto Death*, transl. and with notes by W. Lowrie, with a new introduction by G. Marino. Princeton, NJ: Princeton University Press.
Kleinman, A. (2006). *What Really Matters: Living a Moral Life amidst Uncertainty and Danger*. Oxford: Oxford University Press.
Kleinman, A. (2012). Culture, bereavement, and psychiatry. *Lancet*, 379, 608–9.
Konstan, D. (2006). *The Emotions of the Ancient Greeks: Studies in Aristotle and Classical Literature*. Toronto: University of Toronto Press.
Kosman, L. A. (1980). Being properly affected: virtues and feelings. In Oksenberg Rorty (ed.), 103–16.
Kraut, L. (1991). *Aristotle on the Human Good*. Princeton, NJ: Princeton University Press.
Kristjánsson, K. (2015). *Virtues and Vices in Positive Psychology*. Cambridge: Cambridge University Press.
Lazarus, R. S. (1991). *Emotion and Adaptation*. New York, NY: Oxford University Press.
LeDoux, J. (2016). *Anxious: Using the Brain to Understand and Treat Fear and Anxiety*. London: Penguin Books.
Macaro, A. (2006). *Reason, Virtue and Psychotherapy*. Chichester: Wiley.
MacIntyre, A. (2016). *Ethics in the Conflicts of Modernity: An Essay on Desire, Practical Reasoning, and Narrative*. Cambridge: Cambridge University Press.

Madsen, O. J. (2014). *The Therapeutic Turn: How Psychology Altered Western Culture*. New York, NY: Routledge.
Madsen, O. J. (2018). *The Psychologization of Society: On the Unfolding of the Therapeutic in Norway*. New York, NY: Routledge.
Martin, M. W. (2006). *From Morality to Mental Health: Virtue and Vice in a Therapeutic Culture*. Oxford: Oxford University Press.
Maslow, A. (1954). *Motivation and Personality*. New York, NY: Harper & Brothers.
Matthews, E. (1999). Moral vision and the idea of mental illness. *Philosophy, Psychiatry, & Psychology*, 6(4), 373–88.
McDowell, J. (1994). *Mind and World*. Cambridge, MA: Harvard University Press.
McDowell, J. (1998). Two sorts of naturalism. In *Mind, Value, and Reality*, 167–97. Cambridge, MA: Harvard University Press.
Murphy, D. (2017). Can psychiatry refurnish the mind? *Philosophical Explorations*, 20(2), 160–74.
Natali, C. (ed.) (2009). *Aristotle's Nicomachean Ethics Book VII*. Oxford: Oxford University Press.
Nesse, R. M. (1990). Evolutionary explanations of emotions. *Human Nature*, 1(3), 261–89.
Nesse, R. M. (2000). Is depression an adaptation? *Archives of General Psychiatry*, 57, 14–20.
Nesse, R. M. (2007). Evolution is the scientific foundation for diagnosis: psychiatry should use it. *World Psychiatry*, 6, 160–1.
Nussbaum, M. (1985). The discernment of perception: an Aristotelian conception of private and public rationality. *Proceedings of the Boston Area Colloquium in Ancient Philosophy*, 1(1), 151–201.
Nussbaum, M. (2004). Mill between Aristotle and Bentham. *Daedalus*, 133(2), 60–8.
Nussbaum, M. (2008). Who is the happy warrior: philosophy poses questions to psychology. *Journal of Legal Studies*, 37(2), S81–S113.
Oakley, J. (1992). *Morality and the Emotions*. Oxford: Oxford University Press.
Oksenberg Rorty, A. (ed.) (1980). *Essays on Aristotle's Ethics*. Berkeley and Los Angeles, CA: University of California Press.
Payne, P., P. Levine, and M. Crane-Godreau (2015). Somatic experiencing: using interoception and proprioception as core elements of trauma therapy. *Frontiers in Psychology*, 6, 1–18.
Pearson, G. (2014). *Aristotle on Desire*. Cambridge: Cambridge University Press.
Pestana, M. S. (1998). *Moral Virtue or Mental Health?* San Francisco, CA: San Francisco State University Press.
Pickard, H. (2009). Mental illness is indeed a myth. In M. Broome and L. Bortolotti (eds.), *Psychiatry as Cognitive Neuroscience: Philosophical Perspectives*, 83–101. Oxford: Oxford University Press.
Pickering, N. (2006). *The metaphor of mental illness*. Oxford: Oxford University Press.
Porges, S. W. (2004). Neuroception: a subconscious system for detecting threats and safety. *Zero to Three*, 24(5), 19–24.
Porges, S. W. (2011). *The Polyvagal Theory: Neurophysiological Foundations of Emotions, Attachment, Communication, and Self-Regulation*. New York, NY: W. W. Norton & Company, Inc.
Price, C. (2015). *Emotion* (Key Concepts in Philosophy). Cambridge: Polity Press.
Putnam, D. (2001). The emotions of courage. *Journal of Social Philosophy*, 32(4), 463–70.
Radden, J. (ed.) (2004). *The Philosophy of Psychiatry*. Oxford: Oxford University Press.
Radden, J. (2009). *Moody Minds Distempered: Essays on Melancholy and Depression*. Oxford: Oxford University Press.
Ratcliffe, M. (2015). *Experiences of Depression: A Study in Phenomenology*, 1st edn. Oxford: Oxford University Press.
Rieff, P. (1966). *Triumph of the Therapeutic*. New York, NY: Harper & Row.

Roberts, R. C. (2013). *Emotions in the Moral Life*. Cambridge: Cambridge University Press.
Russell, J. and L. Barrett (1999). Core affect, prototypical emotional episodes, and other things called *emotion*: dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5), 805–19.
Ryan, R. and E. Deci (2001). On happiness and human potentials: A review of research on hedonic and Eudaimonic Wellbeing. *Annual Review of Psychology*, 52, 141–66.
Sankowski, E. (1977). Responsibility of persons for their emotions. *Canadian Journal of Philosophy*, 7(4), 829–40.
Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44, 695–729.
Scherer, K. and P. Ekman (eds.) (1984). *Approaches to Emotion*. New York, NY: Psychology Press.
Scherer, K. and M. Mehu (2015). Normal and abnormal emotions—the quandary of diagnosing affective disorder [special issue]. *Emotion Review*, 7(3), 201–3.
Schliesser, C. (2008). *Everyone Who Acts Responsibly Becomes Guilty: Bonhoeffer's Concept of Accepting Guilt*. Westminster: John Knox Press.
Schriber, C. et al. (2017). Dispositional contempt: a first look at the contemptuous person. *Journal of Personality and Social Psychology*, 113(2), 280–309.
Schwartz, B. and K. Sharpe (2006). Practical wisdom: Aristotle meets positive psychology. *Journal of Happiness Studies*, 7, 377–95.
Seligman, M. (2002). *Authentic Happiness*. London: John Murray Press.
Seligman, M. (2005). "Positive psychology, positive prevention, and positive therapy," in C.R. Snyder and S.J. Lopez (eds.), 3–9.
Sellars, W. (1956a). *Science, Perception, and Reality*. London: Routledge and Kegan Paul.
Sellars, W. (1956b). *Empiricism and the Philosophy of Mind*. Reprinted in W. Sellars (1956a). London: Routledge and Kegan Paul.
Sherman, N. (1989). *The Fabric of Character: Aristotle's Theory of Virtue*. Oxford: Clarendon Press
Sherman, N. (1997). *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge: Cambridge University Press.
Sibson Ayob, G. (2019). Delusions and personal autonomy. *Journal of Applied Philosophy*, 36(5), 737–54.
Snyder, C. R. and Lopez, S. J. (eds.) (2005). *Handbook of Positive Psychology*. Oxford: Oxford University Press.
Solomon, R. (1993). *The Passions: Emotions and the Meaning of Life*. Indianapolis, IN: Hackett Publishing.
Solomon, R. (2004). *In Defense of Sentimentality*. Oxford: Oxford University Press.
Solomon, R. (2006). *The Passions: Philosophy and the Intelligence of Emotions*. The Great Courses. Chantilly, Virginia: The Teaching Company.
Solomon, R. (2008). *True to Our Feelings: What Our Emotions Are Really Telling Us*. Oxford: Oxford University Press.
de Sousa, R. (1988). *The Rationality of Emotion*. Cambridge, MA: MIT Press.
Spalding, T., J. Stedman, and C. Gagne (2019). *The Human Person: What Aristotle and Thomas Aquinas Offer Modern Psychology*. Cham: Springer.
Stark, S. (2001). Virtue and emotion. *Nous*, 35(3), 440–55.
Stein, D. (2021). *Problems of Living: Perspectives from Philosophy, Psychiatry, and Cognitive-Affective Science*. Amsterdam: Elsevier Academic Press.
Stocker, M. (1982). Responsibility especially for beliefs. *Mind*, 91(363), 398–417.
Styron, W. (2001). *Darkness Visible*. London: Vintage.
Sugarman, J. (2007). Practical rationality and the questionable promise of positive psychology. *Journal of Humanistic Psychology*, 47(2), 175–97.

Svenaeus, F. (2011). Illness as unhomelike being-in-the-world: Heidegger and the phenomenology of medicine. *Medical Health Care and Philosophy*, 14, 333–43.
Sweeney, A., B. Filson, A. Kennedy, L. Collison, and S. Gillard (2018). A paradigm shift: relationships in trauma-informed mental health services. *BJPsych Advances*, 24, 319–33.
Taylor, C. C. W. (2006). *Aristotle Nicomachean Ethics Books II–IV: Translation and Commentary*. Oxford: Oxford University Press.
Taylor, C. (2007). *A Secular Age*. Cambridge, MA: Harvard University Press.
Thompson, R. A. (1994). Emotion regulation: a theme in search of a definition. Monographs of the Society for Research in Child Development, 59(2–3), 25–52.
Thornton, T. (2000). Mental illness and reductionism: can functions be naturalized? *Philosophy, Psychiatry and Psychology*, 7, 67–76.
Thornton, T. (2007). *Essential Philosophy of Psychiatry*. Oxford: Oxford University Press.
Thornton, T. (2022). *Mental Illness* (Cambridge Elements). Cambridge: Cambridge University Press.
Tiberius, V. (2013). Recipes for a good life: Eudaimonism and the contribution of philosophy. In A. Waterman (ed.), *The Best within Us: Positive Psychology Perspectives of Eudaimonia*, 19, 1st edn. Washington D.C.: American Psychological Association.
Tolstoy. L. (1882, republished 2005). *A Confession*, transl. A. Maude. Mineola, NY: Dover Publications Inc.
Tomkins, S. S. (1981). The quest for primary motives: biography and autobiography of an idea. *Journal of Personality and Social Psychology*, 41(2), 306–29.
Tomkins, S. (1984). Affect theory. In Scherer and Ekman (eds.), 163–95.
Travis, C. (2004). The silence of the senses. *Mind*, 113, 57–94.
Tsou, J. (2021). *Philosophy of Psychiatry* (Cambridge Elements series: Philosophy of Science). Cambridge: Cambridge University Press.
Urmson, J. O. (1988). *Aristotle's Ethics*. Oxford: Basil Blackwell.
Valtonen, J. and B. Lewis (2023). The brain disorders debate, Chekhov, and mental health humanities. *Journal of Medical Humanities*, 44, 291–309.
Wakefield, J. (1992). The concept of mental disorder. *American Psychologist*, 47(3), 373–88.
Wakefield, J. (2007). The concept of mental disorder: diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, 6(3), 149–56.
Wakefield, J. (2015). The loss of grief: science and pseudoscience in the debate over DSM-5's elimination of the bereavement exclusion. In Demazeus and Singy (eds.), 157–78.
Wakefield, J. and S. Demazeux (eds.) (2016). *Sadness or Depression? International Perspectives on the Depression Epidemic and Its Meaning*. Dordrecht: Springer.
Wakefield, J, D. Wasserman, and J.A. Conrad (2020). Neurodiversity, autism, and psychiatric disability: The harmful dysfunction perspective. In Adam Cureton, and David T. Wasserman (eds.), *The Oxford Handbook of Philosophy and Disability*. Oxford: Oxford University Press.
Waring, D. (2012). The virtuous patient: psychotherapy and the cultivation of character. *Philosophy, Psychiatry, & Psychology*, 19(1), 25–35.
Waring, D. (2016). *The Healing Virtues: Character Ethics in Psychotherapy*. Oxford: Oxford University Press.
Waterman, A. (1993). Two conceptions of happiness: contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, 64(4), 678–91.
Waterman, A. (2007). On the importance of distinguishing hedonia and eudaimonia when contemplating the hedonic treadmill. *American Psychologist*, 62(6), 612–13.
Weiner, N. (1993). *The Harmony of the Soul: Mental Health and Moral Virtue Reconsidered*. New York, NY: State University of New York (SUNY) Press.

Williams, B. (1973). Morality and the emotions. In *Problems of the Self: Philosophical Papers 1956–1972*, 207–29. Cambridge: Cambridge University Press.

Wolf, S. (2007). Moral psychology and the unity of the virtues. *Ratio*, 20(2), 145–67.

Wynn, M. (2005). *Emotional Experience and Religious Understanding: Integrating Perception, Conception and Feeling*. Cambridge: Cambridge University Press.

Zachar, P., M. First, and K. Kendler (2017). The bereavement exclusion debate in the DSM-5: a history. *Clinical Psychological Science*, 5(5), 890–906.

Index

For the benefit of digital users, indexed terms that span two pages (e.g., 52–53) may, on occasion, appear on only one of those pages.

Tables and figures are indicated by *t* and *f* following the paragraph number