IN PURSUIT OF REASON

AN

ESSAY ON RATIONALITY AND EMOTION

SAMUEL HUGHES

SELWYN COLLEGE, CAMBRIDGE

THIS THESIS IS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

MARCH 2020

Declaration

- This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text.
- It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.
- It does not exceed the prescribed word limit for the Philosophy Degree Committee.

Thesis Details

Title	In Pursuit of Reason: An Essay on Rationality and Emotion
Candidate	Samuel Hughes
Department	Philosophy
College	Selwyn
Submitted for	PhD in Philosophy

Abstract

Since the publication of Anthony Kenny's *Action, Emotion and Will*, there has been a consensus that emotions involve representing their objects as mattering to the subject in a certain way. All major contemporary theories aim to accommodate this characteristic, and those that clearly cannot have generally been abandoned. This essay is an investigation of a related but distinct characteristic of emotion, one that has often been mentioned but that is only beginning to be systematically investigated. This is that emotions are candidates for rationality: it is unreasonable to be angry with someone who is blameless, to envy someone who is wretched, or to be outraged about something that is unproblematic. In this essay, I offer an account of exactly what this rationality consists in, I argue that existing theories of emotion cannot wholly accommodate it, and I develop a theory that can. Chapter One is a critique of the currently ascendant perception theories, and Chapter Two is a critique of the formerly ascendant belief theories. Chapter Three outlines a positive account of emotions, which I characterise as a form of 'non-doxastic cognitivism'. Chapters Four and Five look at some more specific features of emotion, namely their hedonic character and their potential for sentimentality.

Acknowledgements

I am grateful to my primary supervisor, Angela Breitenbach, for her support, criticism, advice and patience throughout this project, and to Rae Langton, Thomas Schmidt and Paulina Sliwa for valuable supervisory support at different times. I have also benefitted greatly from the mentorship of Andrew Pinsent and the late Roger Scruton, and from the friendship of Ralph Weir and Mikołaj Sławkowski-Rode. From each of them I have learnt much, and I am suitably grateful. In addition to those already mentioned, special thanks are due to Stefan Riedener, James Hutton, Maxime Lepoutre, Karamvir Chadha and Razvan Sofroni for comments, criticism and advice at various stages. I should also like to thank Benedikt Göcke, Alexander Norman, Nicholas Boys Smith and Stephen Priest for some remarkable opportunities over the last four years. This research was made possible by support from the Arts and Humanities Research Council, for which I am grateful. I thank Alistair Swiffen for organising the Cambridge Development and Training Partnership. I am also grateful for a grant from the Kurt Hahn Foundation, for hospitality from the Humboldt University of Berlin and the University of Notre Dame, and for the support of Selwyn, my college, and David Smith, my college tutor. Finally I thank my parents, John Hughes and Jenny Parkes, for all that they have given me over these many years.

Contents

Introduction 6

1. Perception Theories 11

    1.1 Emotion as Representational Bodily Feeling 12

    1.2 Bodily Feelings and Constitutive Rationality 15

    1.3. Limitations of the Argument 19

    1.4 Constitutively Rational Perceptual States? 23

    1.5 Perceptualist Responses 29

    1.6 Concluding Remarks 37

2. Belief Theories 41

    2.1 The Basic Problem 43

    2.2 A Sophisticated Belief Theory? 49

    2.3 Concluding Remarks 54

3. Towards a Theory of Emotion 55

    3.1 An Emotion Theory of Emotion? 57

    3.2 Emotion as Understanding? 59

    3.3 Emotional Complexity 61

    3.4 Recalcitrance 64

    3.5 Animal Emotion 69

    3.6 Helm's Account 75

    3.7 Concluding Remarks 78

4. Emotion, Pain and Pleasure 79

    4.1 Existing Approaches to Emotional Pain 79

    4.2 The Oblivion Problem 83

    4.3 Concluding Remarks 85

5. Sentimentality 87

    5.1 Tanner's Account 88

    5.2 Sentimentality and Misrepresentation 90

    5.3 A Motivation-Based Account 92

    5.4 Concluding Remarks 97

Conclusions 98

Bibliography 102

Introduction

Until the publication of Anthony Kenny's *Action, Emotion and Will*, the nature of emotion was relatively little studied in the analytical philosophical tradition. If there existed an orthodox view at all, it was that which had been developed by William James and Carl Lange in the late nineteenth century, according to which emotions are feelings of bodily states.[1] Kenny famously argued that this cannot be the whole truth about emotions, for the simple reason that while feelings of bodily states are at most about our bodily states, emotions are about objects in the world: if I admire Julius Caesar, my admiration is directed at the Roman general, and not at the hairs on the back of my neck.[2] Kenny's observation laid the foundation for a great range of theories that aim to accommodate this by characterising emotions as special cases of mental states that are representational in this way, or as hybrids of several mental states including at least one that is like this. For some years, the dominant view characterised emotions as partly or wholly beliefs; more recently, philosophers have tended to characterise emotions on the model of perceptual states.

In this essay I am going to focus on a related but distinct feature of emotion. This is the fact that our emotions are normally what one might call *candidates for rationality*. We criticize emotions as preposterous, sentimental, unwarranted, callous, excessive, self-indulgent or superficial; reflecting on the rationality of one's emotional reactions is an important part of moral maturity, and dialogue about the rationality of our mutual emotional reactions is an important part of social life. This feature has often been mentioned in the literature, and it may seem so familiar as to be uninteresting. But there are puzzles here that have not been systematically investigated. What is it for a mental state to be rational or irrational? Are there different ways in which this can be the case? If so, which of them is applicable to emotion? Are all emotions candidates for rationality in the same way? The way in which we answer these questions has implications for the central question in philosophy of emotion, namely the question of what emotions are. If emotions are bodily feelings, they can only be candidates for rationality in the ways that bodily feelings are candidates for rationality; if they are beliefs, only in the way that beliefs can be; and so on. If it transpires that emotions can be rational in some way that bodily feelings or beliefs cannot be, then accounts of emotions as constituted wholly by those mental states must fail. This line of enquiry has seldom been methodically pursued by philosophers of emotion, but, as we shall see, it is rich in consequences.

---

[1] Carl Lange, 'The Mechanism of Emotions' in *The Classical Psychologists,* ed. Rand Benjamin (Boston: Houghton Mifflin, 1912 [1885]) and William James, 'What is an Emotion?' *Mind* 9:34 (1884): 199-205.

[2] Anthony Kenny, *Action, Emotion and Will*, 2nd edn (London: Routledge, 2003 [1963]), esp. chs. 2 and 9. Kenny was influenced by Franz Brentano, *Psychologie vom empirischen Standpunkt* (Leipzig: Duncker und Humblot, 1874), vol. 1, esp. pp. 101-120, who was in turn influenced by Scholastic philosophy of mind.

Chapters One, Two and Three of this essay are a systematic investigation of these themes. I distinguish several ways in which emotions can be candidates for rationality, and assess whether the various reductions of emotion proposed in the literature are able to accommodate them. The picture that emerges is complex. Emotions are indeed candidates for rationality, and so, I argue, cannot be wholly constituted by perceptual experiences; but the conditions under which emotions are rational are different to those of belief, emotions cannot be beliefs either. The picture that emerges is what I call 'non-doxastic cognitivism', the view that emotions do have a constitutive aim, but that it is not the same as that of belief. In Chapters Four and Five I fill out this picture, looking at how we should understand the phenomena of emotional pain and pleasure, and at the nature of sentimentality.

One preliminary before we proceed to a more detailed outline. If James grievously wronged me many years ago, it might be true to say of me that I am angry with James (or that I am 'still angry', 'will always be angry' etc.), regardless of whether I am actually entertaining my anger towards him at this very moment. But there is another sense in which it might truly be said that I might spend only a very small proportion of my time being angry with James, and nearly all of it feeling other emotions. The usual terminology for this description in the literature is that the latter sense of 'angry' is 'occurrent' and the former is 'dispositional': the idea is that occurrent anger is temporally determinate, starting say at 4:17 p.m. and lasting until 4:27 p.m., while dispositional anger is constituted by something like a disposition to have occurrent anger under certain conditions, along perhaps with related dispositions such as acting aggressively under certain circumstances. This terminology is perhaps regrettable: after all, on the prevalent functionalist theory of mind, occurrent mental states are themselves ultimately constituted by dispositions to behave in certain ways under certain conditions, and the project of philosophy of emotion is that of identifying whether the dispositions in question are those that constitute perceptual states, bodily feelings, beliefs, or something else. But the terminology has nevertheless become entrenched, and I accept it here. Like most philosophers of emotion, my primary interest is in occurrent emotion. So when I refer to emotion unqualifiedly, it will be to occurrent emotion that I am referring.

*****

The easiest way to understand the ways in which emotions are candidates for rationality is by looking at how some existing approaches are not able to accommodate them. I begin in Chapter One with the currently popular family of accounts according to which emotions are perceptual states, or, alternatively, are 'analogous to' perceptual states, though not precisely identifiable with them. This is a broad category, especially given the 'analogous to' clause. I argue, however, that by focussing on some particular accounts we can identify a general problem with any account plausibly characterisable as a perceptual theory, and indeed with some that are not. In Section 1.1, I outline one of the most influential theories of recent years, that of Jesse Prinz. Prinz believes that our emotions are bodily sensations that evolved

to convey information to us about the nature of our environment, and which, given certain accounts of representation and perception, thus constitute perceptual experiences of that environment. In responding to Prinz in Section 1.2, I distinguish two broad ways in which mental states can be rational. One is that they can be mental states that we have reason to cause ourselves to have: we could plausibly have reasons of this kind for Prinz's representational bodily sensations. The other kind is what I call 'constitutive rationality'. Defining this precisely is best left to Chapter One, but the rough idea is easily illustrated. When a whimsical billionaire offers me a great fortune in return for believing that butterflies are lizards, there is an obvious sense in which it becomes rational for me to acquire this belief, but another in which its irrationality is completely unaffected. 'Constitutive rationality' is the latter kind of reasonableness: speaking rather inexactly, one might say that a constitutively rational mental state is an intrinsically rational mental state, rather than a mental state that it is rational to cause oneself to have. I argue that emotions, like beliefs and unlike bodily feelings, have a 'constitutive aim' in virtue of which they are candidates for constitutive rationality: if the whimsical billionaire were to offer one a fortune in return for feeling anger towards all butterflies, then it might become 'extrinsically' rational to cause oneself to be angry with all butterflies, but it would remain constitutively irrational, whereas only the first sort of rationality is applicable to a toothache. Because of this, I argue, representational bodily feelings of the kind that Prinz is concerned with could at most only be a component of emotion: the main representational element of emotion is something else.

We have, then, identified a desideratum for theories of emotion, which is that they be consistent with the fact that emotions have a constitutive aim. The kind of perceptual experience with which Prinz seeks to identify emotion does not have a constitutive aim, so it does not satisfy this desideratum. The general question for perception theories, then, is whether there are kinds of perceptual experience that (a) have constitutive aims, and that (b) have the same constitutive aims as emotion. In Section 1.4 I consider this question. I argue that there is a plausible (though far from universally accepted) account of perceptual experience which yields a positive answer to (a), namely the Kantian account according to which experience itself is the product of judgement. Although to my knowledge nobody has proposed identifying emotion with perceptual experience of this kind, it is possible to reconstruct a perception theory that does so, drawing on the construals theory of Robert Roberts. However, I argue, even on the Kantian theory the answer to (b) turns out to be negative. Part of the constitutive aim of emotion is that it be accurate. If one has good evidence that somebody is not admirable or contemptible, then it is not constitutively rational to admire or despise them, however admirable or contemptible they may seem. I argue that this is not true of perceptual experience: there are perceptual illusions that, though obviously inaccurate, we are under no responsibility to correct. This is true on any plausible theory of perception, including the best-case Kantian one I have reconstructed. Emotions therefore cannot be wholly constituted by perceptual experiences, although it is possible that at least some emotions might be partly constituted by them. I end the chapter in Section 1.5 by looking at some of the ways that

perceptualists have sought to respond to this sort of concern, concluding that none of the attempted responses is successful.

The paradigm of a mental state with a constitutive aim that includes accuracy is belief, and the considerations advanced in Chapter One might lead us to think that we should return to what used to be the orthodoxy in philosophy of emotion, namely some kind of belief theory. In Chapter Two I investigate this possibility. In Section 2.1 I argue that the belief theories that have standardly been presented in the literature, according to which emotions are constituted either by an evaluative belief or by the combination of an evaluative belief with some other elements like feeling or desire, encounter important problems in connection with testimony. In particular, while testimony can always be sufficient for a constitutively rational belief, it is often insufficient for a constitutively rational emotion: it would be constitutively irrational for me to admire someone on the basis of reliable testimony that they are admirable, although it would be perfectly rational for me to believe them to be admirable. With a few possible exceptions, emotions seem to involve some grasp of what *grounds* the evaluative property in question: to admire someone, one must have some kind of grasp on what there is to admire.

One possibility is that this 'grasp' is itself constituted by beliefs: this would yield a sophisticated belief theory, on which emotions are constituted by evaluative beliefs conjoined with beliefs about what grounds the evaluative property in question (and perhaps also feelings, desires and so on). In Section 2.2 I consider this possibility at length. I argue that while a sophisticated belief theory of this kind *might* work for some emotions, there are others that it cannot accommodate. These include many of the emotions felt towards objects in virtue of their aesthetic properties. I argue that to occurrently have many of these emotions requires that we be experiencing their objects at the same time. For me to be experiencing an artwork's beauty, it is not enough that I can remember propositions about what grounds it: I must be experiencing it, at least in imaginative recollection. This suggests that the grasp we have on the ground of the artwork's beauty cannot be constituted by belief, because beliefs are not tied to occurrent experiences in this way. To put the basic thought more intuitively, our emotions about artworks seem to be inextricably fused with our experiences of them, in such a way that we cannot characterise them in the terms of the belief theory. I go on to argue that if an implausibly fractured account of emotion is to be avoided, this must be true of other emotions too.

In the opening sections of Chapter Three, I review what we have learnt about emotion. Emotion cannot be wholly constituted by perceptual experience, or by any other mental state does not have a constitutive aim that includes accuracy. Nor can it be constituted by belief, for its constitutive aim also goes beyond accuracy in a way belief's does not. The true theory of emotion must therefore fall in a logical space I call 'non-doxastic cognitivism', the view that emotion is subject to standards of constitutive rationality, but that it is not belief. These conclusions cast doubt on the possibility of realising one of the central aspirations of much theory of emotion to date, namely that of characterising emotion as being constituted by other, putatively better understood, mental states: the reason for this it is simply not clear there are any other mental states whose constitutive aim includes but goes beyond

accuracy. One option is thus to endorse an 'emotion theory of emotion', according to which emotion can be described in various ways, but cannot ultimately be reduced to any other type of mental state. I discuss this possibility in Section 3.1. Another option is that emotion might after all be constituted by a form of *understanding*, a possibility I discuss in Section 3.2. In the remaining sections of the chapter, I develop and defend this picture. In Section 3.3 I look at some complexities surrounding the kind of information that emotion involves grasping. Sections 3.4 and 3.5 are devoted to revisiting some of the problems that faced classical cognitivism, including the problems arising from the existence of animal and infant emotion, and of recalcitrant emotion, that is, emotions that seem to persist against one's better judgement. I argue that my account has resources to resist these objections, sometimes because those objections were unsuccessful in the first place, and sometimes because of ways in which my account is different from traditional belief theories. In Section 3.6 I look at Bennett Helm's hard-to-categorise account, arguing that although Helm develops a uniquely cogent critique of existing theories of emotion, his own positive account does not succeed in avoiding the problems he himself raises.

In Chapters Four and Five, I look in more detail at some specific features of emotion. In Chapter Four, I examine emotion's hedonic character. Many philosophers, especially those with theories more similar to my own, have held that the hedonic character of emotion is itself 'content-constituted': the painfulness of grief, say, is a matter of the badness of one's loss, and one could in principle settle the question of whether one had the pain associated with grief by settling questions about whether one had lost someone or something important to one. It would be convenient for me if this were true, because it would cohere more elegantly with the cognitivist account of emotion developed in the preceding chapters. Unfortunately, however, it is not: I argue that although emotions certainly may have 'content-constituted' hedonic features, they also have 'content-distinct' ones. To see why, I consider the behaviour of a bereaved person who drinks herself to oblivion to avoid, she says, the painfulness of her grief. If painfulness is constituted just by the judgement that one has lost something, then this behaviour seems hard to make sense of: no adult could think that ending one's consciousness of a loss actually undoes it. The behaviour is only explicable if painfulness is not content-constituted. We are left with the curious, but I think unavoidable, conclusion that the representation of loss involved in grief is logically separable from grief's painfulness.

In Chapter Five I look at a particularly interesting way in which emotion can be rationally defective, namely sentimentality. I begin in Sections 5.1 and 5.2 with an examination of existing accounts of sentimentality. On most existing theories, sentimentality involves misrepresenting the object to provide a pretext for some kind of emotional indulgence. I argue that although sentimentality certainly often involves misrepresentation, this is not strictly necessary: there is nothing to stop one having sentimental emotions about things that really matter to one in the way the emotion represents them as mattering. What makes an emotion sentimental, I argue in Section 5.3, is simply how it was formed: an emotion is sentimental just if it was caused in a certain way by the subject's desire for it. I conclude by exploring an analogy to self-deception, and discussing sentimentality's moral significance.

Chapter One

Perception Theories

The most popular accounts of emotion in the recent literature can be loosely grouped as 'perception theories'.[3] Perception theories characterise emotions either as perceptual states, or as something 'analogous to' perceptual states. Like perceptual states, emotions represent objects, and like perceptual states, emotions are not factive: one can have an emotion whose content is inaccurate.[4] Such accounts have attracted philosophers who want to accommodate this possibility of representing or misrepresenting external objects, but who think that belief theories of emotion are unable to deal with recalcitrant emotions or the emotions of animals and children.[5]

Given the range of things that can count as perceptual states, the category of perception theories is a broad one. The 'analogous to' clause makes it broader still, in ways I return to discuss in Section 1.4. It can also have hazy boundaries. Since there is no philosophical consensus on the nature of perception,[6] it is sometimes unclear of a given mental state with which some philosopher proposes to identify emotion whether it is a perceptual state or not. Happily, however, my argument is structured in such a way that it will not be necessary to settle the question of what counts as a perceptual state. I am going to argue that most emotions have a characteristic that it is pretheoretically clear that perception does not have. So it applies very generally to any theory that successfully qualifies as a perception theory,

---

[3] For some leading perception theories, see Sabine Döring, 'Explaining Action by Emotion', *The Philosophical Quarterly* 53:211 (2003): 214–30; Robert C. Roberts, *Emotions: An Essay in Aid of Moral Psychology* (Cambridge: Cambridge University Press, 2003); Julien A. Deonna, 'Emotion, Perception and Perspective', *Dialectica* 60:1 (2006): 24–49; Christine Tappolet, *Emotions, Values and Agency* (Oxford: Oxford University Press, 2016); Jesse J. Prinz, *Gut Reactions: A Perceptual Theory of Emotion* (Oxford: Oxford University Press, 2004); Cheshire Calhoun, 'Cognitive Emotions?' in Robert C. Solomon (ed.), *What is an Emotion? Classic and Contemporary Philosophers on the Emotions* (Oxford: Oxford University Press, 1983). An arguable perception theory is Ronald de Sousa, *The Rationality of Emotion* (Cambridge Mass: MIT Press, 1987). This is far from exhaustive: further references will be given as we go on.

[4] Or, like apparent perceptual states, if disjunctivism is true. I drop the qualification, which will not be important, in what follows.

[5] For these criticisms, see e.g. John Deigh, 'Cognitivism in the Theory of Emotions', *Ethics* 104:4 (1994): 824-854; Paul E. Griffiths, 'The Degeneration of the Cognitive Theory of Emotions', *Philosophical Psychology* 2:3 (1989): 297-313; Bennett W. Helm, *Emotional Reason: Deliberation, Motivation and the Nature of Value* (Cambridge: Cambridge University Press, 2000), pp. 36-46; Roberts, *Emotions*, pp. 83-106; Claire Armon-Jones, *Varieties of Affect* (New York: Harvester Wheatsheaf, 1991), ch. 2.

[6] For a lack of consensus, see the essays collected in John Hawthorne and Tamar Szabo Gendler (eds), *Perceptual Experience* (Oxford: Oxford University Press, 2006).

including those included only by dint of the 'analogous to' clause. It probably applies to quite a few other theories, too, a point that I touch on in the conclusion.

This chapter will run as follows. In Section 1.1 I explain one particular perception theory, that of Jesse Prinz. I choose Prinz's theory partly because it is especially well-known, but mostly because my objection to perception theories affects it especially clearly, making it helpful for illustrative purposes. I develop this objection in Section 1.2. In Section 1.3 I highlight some of the limitations of my argument, noting especially that there are mental states like moods and reflexes about which it shows little. In Section 1.4 I apply my objection to other perception theories. I do this by creating a 'best case' perception theory, that is as well-fitted as it could possibly be to respond to my objection while remaining a perception theory. This is based on Robert Robert's influential 'construals theory', but optimises it for the purposes of responding to my objection. I argue that even this best-case perception theory is unsuccessful, and that the prospects for perception theories in general are therefore poor. In Section 1.5 I look at a number of responses that perception theorists have offered to something like the problem I raise, arguing that none is successful.

1.1 Emotion as Representational Bodily Feeling

Let us turn, then, to Prinz. Prinz's account is, among other things, an attempt to vindicate the James-Lange theory of emotion, to which I alluded in the introduction.[7] The James-Lange account begins with the claim that the perception of certain sorts of situations typically cause our bodies to react in certain ways: the perception of a dangerous situation, say, causes a quickened pulse, dilating pupils and blood rushing to the extremities. The next, and for our purposes the crucial, claim is that emotion just is the feeling of these bodily states: fear, for example, would be constituted by the feeling of a quickening pulse, dilating pupils, and so on. As mentioned above, the James-Lange theory was generally thought to be inconsistent with Kenny's insight that emotions are about our external environments: the feeling of my pulse quickening seems to be, at most, a representation of the quickening of my pulse, not a representation of the object in the world that caused this to happen. For this reason, it was largely abandoned by philosophers in the later twentieth century, although it continued to attract interest in the empirical sciences.[8]

Prinz's project is to vindicate the James-Lange theory by offering an account of how the bodily sensations that James and Lange described could be representations of our environment after all.[9] The

---

[7] Lange, 'The Mechanism of Emotions' and James, 'What is an Emotion?'.

[8] See e.g. Antonio Damasio, *Descartes' Error: Reason, Emotion and the Human Brain* (New York: Putnam, 1994).

[9] This is developed in his *Gut Reactions,* esp. ch. 3; see also including 'Emotions Embodied' in *Thinking About Feeling*, ed. Robert Solomon (Oxford: Oxford University Press, 2004) and 'Are Emotions Feelings?', *Journal of Consciousness Studies* 12:8-10 (2005): 9-25. 998). For some objections to Prinz's theory, which do not, however, question the

way in which these bodily feelings are representational also makes them a kind of perceptual experience, meaning that Prinz's account is *both* a perception theory and a revival of the old bodily feelings theory. Prinz begins by invoking Fred Dretske's naturalistic account of representation. On Dretske's view, representations are a subset of *information-carrying states of affairs*, or, for short, 'indications'.[10] One state of affairs is an indication of another just if they consistently co-occur. If masses of smoke rise from a town, for instance, this is an indication that the town is burning. What an indication needs in order to be a representation is that it must also be possible for it to be a *mis*representation: although it is possible for the smoke rising from the town to indicate something that is not the case (the townspeople might be burning bonfires), it is impossible for the smoke to be *inaccurate*, and this seems to preclude its being a representation of its object. What distinguishes representations from mere indications, Dretske proposes, is that they have the *function* of carrying certain information. In the most obvious case, this function is a matter of something's having been *designed* for that purpose: topographical maps are designed to carry certain information about landscapes, ground plans are designed to carry certain information about buildings, and so on. Requiring that all representations be designed would however be inconsistent with the possibility of representational mental states, because, at least on a naturalistic picture, mental states were not designed at all. Instead, Dretske thinks, the function of mental states should be understood in evolutionary terms: representational mental states like beliefs and perceptual experiences have the function of carrying information because it was in virtue of their capacity to do this that they conveyed a survival advantage, and hence in virtue of this that evolution 'selected' them.

If this Dretskean account of mental representation is right, Prinz argues, it turns out that the bodily states described by the James-Lange theory are representations after all. The bodily states that James and Lange identify with emotion reliably co-occur with certain kinds of important situations ('core relational themes'), and the feelings of those bodily states reliably co-occur with those bodily states.[11] This means that these feelings also reliably co-occur with facts about our environment, thereby carrying information about both the environment and our bodily states. Their evolutionary function, however, must be to carry information about our environment, rather than about our bodily states, since learning

---

features primarily criticized here, see Peter Goldie, 'Review: Gut Reactions: A Perceptual Theory of Emotion', *Mind* 115:458 (2006): 453-457 and Paul E. Griffiths, 'Review: Gut Reactions: A Perceptual Theory of Emotion', *British Journal of the Philosophy of Science* 59:3 (2008): 559-567. Prinz has recently distanced himself from his account in *Gut Reactions* and expressed an interest in an 'enactivist' theory, on which I comment briefly in Section 1.6. See Daniel Shargel and Jesse Prinz, 'An Enactivist Theory of Emotional Content' in Hichem Naar and Fabrice Teroni, *The Ontology of Emotions* (Cambridge: Cambridge University Press, 2018).

[10] See Fred Dretske, *Knowledge and the Flow of Information* (Cambridge: MIT Press, 1982) and 'Misrepresentation' in *Belief: Form, Content and Function,* ed. Radu Bogdan (Oxford: Oxford University Press, 1986).

[11] Prinz, *Gut Reactions,* pp. 52-60. The evidence Prinz adduces that each emotion has a distinctive phenomenology independently of its content seems to me less than conclusive, but I shall not press this point here.

about the latter has no obvious survival value.[12] Hence what our feelings *represent* is the sort of situation we are in, although they do this by means of carrying information about bodily states. Prinz claims to have thereby accommodated the evaluative representational content of emotion within the James-Lange theory of emotion, as well as vindicating our common-sense talk of 'gut reactions'. Prinz calls these representational bodily feelings 'perceptions' because they have what he argues to be the central features of paradigmatic perceptual experiences, including the facts that they take place in sensory systems, that they generate internal representations, and that they can be consciously experienced.[13]

Presented thus, Prinz's account may seem to have a conspicuous problem, namely that it makes extremely contentious presuppositions about mental content. Dretske's account, though influential, is only one in a crowded field of theories of mental content, and one that has faced cogent objections.[14] If Prinz's theory really does entail Dretske's, then those who believe the latter to be false should believe Prinz's theory to be false too, and those who are undecided about mental content should at best afford Prinz's theory tentative interest. It is of course no part of my project to show that Prinz's theory is faultless, but it is nevertheless worth saying that Prinz has undersold himself here. Although his account of emotion is not consistent with all theories of mental content, it is consistent with many more than one; or to put this another way, if there are indeed bodily feelings that correlate with certain core relational themes in the way he describes, many theories of mental content will yield the result that those feelings represent those themes. Consider for instance one of the main rivals of the Dretskean account, Jerry Fodor's 'asymmetric dependence' theory.[15] On this view, my mental representation of a dog is a representation of a dog in virtue of the facts that (1) it is caused by dogs and (2) it can be caused by things other than dogs (foxes, wolves etc.) only because of its disposition to be caused by dogs - that is, one is capable of taking a fox for a dog only because one has a more fundamental disposition to take dogs for dogs. Fodor's account would, so far as I can see, give the same results as Dretske's concerning Prinz's gut reactions: certain physical feelings represent danger because they are caused by danger, and because they can be caused by non-dangerous situations only due to those situations' superficial

[12] Prinz, *Gut Reactions,* pp. 66-7.

[13] See esp. Jesse J. Prinz, 'Is Emotion a Form of Perception?', *Canadian Journal of Philosophy* 36 (2006): 137-160.

[14] e.g. Donald Davidson, 'Knowing One's Own Mind', *Proceedings and Addresses of the American Philosophical Association* 60 (1987): 441-458. A well-known problem: Dretske's theory seemingly implies that if it turned out that our mental states had not developed through evolution, as it presumably could, this would mean that they represent nothing, despite appearances to the contrary.

[15] Fodor develops various permutations of this account in, *inter alia,* his *Psychosemantics* (Cambridge MA: MIT Press, 1987) and *The Elm and the Expert* (Cambridge MA: MIT Press 1994).

resemblance to dangerous ones. It seems to me that something similar may be true of several other leading theories, although I shall not argue this here.[16]

Something more should be said in Prinz's support, which is that it is surely pretheoretically plausible that our 'gut feelings' tell us something about the situation we are in. Even if there are defects in the story Prinz tells, we might therefore suspect that there is *some* account that could be given of how this happens. This might be one reason why Prinz's account has had so much influence. In what follows, in any case, I accept for the sake of argument that there might be bodily feelings that represent core relational themes. What I want to focus on is, rather, whether or to what extent these feelings could be what constitutes our emotions.

1.2 Bodily Feelings and Constitutive Rationality

Let us turn, then, to investigate whether Prinz's account can accommodate the ways in which emotions can be rational or irrational. It will be helpful to start with an illustration. Suppose the eccentric billionaire offers to donate a fortune to curing neglected tropical diseases if I form the belief that pigs can fly. In doing this, he gives me a powerful reason to form this belief: plausibly, there is a sense in which he gives me overall reason to form this belief. But there is obviously another sense in which a belief that pigs can fly is irrational, because it is not supported by the available evidence. Speaking paradoxically, we might say that this is an 'irrational belief that it would be rational to have'.[17] Interestingly, it is not going to be at all easy for me to meet the billionaire's condition. I cannot acquire the belief at will, and I cannot in any normal way reason myself into it. To acquire it, I am going to have to somehow subvert or circumvent my normal rationality, perhaps by employing a hypnotist or through some very elaborate process of self-deceptive reasoning.[18]

Epistemologists normally describe this kind of situation by distinguishing between 'constitutive' and 'extrinsic' reasons. A constitutive reason is a reason in virtue of which the mental state tends to

---

[16] Examples include, I suspect, the accounts of Ruth Millikan, *Language, Thought and Other Biological Categories* (Cambridge MA: MIT Press, 1984) and Daniel Dennett, *The Intentional Stance* (Cambridge MA: MIT Press 1987).

[17] Examples of this kind have been much discussed. See canonically Justin D'Arms and Daniel Jacobson, 'The Moralistic Fallacy: On the "Appropriateness" of Emotions', *Philosophy and Phenomenological Research* 61:1 (2000): 65–90 and Roger Crisp, 'Review: Value... And What Follows', *Philosophy* 75:3 (2000): 452-462.

[18] For a rich discussion, see Richard Moran, *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton: Princeton University Press, 2001), esp. ch. 5; see also his 'Frankfurt on Identification'. The distinction has a long history in the post-Kantian tradition; see esp. Jean-Paul Sartre, *L'être et le néant: Essai d'ontologie phénoménologique* (Paris: Gallimard, 2001 [1943]), part 1, ch. 2.

fulfil its 'constitutive aim'. In the case of belief, this is, at a minimum, truth.[19] What this means is that it is part of what it is for a mental state to be a belief that it aims at truth: if it did not aim at truth, it would not be a belief. Reasons that are not constitutive are extrinsic. The nature of this 'aim' is an important controversy in contemporary epistemology. Some philosophers think that it is a matter of intention: very roughly, to believe that *p* just is to represent *p* as true and to intend to represent *p* as true only if *p* is true.[20] Others think that it is a matter of a norm: very roughly, to believe that *p* just is to represent *p* as true and to be obliged to represent *p* as true only if *p* is true.[21] Philosophers in these groups have correspondingly differing accounts of why it is normally impossible to consciously believe something solely on the basis of extrinsic reason. On an intention-based account, this is because consciously representing *p* as true despite consciously knowing *p* is false means that one lacks the intention that is constitutively necessary for belief, namely to represent *p* as true only if it actually is. On a norm-based account, this is because to believe *p* is to be normatively committed to representing *p* as true only if it is, so that any subject who understands the concept of belief must recognise that the question of whether to be believe *p* can only be settled by reference to the question of whether *p* is true. Fortunately, we do not need to settle this controversy here: either kind of account of constitutive aims is consistent with the present argument. We need only to remember that constitutive reasons are reasons in virtue of which a mental state tends to fulfil its constitutive aim, and that acquiring a belief on the basis of extrinsic reasons is difficult and requires some kind of bypassing of normal reasoning processes.

It is clear that we can have both extrinsic and constitutive reasons for emotions.[22] Suppose the billionaire offers to donate the fortune not in return for my believing in flying pigs, but rather in return

[19] A number of philosophers have argued that belief aims at something more than truth, the most common proposal being knowledge. See esp. Timothy Williamson, *Knowledge and its Limits* (Oxford: Oxford University Press, 2000). This would not affect the present argument, and I set it to one side for clarity of exposition.

[20] Asbjørn Steglich-Petersen, 'No Norm Needed: On the Aim of Belief', *Philosophical Quarterly* 56:225 (2006): 499-516; David Velleman, 'On the Aim of Belief' in David Velleman (ed.), *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000).

[21] This is the more common view. See Peter Boghossian, 'The Normativity of Content. *Philosophical Issues* 13:1 (2003), 31–45; Nishi Shah, 'How Truth Governs Belief', *Philosophical Review* 112:4 (2003), 447–482; Allen Gibbard, 'Truth and Correct Belief', *Philosophical Issues* 15 (2005): 338-350; Nishi Shah and David Velleman, 'Doxastic Deliberation', *Philosophical Review* 114 (2005): 497-534

[22] That we have reasons for or against emotions has often been noticed in the literature, and the examples given are sometimes clearly constitutive, though the distinction between extrinsic and constitutive reasons has seldom been explicitly drawn. Brentano, incisive as ever, touches upon the issue in his *Vom Ursprung sittlicher Erkenntnis* (Leipzig: Duncker & Humblot, 1889), pp. 75-76. In the heyday of the 'cognitivism' about emotion, constitutive rationality, or something very like it, was mentioned fairly often, e.g. in Martha Nussbaum, *Upheavals of Thought* (Cambridge: Cambridge University Press, 2001), pp. 46-47; Robert Solomon, 'Emotions as Judgements', *American*

for my feeling angry at a teacup. In so doing he gives me strong extrinsic reasons for being angry at the teacup, but there remains a sense in which such anger would be irrational, namely that I have no constitutive reasons for it. If we can have constitutive reasons for emotion, then emotion must have a constitutive aim. We shall have much to say about this aim in Chapter Two, but for the time being we should note only that at least *part* of that aim seems to be something like truth or accuracy: an emotion is irrational if we have good evidence that its object does not possess the evaluative property that the emotion involves representing it as having. A theory of emotion needs to characterise emotion in a way that is consistent with this: it must characterise emotion as the kind of mental state for which we can have constitutive reasons, that is, the kind of state that has a constitutive aim.

The question we need to address now is: does Prinz's account do this? The answer to this is best addressed by looking at a parallel case. Suppose I have a kind of toothache that reliably co-occurs with my tooth's being infected, and whose evolutionary function is to inform me about this infection. On Prinz's view of representation, as on many others, this pain represents my tooth as infected. If I anomalously have the toothache without there actually being anything wrong with the tooth, then the toothache misrepresents its object. Furthermore, we can certainly have extrinsic reasons for or against toothaches, to the extent that we are able to control them. The most obvious extrinsic reason we have against toothaches is their painfulness of toothaches, but in the case of the inaccurate toothache, it might be that this inaccuracy is part of an extrinsic reason against having them too. We can imagine a doctor saying: 'Normally I would say that this pain is a useful way of reminding you not to put any pressure on a damaged tooth, and that you should therefore not take painkillers. But since this tooth is actually not damaged in any way, taking painkillers is certainly right for you'.

What seems less clear, however, is that the inaccuracy of a toothache is a *constitutive* reason against having it. As mentioned above, it is difficult to acquire a belief if one does not believe oneself to have constitutive reason for it: one has to circumvent one's normal rationality in some way, as with the hypnotist. This is simply untrue of a toothache that misrepresents the tooth. Suppose the eccentric billionaire has a device that induces the painful feeling of a damaged tooth without actually damaging the tooth in any way. I could operate this device and suddenly feel a toothache, without any of the self-

---

Philosophical Quarterly 25:2 (1988): 183-191 at p. 183; Gabriele Taylor, 'Justifying the Emotions', *Mind* 84:1 (1975): 390-402; Justin Oakley, *Morality and the Emotions* (London and New York: Routledge, 1992), ch. 1; cf. also Patricia Greenspan, *Emotions and Reasons* (London: Routledge, 1988), chs. 1 and 6. It has received less attention in recent years, but Bennett Helm has noted that it constitutes a problem for perception theories. See esp. his *Emotional Reason*, pp. 41-46 and 'Emotional Recalcitrance: Re-Evaluating the Perceptual Model', *Dialectica* 69:3 (2015): 417-433; cf. Hagit Benbaji, 'How is Recalcitrant Emotion Possible?', *Australasian Journal of Philosophy* 91:3 (2013): 577-599. I shall discuss Helm's arguments in Sections 1.5 and 3.6.

deception or bad faith involved in acquiring a belief I know to be false.[23] This indicates that toothaches, though they may be inaccurate, and though their inaccuracy may be an extrinsic reasons against them, cannot be constitutively irrational. It should be stressed that this is hardly a surprising conclusion: consider how strange it sounds in normal English to speak of an 'irrational' or an 'unjustified toothache'.

Return now to Prinz's account of emotion. Prinz thinks that emotions are bodily feelings whose evolved function is not only to carry information about the body, but also to carry information about the environment. Although external situations, physical events and physical sensations must reliably co-occur in order for the sensations to represent the external situations, this reliability does not have to be absolute—it is perfectly possible for the physical events and sensations associated with a given situation to occur anomalously without it. Indeed, according to the evidence Prinz cites, this actually happens sometimes.[24] Hence it is possible that, say, upon my friend's approach I will feel those sensations typically associated with the approach of someone hateful and dangerous. On Prinz's view, this is the same thing as hating and fearing my friend.

Now, as we have seen, Prinz is able to say that this mental state is a misrepresentation. He is also able to say that we have extrinsic reasons to cause it to go away, if we are so able: such inaccurate bodily feelings might well consume attention better engaged elsewhere, or be unpleasant to have, or have troublesome motivational effects. But is he able to say that it is constitutively irrational? And here the answer is no. This is just the same sort of mental state as the toothache was - a bodily feeling that represents an object as having a certain feature in virtue of its having the function of carrying this information. All that has changed is the information in question. As we saw, toothaches cannot be constitutively rational or irrational, because they do not have constitutive ends. The same follows in the case of Prinz's emotions. And as we have also seen, this is the wrong result: hating my innocent friend *is* constitutively irrational. This point can be drawn out by returning to the idea of self-deception. Intending to become angry at a friend or frightened of a teacup has, I submit, a similar paradoxical character to that which intending to believe in flying pigs does. Just as one can only form a belief that one knows to be false by somehow circumventing one's rationality, one cannot cause oneself to have an emotion without doing something similar. But as we have seen, this is not true of bodily feelings: one can cause an inaccurate bodily feeling in oneself without any kind of self-deception. If hatred for my friend were just the churning of the stomach and such like, then provided I had the means, I could induce myself to hate my friend while knowing my friend to be blameless without self-deception of any kind.

It is important here to be clear about exactly which feature of Prinz's view is implausible. The problem is not the claim that the physical sensations that are associated with (and perhaps partly

[23] Cf. local anaesthetics during an operation: here the *lack* of pain is inaccurate, since in fact the body part in question is being damaged, but no self-deception is involved.

[24] Prinz, *Gut Reactions,* p. 58.

constitute) a given emotion should sometimes occur even when we know that that emotion would be inappropriate. There is nothing at all implausible about that: we can perfectly well imagine that, for instance, some peculiar injection should cause one to feel the coursing blood and raised hairs associated with hatred at the moment of one's friend's arrival. Nor is it implausible that those feelings represent 'core relational themes': although I am not committed to this claim, I have no objection to the idea that gut reactions represent the world as being a certain way. Nor, finally, is it necessarily implausible that bodily feelings are a component of emotion, or perhaps even a *necessary* component. I have said nothing yet concerning that thesis, although I shall discuss it in section 1.3 below. The real problem with Prinz's view is that the whole emotion, including *all* its representational elements, has been identified with these feelings, thereby implying that I have no constitutive reasons against hating my friend, or, to put this differently, that my hatred of my friend is not constitutively irrational. What is suggested by this, I submit, is that even if bodily sensations can represent evaluative features of our environment, and even if emotion is partly constituted by such sensation, there must be more to the representational elements of emotion than this: in particular, emotion must be at least partly constituted by one of those kinds of representation for which we can have constitutive reasons.

This concludes my basic argument against Prinz's account: in the next section I shall examine some of this argument's limitations, and in Section 1.5 I shall examine some of the ways that perception theorists have responded to this sort of objective. Before moving on, however, I want to introduce some terminology to ease exposition in what follows. Referring to 'the kind of mental state for which we can have constitutive reasons' is cumbersome, so instead I will usually refer to them as states that are 'subject to reason' or 'rational states'. In the case of the last term the intended contrast is not of course to irrational states, but to non-rational ones: it is only in the case of what I am calling rational states that the question of whether a given mental state is rational or irrational is intelligible.

1.3 Limitations of the Argument

So far we have looked only at anger, and it is worth looking at the rational character of other emotions. Consider a standard list of basic emotions: aesthetic experience, anger, anxiety, compassion, depression, envy, fear, gratitude, guilt, happiness, hope, jealousy, love, pride, relief, sadness, and shame.[25] In the

[25] This list is from Richard S. Lazarus and Bernice N. Lazarus, *Passion and Reason: Making Sense of Our Emotions* (Oxford: Oxford University Press, 1994). Other writers have offered shorter or longer lists. Cf. e.g. Paul Ekman, 'Are There Basic Emotions?', *Psychological Review* 99 (1992): 550-552 and the later longer list in his 'Basic Emotions' in Tim Dalgleish and Mike Power (eds.), *Handbook of Cognition and Emotion* (New York: Wiley, 1999). 'Basic emotions' are supposed to be emotions from which all other emotions are constituted. Many researchers argue that they do not exist: nothing in my account presupposes that they do, and if not, the above list can just be treated as a list of some common emotions.

cases of aesthetic experience, compassion, envy, fear, gratitude, guilt, hope, jealousy, love, pride, relief and shame it seems clear that our reactions to thought experiments parallel to the one given above for anger are the same: these emotions have constitutive aims, and in a case like the one above, they would be constitutively irrational. If one hopes for a certain party to win the election, only to discover that the leaders of that party secretly harbour repugnant plans that make them much worse than their opponents, it would be irrational to continue hoping for their victory. If one is grateful to someone for having helped one's grandmother with the shopping, only to learn that they were trying to win her trust in order to swindle her, one's gratitude becomes irrational. In each of these cases, Prinz's account looks no more plausible than it did in the case of anger.

In the cases of depression, happiness, sadness and anxiety, however, matters are less clear. Each of these mental states seems to be distinguishable into two variants. One has the same sort of object as other emotions. I can be happy that my friend has recovered from his illness, sad that my brother failed his exam, and so on. The constitutive standard here seems to be just the same as those in the cases of other emotions: my happiness for my friend becomes constitutively irrational if I learn his condition has taken a new turn for the worse, my sadness about my daughter becomes constitutively irrational if I learn that a remark has raised her to a much higher result. It seems, however, that these mental states also occur in another variant, which does not have an object in the same way as other emotions. I can be happy, sad, depressed and perhaps anxious without being able to say what exactly I am happy, sad, depressed or anxious about. These feelings are typically distinguished as 'moods'.

Are moods candidates for constitutive rationality? On the whole, the answer seems to be *no*. When one is in a depressed mood, one is certainly disposed to find things bad: thoughts of the 'Oh, everything is so awful', or 'Why must my manager be such a ghastly man?' variety come readily. When one is in a happy mood, one is correspondingly prone to make more charitable and optimistic judgements of people and situations. All these judgements are certainly candidates for constitutive rationality. But it seems to be possible to be in a bad mood without making any of these judgements: it is surely possible to be fully aware that there is nothing wrong, that one's feeling of exhausted gloom is purely the consequence of imprudent lunchtime drinking, and yet for the mood to continue. It seems, furthermore, that there is nothing constitutively irrational about this mental state, so long, of course, as one resists the temptation of forming unjustifiably bleak judgements about one's life. Even an ideally rational person might sometimes feel upbeat or downbeat for no reason.

I am not going to give any further argument for this conclusion. If readers disagree with it, that is all to the good, because the falsity of this conclusion would remove an obstacle to my integrating moods into my theory of emotion.[26] If it is right, however, then so far as the present argument is

[26] Judgement theories of moods have been offered by Richard Lazarus and Robert Solomon, who believe moods to be judgements about 'one's existential condition' or 'the world as a whole' respectively. See Richard Lazarus, 'The Stable and Unstable in Emotion' in Paul Ekman and Richard J. Davidson (eds), *The Fundamental Nature of*

concerned, Prinz might be right about moods, which he characterises in the same way as he characterises emotions, but with more general objects.[27] In fact, Prinzianism about moods might actually be quite plausible. Moods, like representational bodily feelings, plausibly represent one's situation as having a certain core relational theme; both are often caused by events that correspond to that theme (for instance, a good mood by some good news), but both can also occur independently of such events; neither has a fine-grained object beyond its core representational theme; and neither, I think, is a candidate for constitutive rationality. In other words, 'gut feelings' might be rather a good model for understanding at least some moods, but not for understanding paradigm emotions. I should be clear, however, that I am not committed to this. There are other theories of moods that meet the same desiderata,[28] or that may have other advantages that Prinz's lacks.[29] My only official claim in this section is the concession that my main argument of this chapter does not clearly tell against Prinzianism about moods. Here and hereafter I am officially neutral on what moods are constituted by.

Another point at which my argument seems to lose force is the mental state of being startled, inelegantly nominalised in the literature as 'startle'.[30] Suppose one is winding up a jack-in-the-box. One knows exactly what will happen, and one knows that the jack is completely harmless. Yet it is perfectly normal to be startled when it bursts out, and it seems strained to describe this as irrational. Nor indeed would it be irrational if someone with very strong nerves were *not* startled by it. Judgements of rationality and irrationality just do not apply here: startle appears to be a sort of mental reflex. So startle seems to be another case in which Prinzianism could be right, so far as the present argument is concerned.[31] Perhaps there are other states of mind that are like this too. Two possible examples are shock and

---

Emotion (Oxford: Oxford University Press, 1994) and Robert C. Solomon, *The Passions* (South Bend IN: University of Notre Dame Press 1976), p. 133-5.

[27] Prinz, *Gut Reactions*, pp. 182-188.

[28] e.g. Carolyn Price's theory that moods are 'states of vigilance' or Achim Stephan's interesting theory that there are several kinds of dispositional mood 'layered' in our 'manifest mood'. See Carolyn Price, 'Affect Without Object: Moods and Objectless Emotions', *European Journal of Analytical Philosophy* 2:1 (2006): 49-68 and Achim Stephan, 'Moods in Layers', *Philosophia* 45 (2017): 1481-1495.

[29] e.g. the view that emotions are objectless, defended by in Peter Goldie, *The Emotions*: *A Philosophical Exploration* (Oxford: Oxford University Press, 2000), esp. ch. 6; Eric Lormand, 'Towards a Theory of Moods', *Philosophical Studies* 47 (1985): 385-407, Nico Frijda, 'Varieties of Affect: Emotions and Episodes, Moods and Sentiments' in Paul Ekman and Richard J Davison (eds), *The Nature of Emotion: Fundamental Questions* (New York: Oxford University Press, 1994); Laura Sizer, 'Towards a Computational Theory of Mood', *British Journal for the Philosophy of Science* 51 (2000): 743-769.

[30] For discussion, see esp. Jenefer Robinson, 'Startle', *The Journal of Philosophy* 92:2 (1995): 53-74.

[31] Note that though being startled obviously has much in common with being surprised, there seems to be a distinction: paradigmatic surprise occurs when something happens that one did not expect, but of course one knows exactly what is going to happen when one winds up a jack-in-the-box.

nausea, at least in certain forms. I will not discuss whether they should be categorised this way, but it is at least possible that there are several more members of the 'mental reflexes' group.

A terminological note is due here. Moods are sometimes spoken of as kinds of emotion, and sometimes as distinct from emotions. If it is true, as I have tentatively suggested here, that they are in fact different in at least one important way from other emotions, then we are left with a terminological choice: either we say that the analysis I shall go on to give is an analysis of those emotions that are not moods, 'non-mood emotions', or we say that it is an analysis of emotions and not of moods. The same is true of the mental reflexes. Nothing substantive rides on this. I am going to refer to the core group of non-mood and non-reflex emotions as 'emotions', because it is less cumbersome. We have found that this group is a reference magnet, being distinguished from its neighbours by its being the sort of mental state that is a candidate for constitutive rationality. And there seems to be no alternative to the word 'emotion' for referring to them. So this seems to be a sensible way to talk.[32]

There is a second dimension in which we might look for limits to the argument. I have argued that emotion must be constituted at least in part by a mental state which is a candidate for constitutive rationality. The question then arises: is it the sole component of emotion; the sole *necessary* component of emotion; or just one of a number of necessary components? The answer is that the argument so far has not ruled out any of these options: it is perfectly possible that emotions are candidates for constitutive rationality while having components that would not, by themselves, be like this. Indeed, given what we have said in this section, this might seem to be quite plausible. We have found here that at least two groups of mental states that are pretheoretically continuous with emotion, moods and mental reflexes, are not candidates for constitutive rationality. If that is right, we need some way to explain their 'pretheoretical continuity' - why do they seem to be similar sorts of mental states, if one group is rational and the other is not? An attractive answer to this would say that emotions, moods and startles have some shared component - a representational bodily feeling, perhaps - and that emotions have, in addition to this, a more cognitive component that makes the candidates for constitutive rationality.

For the time being, I am going to leave that as an interesting possibility, without making a serious argument for it. In Chapter Three, we are going to see some further reasons why positing that emotions have further non-rational components is plausible. Finally in Chapter Four I will present an argument showing that many or all emotions must include non-rational hedonic components. But for the time being, my official position is just to leave this possibility open. Emotions are definitely partly constituted by a kind of mental state that is a candidate for constitutive rationality. They might well have other components, but until Chapter Four, we cannot be sure.

---

[32] One further area in which we might want to explore the scope of this argument is the emotion of animals and infants. I look at this at length in Section 3.5, so for the time being I am just going to issue a promissory note.

It is worth saying that, if this *is* true, then in one sense my picture turns out to be fairly similar to Prinz's, or at least to that of the Neo-Prinzian that I invented in Section 1.1, who thinks that rational mental states accompany and are causally necessary for all emotions. The Neo-Prinzian and I would both claim that, where there is emotion, there is a rationally assessable mental state as well as some kind of feeling. The difference is that I claim that both are necessary components of the emotion, whereas Prinz claims that the emotion is wholly constituted by the feeling. Nevertheless, viewed from a certain angle, our positions are not radically at odds.

1.4 Constitutively Rational Perceptual States?

It may at this point be objected that Prinz's account is atypical among perception theories in identifying emotions with bodily feelings, albeit with bodily feelings of a special kind. Most perception theorists identify emotion with other kinds of perception, or indeed, as noted above, with states of mind that are only 'analogous to' perceptual states. It does not follow from Prinz's kind of perception's being non-rational that all kinds of perception are, let alone all states that are 'analogous to' perception. Indeed, given that bodily feelings are in some vague intuitive way more primitive than, say, vision, we might think this is rather plausible.

What I therefore propose to do in this section is to reconstruct what I take to be that perceptual theory that is best-placed to accommodate the reason-governed character of emotion. This is an account of emotion that (a) identifies it with a mental state analogous to perception and (b) characterizes this mental state in such a way that it is a candidate for constitutive rationality. I shall argue that although there is a reasonably plausible account of perceptual experience on which it is a candidate for constitutive rationality, the rationality conditions of perceptual experience *even so characterized* are not the same as those of emotion, and that the identification of the two is therefore unsuccessful. This section thus has a slightly odd task: we will be discussing a perception theory that nobody has actually expounded, but that would be ideally placed for dealing with my objection; and I am going to argue that even this theory is not successful.

The question for the perception theorist is: are there any accounts of perceptual experience according to which at least some kinds of perceptual experience are subject to norms of constitutive rationality? And the answer to this question is yes. The clearest example is the Kantian theory of perceptual experience, represented in the analytical tradition by philosophers like Peter Strawson and John McDowell.[33] On this theory, my experience of a table as a table involves a judgement, in which I

[33] Peter Strawson, 'Perception and its Objects' in his *Perception and Identity: Essays Presented to A.J. Ayer,* ed. G.F. McDonald (London: Macmillan, 1979) and John McDowell, *Mind and World* (Harvard MA: Harvard University Press, 1994). There is a much wider family of views according to which experience is 'cognitively penetrated'. Cognitive penetration is consistent with experience's being a candidate for constitutive rationality, but it does not entail it: a cognitive penetration theorist might hold that beliefs causally influence the content of perceptual

bring the object under the concept of a table. It follows that one could point to features of the object in virtue of which it is or is not properly subsumed under a given concept: if the interlocutor accepted these reasons, then that acceptance would constitutively change the character of his or her perceptual experience of the object. I might be experiencing an object as an abstract sculpture, and you might then point out to me that the four protuberances are in fact legs and that the curious disk is a back, at which point I might suddenly perceive the object as a chair, produced by an avant-garde designer. What exactly this Kantian judgement is lies beyond our present concern, but its crucial feature is that the experience that results from it is understood as having been justified, not merely caused, by your observations. My experience of the object as a chair could thus, on this view, count as a constitutively rational mental state. In this case, we can temporally distinguish experiences before and after the judgement in question, but this is exceptional: in most cases, a 'pre-judgement' experience is only logically distinguishable.

Some people think that accounts like these are implausible. That is fine for me: if so, then this is not a plausible way for perception theories to address the constitutive rationality problem. But suppose *arguendo* that something like the Kantian theory is right, at least about some kinds of perceptual experience. An interesting option for the perception theory would be to identify emotion with perceptual experiences thus characterised. To my knowledge, this has not yet been attempted. But is at least one perception theory that seems to teeter on the edge of this, namely Robert Roberts's well-known account of emotions as 'construals'.[34] A construal, as Roberts characterises it, is a way of experiencing an object in terms of a certain feature, where this 'experiencing as' constitutively alters the character of the experience itself. The simplest examples are objects that are devised to illustrate this feature of perception, like the duck-rabbit or the Rubin Vase, which can be seen alternately as a vase and or as two faces in profile: Roberts thinks that this is just a special case of a pervasive phenomenon in perceptual experience: what makes it special is that there are two possible construals between which one can shift, whereas normally there is only one. What distinguishes emotion, on Roberts's view, is that emotions are *concernful* construals, that is, that the object is construed in terms of some concern that the subject has. When I fear a bear, I construe it as dangerous, experiencing its powerful claws, merciless eyes and so forth as collectively constituting its dangerousness; when I am grateful to a friend, I construe her as having given me something that mattered and that I was not owed, seeing her in terms of those

experience but that perceptual experience is itself non-rational. For cognitive penetration, see N.R. Hanson, *Patterns of Discovery* (Cambridge: Cambridge University Press, 1958) and Paul M. Churchland, *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press, 1979).

[34] Roberts initially outlined his view in 'What is an Emotion: A Sketch', *Philosophical Review* 97:2 (1988): 183-209, and has defended it at length in *Emotions: An Essay in Aid of Moral Psychology* and *Emotions in the Moral Life* (Cambridge: Cambridge University Press, 2013). See also his useful discussion of animal emotions, 'The Sophistication of Non-Human Emotion' in *The Philosophy of Animal Minds,* ed. Robert W. Lurz (Cambridge: Cambridge University Press, 2009).

features; and so forth. Roberts is keen to emphasize that we can concernfully construe objects in memory and imagination, and even ideas and arguments: perception, he thinks, does not necessarily mean *sense* perception.[35]

I think it is obvious that there are similarities between construals as Roberts describes them and perceptual experience on the Neo-Kantian model: in both cases, experience is constituted by subsuming an object under concepts. It seems to me, therefore, that it is open to Roberts to take advantage of the rationality conditions that perceptual experience has on the Neo-Kantian account in order to accommodate the fact highlighted in Section One that emotions are candidates for constitutive rationality. In fact, Roberts explicitly rejects this possibility.[36] I believe this means that Roberts's account suffers from the same underlying problem as Prinz's; in Section 1.5, I return to look at Roberts's rebuttal of this objection, arguing that it is unsuccessful. But since our principal interest here is not with Roberts's particular account, but with reconstructing the perception theory of emotion best-placed for dealing with the constitutive reasons objection, let us try modifying Roberts's theory, understanding construals as a form of judgement which is constitutively rational if the object of the judgement, as experienced by the subject, falls under the concept in question. Would this modified construals theory (the 'Neo-Robertsian theory') be able to accommodate the rationality conditions of emotion?

In many cases, the answer is yes. In fact, in many cases, the Neo-Robertsian theory seems to get a lot right. Remember, the distinctive feature of this theory is that emotions *are* judgements, but judgements that partly constitute our perceptual experience - we might say that they are *woven into* our experience, or that they *colour* and *imbue* it. Suppose we have just met Black. I find Black aloof and standoffish, but I am mistaken: in fact he is only shy and nervous. You may point out features of Black's behaviour that I overlooked, facts about the context of which I was unaware, or interpretations of things that Black said that I had not thought of. On the Neo-Robertsian theory, it may well become the case that I have constitutive reason to construe Black differently: I ought to have something like a *Gestalt* shift, in which I switch from viewing Black with resentment and view him with solicitude instead, suddenly seeing his behaviour 'in a different light'. This is not just a matter of *believing* something different about Black: to judge differently, when the judgement is of this kind, is to experience differently.

I think there is something that seems right about this, and I am going to come back to it later in this essay. But, for the purposes of the present chapter, nothing rides on there being something promising about this, because I am also going to argue that, as it stands, the Neo-Robertsian theory cannot be right. The reason is that the rationality conditions of perceptual experience, even on this generous account, are different to those of emotion. Although some inaccurate perceptual experiences

---

[35] For this and other nuances of Roberts's rich and sophisticated account, see esp. *Emotions,* pp. 69-83.

[36] Roberts, *Emotions*, pp. 83-106. There are, though, times where he seems to hint at a more Kantian sort of account, such as in his 'Propositions and Animal Emotions', *Philosophy* 71:275 (1996): 147-156 at p. 152.

are fairly plausibly characterised as constitutively irrational, given something like the Neo-Kantian account, there are others that are unambiguously not like this: there are, very clearly, inaccurate perceptual experiences that are *not* constitutively irrational, which even the Neo-Kantian is going to have to accommodate as such. Many perceptual illusions are like this. If I experience a stick as bending when it enters water, my experience is inaccurate, but it is not constitutively irrational: there is no sense in which I ought to have a different experience.[37] In a like fashion, the experience that a colour-blind person has of a surface as green does not become constitutively irrational when he or she learns that the surface is in fact red. Trying to persuade him or her to perceive a different colour would just reveal a misunderstanding of what colour-blindness involves.[38] This can be drawn out further by returning to the theme of self-deception mentioned in Section 1.2. I can experience perceptual illusions intentionally without the need for any kind of self-deception: none of the strains and paradoxes that arise in connection with intentionally having a perceptual illusion arise in connection with intentionally acquiring a false belief.

What this shows is that although perceptual experience might have a constitutive rationality condition of some kind, it is not one of accuracy. What the rationality conditions of perceptual experience might actually be is not my chief concern. But my impression is that the best possible Neo-Kantian theory would say that it is constitutively rational to experience objects as having those features whose grounding properties are in some sense already present in our perceptual experience. This theory would accommodate cases like that of the avant-garde chair: it should be construed as a chair because one can see already that it is clearly a manufactured object with a stable flat surface at a level convenient for sitting, and so on. And it would rightly preclude classifying the green experiences of colour-blind people as constitutively irrational, because there are no properties in their experience of the red object that ground redness. We might say that the judgements that partly constitute perception, on this theory, have a constitutive aim of 'correspondence to *appearances*': they must correspond to how things appear, prior to their being made. As I say, I am not committed to this theory of perception: my claim is just that something like this is as far as the Neo-Kantian theory of perception will be able plausibly to go.

The problem for the perceptual theory of emotion is that emotions *do*, by contrast, have a constitutive aim that at least includes accuracy. Suppose I am introduced to Green, who seems immensely admirable: he seems heroic, courageous and humble. But I know by reliable testimony that Green is in fact not admirable at all: I have been told by a reliable source that Green is actually a monster, although I know nothing about what his monstrousness consists in. Clearly it would not be rational to

---

[37] Bennett Helm has raised a related point in *Emotional Reason*, pp. 41-46.

[38] If disjunctivism is true, then the claims made in this paragraph would have to be reworked in terms of 'apparent perceptual experiences', but the argument would not be substantively affected. For disjunctivism, see classically Paul Snowdon, 'Perception, Vision and Causation', *Proceedings of the Aristotelian Society* 81 (1980-1981): 175-192 and Michael Martin, 'The Transparency of Experience', *Mind and Language* 17 (2002): 376–425.

admire Green here, however admirable Green may seem. If the constitutive aim of emotion were solely one of correspondence to appearances, this would not be true. Suppose again that I pity Brown, on account of his seeming to be in wretched circumstances through no fault of his own. But I know on testimonial grounds that Brown's circumstances are not in fact wretched at all, and that their appearance as such must be some kind of illusion. Again, it is clear that pity here would be constitutively irrational, and that this is inconsistent with emotion's having a constitutive aim of correspondence to appearances.

The Neo-Robertsian might reply that I have missed an important feature of their view. As noted above, the Neo-Robertsian view is only a perception theory in the looser sense that it takes emotion to be analogous to perception. But it accepts there are some important differences. In particular, it accepts that we have emotions about predictions, memories, and facts. Suppose I am anxious about the outcome of an upcoming election. The idea is not that literally I sense-perceive the outcome of the election as possibly disastrous, which would be absurd. When it is claimed that I 'see' the outcome as potentially disastrous, only an analogy to visual sight is in fact employed: what is meant is that I consciously reflect on how a collection of facts about the election 'fit together' in a way that is analogous to seeing a visual *Gestalt*. For this reason, the response goes, the Neo-Robertsian has no problem responding to my objection. The testimony I have received about Green's monstrousness and Brown's self-sufficiency is *itself* part of how Green and Brown seem prior to my construing them one way or the other. And so they make it rational to construe Green as monstrous and Brown as self-sufficient; and they make it irrational to admire Green and to pity Brown.

It is obvious where this leads. To preclude any cases of inaccurate emotions failing to count as constitutively irrational, the Neo-Robertsian will need to say that everything I know about the object is part of how it pretheoretically appears to me, regardless of whether it is known by occurrent sense experience, by testimony, or by any other means. And so the constitutive aim of correspondence to appearances will entail a constitutive aim of accuracy. To put this slightly differently, Neo-Robertsians face a dilemma. They can say that there are sometimes emotions that are irrational even though one knows that they are inaccurate. If they do this, they would have avoided simply adopting an accuracy condition, but there would still be counterexamples to their view, namely emotions that are constitutively irrational because they are inaccurate, but which their view counts as constitutively rational. Or they can make their conception of correspondence to appearances so expansive that all inaccurate emotions count as constitutively irrational. But then they have adopted accuracy as a constitutive aim of emotion.

Is this a problem for the Neo-Robertsians? Can they just accept that accuracy is a constitutive aim of emotion ('Ultra-Neo-Robertsianism')? In a way, the answer is yes. If they do, they avoid the objection raised in this chapter. And what they would be left with would be a theory that was in many ways highly attractive, a point to which I shall return in Section 3.1. But it is not clear that there is anything left that makes it a *perception* theory. The standard definition of perception theories, mentioned above, is that they understand emotions as perceptual states, or as analogous to such states. Now,

introducing accuracy as a constitutive aim of emotion clearly rules out emotions' actually being perceptual states, because perceptual states can be inaccurate without being constitutively irrational.[39] But might emotions thus characterised still count as 'analogous to' perceptual states?

Before answering this question, we need to subject this phrase 'analogous to' to some scrutiny.[40] At first sight, what it would mean for emotion to be 'analogous to' perception is that emotion is like a kind of perception, except in certain specified respects, such that perception is still a useful model on which to understand it: one starts with perception, changes some things, and then arrives at emotion. The problem is that this makes the class of things that are 'analogous to perceptual states' uselessly capacious. After all, belief is plausibly 'analogous to' perception in this sense, and perception analogous to belief: they are both representational mental states, they differ from one another in certain respects, and specifying the respects in which one of them differs from the other could be a useful model for understanding it. So if we understand the 'analogous to' clause in the obvious way, it would entail that belief theories are perception theories! Indeed, virtually any plausible theory of emotion that includes this sort of 'analogous to' clause will cover virtually any other plausible theory of emotion, because virtually any mental state with which emotion can be plausibly identified will be 'analogous to' any other in this sense. This is obviously not a useful way of defining our terms.

One alternative would be to drop these 'analogous to' clauses altogether, and just call a theory an 'x theory' if it identifies emotion with mental state type x. Another option is to define 'analogous to' in some narrower way, such that 'x theories' include theories that identify emotion with some mental state that is 'very similar to' x. I think it is going to be difficult to rigorously quantify the similarity of mental states from one another, and to plausibly fix a precise quantum of similarity at which theories stop counting as x theories. I am not aware of any systematic existing attempts to do this, and I am not really sure how one would go about doing it. So I am tempted by the first option, of keeping our terms precise by just saying that perception theories identify emotion with perception states. If that is right, Ultra-Neo-Robertsianism - the view that emotions are construals that are constitutively irrational when they are inaccurate - is not a perception theory.

Suppose though that we want to keep the 'analogous to' clause, defining it instead as 'very similar to' on the assumption that we have a clear enough intuitive grasp of the similarity of mental states to one another to make this workable. What then would we make of Ultra-Neo-Robertsianism? The answer is that it remains implausible that it is a perception theory. What makes a perceptual state a perceptual state is controversial among epistemologists, and I shall make no attempt to resolve it here. But one thing that is clear is that, for all paradigm kinds of perception, illusions do not normally go away when you learn what they are, whereas one's belief in the content of that illusion does. When the stick

[39] Again, 'apparent perceptual states' for the disjunctivists.

[40] For discussion of the use of this distinction in existing literature, see Michael S. Brady, *Emotional Insight: The Epistemic Role of Emotional Experience* (Oxford: Oxford University Press, 2013), ch. 2.

enters water, I experience it as bending, despite knowing that it doesn't really - and there is nothing irrational about this. Perceptual illusions just are not the sort of thing we try to rationally persuade one another not to have: this is one of the big, clear, straightforward differences between perception and belief. And here, where a big, clear, straightforward difference has opened up, we find Ultra-Neo-Robertsian emotions on the side of belief, not of perception: like beliefs, the theory goes, emotions are constitutively irrational when they are inaccurate: they are the kind of thing we try to persuade one another not to have. So in this big, clear, straightforward respect, emotions are like the belief that the stick has not bent, and not like the perceptual experience that it has. Now, this does not mean Ultra-Neo-Robertsian emotions are beliefs, or even that they are 'very similar to' beliefs. But it should make us doubtful that they will count as 'analogous to' perceptual states on any useful definition of that phrase.

Maybe you are not convinced by this: maybe you think that emotions could have a constitutive aim of accuracy while still being intuitively 'very similar to' perception states. Because the concept in play here is, as I have said, a hazy one, I cannot rule this out. My final response, then, is that this does not really matter. It is, after all, ultimately just a semantic question. The substantive facts are these: emotions as characterised by Ultra-Neo-Robtersianism have a constitutive aim in a sense that perceptual states do not: to avoid constitutive irrationality, they must be accurate, which is not true of perceptual states on any plausible theory. So they are in an important and striking way different to perceptual states. If someone wants to stretch the idea of 'perception theories' so that it can still encompass them, they can do that. But it is unclear what they will have gained by this, instead of using a more finely grained terminology to define a finely grained mental landscape.

1.5 Perceptualist Responses

It should be clear by now why the objection to the construals theory, is in fact a general objection to perception theories. The construals theory encountered a dilemma, which is that emotions have a constitutive aim of at least accuracy, whereas perceptual states have at most a constitutive aim of 'correspondence to appearances'. Either the construals theory identified emotion with a kind of construal that has a constitutive aim of accuracy, in which case it was not plausibly a perception theory, or it identified emotion with a kind of construal that can be inaccurate without being constitutively irrational, in which case the identification with emotion fails. The general point here is that all perception theories will face the same dilemma. Learning that sticks do not really bend on contact with water does not make it constitutively irrational to see then as bent. Learning that someone one had admired is really a monster does make it constitutively irrational to continue admiring them. So they are not the same sort of thing.

Since concerns about emotional rationality began emerging in the literature, perceptualists have developed a number of responses. Roughly speaking, these responses all have the same shape: they acknowledge that there is something that makes recalcitrant emotion appear more rationally problematic

than perceptual illusions, but they deny that inaccurate emotion is actually constitutively irrational, seeking instead to identify some other way in which inaccurate emotions are rationally troublesome which is sufficient to explain appearances. Here I shall discuss the four such responses that seem to me to be most initially compelling, arguing that none is ultimately successful.[41]

1.5.1 Divisions of the Self

The earliest defence of the perception theory against something like the objection given in this chapter seems to have been given by Robert Roberts in his book *Emotions*. It should be remembered that he is of course defending the original construals theory, on which construals are not judgements of any kind, and not either of the Kantianised versions of his theory that I developed in Section 1.4. He is specifically discussing the case of people who suffer from phobias while knowing that their fear misrepresents its object, often treated as a paradigm case of recalcitrance.

> Why is the knowing phobic's state of mind irrational, but not that of the knowing subject of an optical illusion? I say it is because the knowing phobic feels torn between his judgement and his emotion in a way the knowing stick-viewer does not feel torn between his judgement and his visual experience. The latter is complacent and normal, taking the illusion in his stride; the former is in trouble and goes to a therapist. Where does this trouble come from? It comes from the fact that, unlike the illusion of the stick, the impression of the phobic object is a *concern*-based construal. The fear has a personal depth and life-disrupting motivational power that the illusion lacks. The bent stick is, at most, puzzling; the fear is personally compelling. This means that when the subject dissociates from his fear by denying its propositional content, it is like denying a part of himself, whereas denying his visual impression is not.[42]

Roberts accepts, then, that there is a kind of irrationality to recalcitrant emotion that distinguishes it from known perceptual illusion, but denies that it is constitutive irrationality of the kind that I have proposed. As I understand, this irrationality is grounded in the motivating character of emotion. A known perceptual illusion is normally motivationally inert, whereas a recalcitrant emotion is normally motivationally problematic: it normally involves a desire to do things that we do not believe we should do. In some sense it thus involves a division of the self that known perceptual illusions do not: one must struggle against the desires involved in one's recalcitrant emotions, whereas known perceptual illusions normally involve no struggle at all.

The problem with Roberts's account is that, although it might be able to explain why recalcitrant emotion is *problematic* for us, it is not able to explain why it is *irrational*. Suppose I am travelling through

---

[41] For an excellent critical discussion, see Helm, 'Emotions and Recalcitrance'. See also Benbaji, 'How is Recalcitrant Emotion Possible?'.

[42] Roberts, *Emotions in the Moral Life*, p. 92.

a desert with a limited stock of water. I feel an intense desire to drink, but manage to resist, knowing that I must conserve my supply if it is to last until I reach my destination. Here we have a mental state with 'life-disrupting motivational power'; there is an obvious sense in which I am 'torn between my judgement and my desire'. But there is no question of my thirst's being irrational: we can adduce facts against acting on my thirst, but not against actually having it. This shows that the fact that a state motivates one to do something that one ought not do is not sufficient to make that state irrational. It is therefore unclear where, on Roberts's account, the irrationality of recalcitrant emotion is supposed to arise. As Roberts acknowledges, the fact that recalcitrant emotions misrepresent their objects is insufficient, because perceptual illusions misrepresent their objects too without irrationality arising. The fact that recalcitrant emotions motivate us to perform actions that we know we should not perform is insufficient too, because the same is true of appetites like thirst. So although Roberts can explain why we might find recalcitrant emotions more troubling than known perceptual illusions, he cannot explain their irrationality.

One more radical option for Roberts here would be to repudiate his concession that recalcitrant emotions are irrational, arguing instead that this apparent irrationality is an illusion generated by the fact that they often motivate irrational actions. A version of this strategy has in fact been pursued by Sabine Döring, and I consider it below in Section 1.5.3.

1.5.2 The Plasticity of Emotion

Christine Tappolet has developed a different response to this problem.[43] Like Roberts, Tappolet agrees that there is some sense in which criticism is more appropriate in the case of a recalcitrant emotion than there is in the case of a known perceptual illusion. Like Roberts again, she seeks to explain away this feature of emotion in a way that does not make it a candidate for constitutive rationality. She also holds, crucially, that there is little we can do to control particular emotion episodes: when one is actually experiencing recalcitrant anger, there is nothing to be done about it save to wait until it passes. Tappolet thinks that this is what makes her account a perception theory rather than some kind of judgement account, a view which, Kantian perception theories aside, seems plausible.

What distinguishes emotion from other kinds of perception, on Tappolet's view, is that our 'emotion systems' are more 'plastic', by which she means that we can do more to change our dispositions to have certain emotions under certain circumstances than we can do to change our dispositions to have perceptual experiences. For example, we cannot do anything to affect our tendency to see sticks bend

---

[43] Christine Tappolet, 'Emotions, Perceptions and Emotional Illusions' in Clotilde Calabi (ed.), *Perceptual Illusions: Philosophical and Psychological Essays* (Basingstoke: Palgrave Macmillan, 2012). Cf. her *Emotions, Values and Agency*, pp. 31-38. For critical discussion, see Sabine A. Döring, 'Why Recalcitrant Emotions are not Irrational' in Sabine Roeser and Cain Todd (eds), *Emotion and Value* (Oxford: Oxford University Press, 2014), esp. pp. 126-130.

when they enter water, but we can typically do a lot to manage a tendency to excessive anger, or a tendency to be too easily hurt. This means that it makes sense to tell people to change their emotional dispositions in a way that it does not make sense to tell them to change their perceptual ones, in virtue of something like the principle that ought implies can. And it is this that justifies our practice of criticising emotions that misrepresent their objects. We are right to criticise emotions that misrepresent their object, she thinks, but such criticism aims at persuading us to rephase our *dispositions* to have inaccurate emotions, not at highlighting the constitutive irrationality of emotion episodes. The reasons in question are practical, not theoretical: they are reasons to do something to ourselves so that our emotions are more accurate, rather than evidence in virtue of which the emotion is irrational, regardless of what we ought to do about it.

One difficulty with this view is that it is not clear that our emotional dispositions always are more plastic than our perceptual ones. Although we cannot learn not to have the classic optical illusions, people can and do sharpen many perceptual faculties, from learning to see snakes disguised against tree bark, to learning to distinguish different grape varietals, to learning to tell what instrument is playing or what key a piece of music is in. Some blind people can cultivate their faculty of echolocation so well that they can use it to ride bicycles and play basketball, thereby revealing an astonishing degree of plasticity in a sense modality that most people are not even aware of. There are sometimes good practical reasons to make these perceptual faculties more accurate. But even when they are inaccurate, it does not seem that they are irrational in the way that recalcitrant emotions are. I might have good reason to learn to tell a Sangiovese from a Corvina, but if I just have a coarse-grained experience of 'red wine', there is plausibly nothing constitutively irrational about that.[44]

A more fundamental difficulty with Tappolet's response should be clear based on the discussion in foregoing sections. This is that practical reasons simply do not exhaust the reasons we have for emotions. A clear illustration of this is given by cases in which practical reasons pull apart from constitutive reasons. If the whimsical billionaire promises someone who knowingly suffers from a phobia of dogs that he will eradicate a widespread tropical disease in return for the phobia-victim's not doing anything to eliminate their phobia, then plausibly the phobia-victim ought not to do anything to eliminate their phobic disposition. But there remains an obvious sense in which the phobia-victim's emotion remains completely unjustified.

Perhaps Tappolet might say that what is actually going on here is just that the phobia victim has conflicting practical reasons: there are still some reasons that weigh in favour of eliminating the phobia, but they are outweighed by stronger reasons for keeping it. I do not find this very plausible: it seems to me that there is a sense in which the phobic fear is irrational, not just that there are some reasons against having it that are outweighed by reasons in favour. But the example can in any case be modified to preclude this kind of response. Suppose someone who knows they are terminally ill has a

[44] Cf. Helm, 'Emotions and Recalcitrance', pp. 421-422.

knowing phobia of dogs. They have no time to habituate different emotional dispositions, and they know it. So these emotional dispositions are completely non-plastic: so far as Tappolet is concerned, there is thus nothing to distinguish the rationality of their emotion from that of the perceptual illusion. But of course this is the wrong result. There is still an obvious sense in which the dying phobia-victim's fear of harmless toothless Fido is irrational, a sense which they will of course acknowledge.[45] Emotional dispositions often are plastic, and we surely do typically have practical reason to habituate dispositions to have accurate emotions. But this does not exhaust the ways in which emotion is a candidate for reasonableness.

1.5.3 Motivating Practical Irrationality

A related defence of the perceptual theory has been developed by Sabine Döring.[46] To some extent, Döring's defence of the perceptual theory seems initially to rest on a distinction between 'appearances of truth' and 'mere appearances'. Appearances of truth are generated by systems that are generally reliable at providing information about the environment and whose guidance we tend to follow while 'in default mode', that is, until we have some special reason to suppose they are unreliable. Even when we do have such reason, an 'epistemic inclination' remains to believe their content, which is what results in conflict. Mere appearances, by contrast, are not generally reliable, and are not believed even in default mode.

Döring concedes, however, that this is not enough to explain the normatively problematic character of recalcitrant emotions, because paradigm perceptual states seem to count as appearances of truth: they are generally reliable, and we normally believe their content unless we have some special reason not to. If, as Döring accepts, it is true that there is something more problematic about recalcitrant emotions than there is about perceptual illusions, they must have some further character that explains this. Döring's solution is that emotions, unlike paradigmatic perceptual states, motivate us to do things that we know we should not. So far, her strategy is similar to that pursued by Roberts and discussed in Section 1.5.1. But in contrast to Roberts, Döring accepts that this does not actually make emotions irrational, claiming that this is a sort of illusion. She puts it as follows:

> I think that recalcitrant emotions are intuitively sensed as irrational because, in addition to cognitive conflict, they typically involve practical conflict. Typically, we are poised to act out of our emotions—'act' being here understood in a broad sense so as to include irrational and even so-called 'arational' expressive actions. This brings the avid hiker or the oversensitive person

[45] The latter counterexample is inspired by that given by Roger Crisp in his 'Compassion and Beyond', *Ethical Theory and Moral Practice* 11:3 (2008): 233-246 at 237-238, although the target of Crisp's objection is quite different.

[46] Döring, 'Why Recalcitrant Emotions are not Irrational'. Cf. her 'What's Wrong with Recalcitrant Emotions? From Irrationality to Challenge to Agential Identity', *Dialectica* 69:3 (2015): 381-402.

> into trouble. These persons feel torn between judgement and emotional experience and should perhaps see a therapist, because their recalcitrant emotions interfere with the reasoned pursuit of their goals by moving them so as to act against those goals. But this means that, strictly speaking, the action is irrational—not the emotion itself.[47]

What we should make of this should be clear from what has been said previously. It is very plausible that emotions are more motivating than perceptual states are, and that this adds to the conflict that subjects experience when suffering from recalcitrant emotions. But it is not plausible that this exhausts the normative difference between recalcitrant emotions and perceptual illusions. To see why, we can just return to a variant on our standard example. Suppose the whimsical billionaire asks me not only to feel angry at the teacup, but also to do whatever that anger motivates me to do. I then undergo hypnosis, succeed in becoming angry at the teacup, and smash it, thereby prompting the billionaire to honour his promise and save thousands of lives. Supposing that radical evidentialism is false, my actions are rational here at every stage: I am right to undergo hypnosis, and right to smash the teacup. But something irrational obviously remains, an irrationality that distinguishes anger at teacups from seeing sticks bending when they enter water. Döring's response cannot explain this.

My example is an improbable thought experiment, but it illustrates a more general point. Perhaps it is not so uncommon that irrational emotions motivate us to perform actions which we have good reason to perform, and which indeed we might not have performed without such irrational emotions. Surely irrational confidence sometimes spurs people to make just the right audacious claim in an interview; surely an irrational lack of fear sometimes permits people to undertake risks on behalf of others that would otherwise be very difficult. These cases are no doubt a small minority: on the whole, irrational action accompanies irrational emotion. But this is clearly not always so, and hence it cannot be that the irrationality of emotion is wholly constituted by motivating effects.

1.5.4 Pointless Focussing of Attention

A fourth attempt to accommodate the irrationality of recalcitrant emotions has been developed by Michael Brady. Brady has presented his account in two slightly different versions, which it is useful to distinguish. According to the first, emotion is constituted by some kind of evaluative perception together with the focussing of our attention on trying to determine whether the evaluative perception is accurate.[48] Recalcitrant emotions obtain when we believe the evaluative perception to be inaccurate, but the attentional focus persists. Brady's key claim is that although evaluative perceptions are not rationally assessable, the focussing of attention is. The reason for this is that

---

[47] Döring, 'Why Recalcitrant Emotions are not Irrational', p. 135.

[48] Michael S. Brady, 'Recalcitrant Emotions and Visual Illusions', *American Philosophical Quarterly* 44:3 (2007): 273-284. The brief discussion in his *Emotional Insight*, pp. 161-163 is on the same lines. Brady's account shares much with that in de Sousa, *The Rationality of Emotion*.

> the persistence of attention in recalcitrant emotion is, by the subject's own lights, a waste of his attentional resources. If the point of attentional capture and consumption is to facilitate an enhanced perception of emotional stimuli, and in doing so to enable the subject to determine for himself whether or not something has the significance that he perceives it as having, then the persistence of attention will be pointless in a situation where the subject has already determined that it lacks that significance.[49]

The idea is that this provides an explanation of the contrast between the irrationality of recalcitrant emotions and the lack of rational assessability of perceptual illusions:

> While visual illusions are regarded by the subject of such illusions as mistaken or inaccurate, they do not involve the pointless mobilization of attention, and hence do not count as irrational in this way.[50]

It is plausible that emotion involves focussing one's attention on whether the object really matters in the way that the emotion represents it as mattering. But does that really explain emotions' rational assessability? One problem with this has been raised by Bennett Helm.[51] Helm notes that, whereas recalcitrant emotions are always in some way irrational, wasting one's attentional resources is not: I may watch a rerun of a banal television programme precisely because I have excess attentional resources and prefer wasting them to being bored. This indicates that there is a kind of irrationality involved in recalcitrant emotion that is of a different kind to that often involved in wasting attentional resources.

There is also a more fundamental problem here. Brady speaks of the 'pointless mobilisation of attentional resources', and of course it is in some sense correct that attending to whether a construal is true when one knows it to be false is pointless. But we need to scrutinise what exactly the 'pointlessness' of a mental state could be constituted by. On one interpretation, a mental state is pointless because it has no good effects; on another interpretation, it is pointless because it does not perform its evolutionary function.[52] It is plausible that the attentional focus involved in recalcitrant emotion is normally pointless in both of these ways. But neither of these is sufficient to explain the irrationality of recalcitrant emotions, because both of them are also true of mental states that are obviously not rationally assessable, like perceptual illusions or indeed some non-intentional bodily states. In the case of having no good effects, this is straightforward: perceptual illusions normally do no good. The matter of evolutionary functions is slightly more complex. Perceptual illusions are of course instantiations of systems that have evolutionary functions: perceptual illusions involve perceptual systems that have the evolutionary function of informing us about our environment misfire under idiosyncratic circumstances. But the same is true of focussing attention: our capacity to focus attention has an evolutionary function, and it

---

[49] Brady, 'Recalcitrant Emotions and Visual Illusions', p. 281.

[50] Ibid., p. 281.

[51] Helm, 'Emotions and Recalcitrance', p. 423.

[52] Cf. Brady, 'Recalcitrant Emotions and Visual Illusions', p. 278.

is only in idiosyncratic circumstances like those of recalcitrant emotions that it fails to fulfil it. Since Brady's account is supposed to provide some way in which attentional focusses can be *more* irrational than perceptual illusions, this means that neither failure to do good nor failure to carry out an evolutionary function can be sufficient.

A seemingly plausible alternative interpretation is that this pointlessness is a matter of having no practical reasons. On this interpretation, it is pointless to focus on whether a construal is accurate when one knows it to be inaccurate in the sense that one has no practical reason to do it. The problem with this is that the attentional focus in question is normally involuntary. It makes sense to speak of having practical reason to focus or not to focus on a question whose answer one knows when the focussing in question is voluntary, which of course it sometimes is. But the focus we are talking about here is a necessary part of a mental state that Brady agrees to be normally involuntary.[53] Given the principle of 'ought' implies 'can', practical reasons cannot get a grip here. In the absence of a more plausible interpretation of 'pointless', I think we should therefore conclude that Brady has not succeeded in explaining the irrationality of recalcitrant emotion: involuntarily focussing on whether a construal is accurate when one knows it to be inaccurate is no doubt regrettable, and if one were able to control it one should try to do so, but that does not explain why it is itself irrational.

In a later paper, Brady offers a slightly different account of emotion, and of its potential irrationality.[54] The basic structure of the account is similar: emotion is constituted by evaluative perceptions (which he characterises this time as something like Robertsian construals) and by attentional focus on whether those perceptions are accurate. This time, however, he adds the claim that this focus makes the subject more 'inclined' to 'assent' to the construal in question. He stresses that 'both the evaluative construal and the inclination to accept the construal are necessary constituents of emotions, rather than being related effects or typical accompaniments of emotions'.[55] It is this fact of being inclined to accept the construal that makes emotions irrational:

> It seems to me that the irrationality of recalcitrant emotions can be explained on two dimensions: (a) being primed to act and believe has significant practical and cognitive costs, which render recalcitrant emotions problematic; and (b) being primed to assent to a construal in light of a conflicting evaluative belief violates a substantive epistemic norm.[56]

---

[53] e.g. at 'Recalcitrant Emotions and Visual Illusions', p. 278.

[54] Michael S. Brady, 'The Irrationality of Recalcitrant Emotions', *Philosophical Studies* 145:3 (2009): 413-430. In this paper Brady characterises his theory as 'neojudgmentalist', but defines 'neojudgmentalism' to include all theories on which emotion involves evaluatively representing the object: he explicitly notes that this covers perceptualist theories like Roberts's (ibid., p. 415), and presumably his own earlier account. I have therefore not thought it necessary to torture the exposition by placing discussion of Brady's second account in a different chapter.

[55] Ibid., p. 421.

[56] Ibid., p. 427.

Brady's gloss of being 'inclined to' or 'primed to' (he uses them synonymously) believe a certain construal is as follows:

> [the subject experiences] some kind of *pressure* to accept the relevant appearance; she *leans towards* accepting the evaluative thought, is *tempted* by this way of seeing the evaluative situation, is *moved* to endorse her evaluative construal.[57]

The question, then, is whether recalcitrant emotions' constitutively inclining one to do or think things in this sense can explain their irrationality. And the answer is, I think, no. Recalcitrant emotions do of course often have practical and cognitive costs (Brady's 'dimension (a)') but, as we saw in the preceding section, it is possible for them to offer far greater benefits: when the billionaire offers a fortune in return for anger at a teacup, the practical benefits are enormous, but the emotion is, in an important sense, just as irrational. As for being inclined to believe something ('dimension (b)'), it is not clear that that involves any irrationality at all, provided one does not give in to the inclination.[58] Compare: I am in the greengrocers and I see a punnet of delicious-looking strawberries. I have a strong inclination to eat them, but I do not, because they are not mine. I have not, surely, violated any practical norm: I have merely experienced a desire to do something that would have involved violating a practical norm, had I done it. Why should the parallel situation involving an epistemic norm be any different? Someone who is inclined to believe that flying is dangerous, but who resists that temptation and continues to believe according to the evidence, has surely set a fine example of *respecting* epistemic norms, not of violating them. 'Laws and principles are not for the times when there is no temptation: they are for such moments as this, when body and soul rise in mutiny against their rigour; stringent are they; inviolate they shall be.'[59]

1.6 Concluding Remarks

A clear pattern has now emerged with the perceptualist responses to the rationality problem. In each of them, the perceptualists deny that the misrepresentation involves violating any epistemic norm in itself, while nominating some other way in which emotions putatively involve irrationality in order to accommodate our intuition that something rationally problematic is involved. In each case, we have seen that the accommodation fails: something irrational remains about the misrepresentation itself. I think there is no way of explaining this without acknowledging that emotions have a constitutive aim, and that at least part of that aim is accuracy, such that it is constitutively irrational to have an emotion if we are aware that it is inaccurate. Even as subjects of emotion we are, in this sense, in pursuit of reason.

---

[57] Brady, 'The Irrationality of Recalcitrant Emotions', p. 421.

[58] Cf. Helm, 'Emotions and Recalcitrance', pp. 423-424.

[59] Charlotte Brontë, *Jane Eyre*, ch. 27.

I have been concerned in this chapter with perception theories. But it will readily be seen that the same argument can be run against any other theory that does not characterise emotion as a candidate for constitutive rationality. It would be tedious to trawl through the literature for every account that fits this description, but it will nevertheless be worth illustrating it. One example is the account of Peter Goldie, according to which emotions consist in non-bodily intentional feelings.[60] Goldie says that an emotion 'can be irrational if it is based on irrationally formed beliefs'.[61] But as we have seen, it is not enough to outsource the irrationality of emotions to the beliefs that putatively cause them: recalcitrant emotions are still irrational, even though there are no irrational beliefs around to take the blame. Another, more complex, example is Patricia Greenspan's account, according to which emotions are feelings of comfort or discomfort whose object is a 'thought' about an object.[62] 'Thoughts' consist simply in representational content, which the subject may or may not assent to. Recalcitrant emotions, Greenspan thinks, obtain where the subject does *not* assent to the content of the thought, but the feeling of comfort or discomfort occurs anyway. The problem is that this leaves nowhere for the relevant kind of irrationality to get in. My phobic fear of Fido, on this account, comprises a representation of Fido as dangerous, a belief that he is not dangerous, and a feeling of discomfort about the dangerous representation. But although these states might be uncomfortable, none is actually supposed to be irrational, so Greenspan ends up with the same problem as the perception theorists: recalcitrant emotion may typically be unpleasant and worth avoiding, but there is nothing that is inherently irrational about it.[63] And so the problem we have observed in this chapter is quite a general one for the many theories that seek to preserve the representational character of emotion while denying that emotions are candidates for constitutive rationality, whether or not they characterise themselves as perception theories.[64] The most obvious exception to this among canonical theories are those that characterise

---

[60] Goldie, *The Emotions*, esp. chs. 2 and 3. Goldie's account is sometimes characterised as a perception theory, but this is not the way he normally describes it, and there seems no obvious reason to do so.

[61] Ibid., p. 48. Cf. ibid., pp. 23-24.

[62] Greenspan, *Emotions and Reasons*, esp. chs. 1 and 2. Like Goldie's, Greenspan's account is sometimes categorised as a perception theory, but without obvious reason.

[63] Cf. Helm, *Emotional Reason,* pp. 41-46.

[64] Further examples are the affordance and enactivist theories that writers in the Neo-Jamesian tradition have recently moved towards. See e.g. Paul Griffiths and Andrea Scarantino, 'Emotions in the Wild: The Situated Perspective on Emotion' in Philip Robbins and Murat Aydede (eds.), *The Cambridge Handbook of Situated Cognition* (Cambridge: Cambridge University Press, 2009) and Shargel and Prinz, 'An Enactivist Theory of Emotional Content'. I think the same is ultimately true of the positive account that Bennett Helm develops in *Emotional Reason*. But it will be easier to understand how we should categorise Helm's theory after we have looked at belief theories, so I postpone discussion of it to Section 3.6.

emotion in terms of belief, the paradigmatic mental state with a constitutive aim. We therefore turn to consider this possibility in the next chapter.

A brief word of qualification is merited here concerning emotions towards fictional objects, such as pity for characters in novels or fear of monsters in films.[65] There is controversy regarding whether these mental states are strictly emotions, and whether, if so, they are reasonable: some philosophers argue that they are not in fact emotions, but 'quasi-emotions' or 'imaginatively entertained emotions';[66] others argue that although they are emotions, they are pervasively constitutively irrational.[67] Some philosophers do however argue that these are genuine emotions, and that they are not unreasonable: although the distinction between constitutive and extrinsic rationality is not always drawn in the literature, it is clear from context that in least in some cases it is constitutive rationality that is in question.[68] If this last view is correct, then there will after all be some cases in which emotions have a constitutive aim of something like correspondence to appearances rather than accuracy or truth.

Although it may initially appear that this would be a troublesome counterexample to the view defended in this chapter, this is in fact far from clear. The reason for this is that these emotions, if emotions they be, are clearly emotions of a special kind, with exceptional motivational and hedonic features as well as an exceptional constitutive aim.[69] For example, the grief that people apparently feel for fictional characters is frequently pleasurable, and it almost never makes them feel as though they no longer recognise themselves. But this does not lead us to think that the grief one feels for one's dead friend is also pleasurable, or that it will not lead to a sense of existential crisis: it leads us simply to think that grief might come in two varieties. In like fashion, the fact, if fact it be, that it can be constitutively rational to feel a form of grief for people who are not real and with whom one has no prior relationship does not mean that, when one learns that one's friend is not dead after all, one's grief for him does not become irrational: it would mean only that grief is in fact made up of two related mental states, different in their aims as they are different in other respects. The true theory of emotion may have to

---

[65] Cf. Roberts, *Emotions*, pp. 94-97.

[66] See e.g. Kendall L. Walton, 'Fearing Fictions', *The Journal of Philosophy* 75 (1978): 5-27 and Roger Scruton, 'Fantasy, Imagination and the Screen', *Grazer Philosophische Studien* 19:1 (1983): 35-46.

[S] See originally Colin Radford, 'How Can We Be Moved by the Fate of Anna Karenina?', *Proceedings of the Aristotelian Society*, supplementary vol. 49 (1975): 67-80.

[68] See e.g. Peter Lamarque, 'How Can We Fear and Pity Fictions?', *British Journal of Aesthetics* 21 (2000): 291-304 and Derek Matravers, 'The Challenge of Irrationalism, and How Not to Meet It' in Matthew Kieran (ed.), *Contemporary Debates in Aesthetics and the Philosophy of Art* (Oxford: Blackwell, 2006). Matravers explicitly supports the claim that such emotions are not only rational but constitutively rational, although he does not use this terminology.

[69] For a useful discussion of this point, see Tamas Szabó Gendler and Karson Kovakovich, 'Genuine Rational Fictional Emotions' in Matthew Kieran (ed.), *Contemporary Debates in Aesthetics and the Philosophy of Art* (Oxford: Blackwell, 2006)

accommodate emotions towards fictional characters. But their existence, and their arguably anomalous constitutive aims, would not defuse the counterexample to perception theories given in this chapter.

Chapter Two

Emotion and Belief

We have established that there is a central component of emotion that has a constitutive aim of accuracy. It followed that emotion cannot be completely constituted by perceptual states. It might seem that this leads us in a clear direction. The most obvious mental state with a constitutive aim that is or includes accuracy is belief, and theories according to which emotion is constituted partly or wholly by belief were the old orthodoxy in philosophy of emotion. So surely, we might suppose, the argument of the last chapter should lead us back to the belief theory. It is to this possibility that we now turn.

A little context about the literature might be helpful here. I have said that the old orthodoxy was the belief theory, where a belief theory is one on which emotions are constituted partly or wholly by belief. And so it was, as we shall see in a moment. But it did not usually go by this name. The most popular term at the time was 'cognitivism', an imprecise word that seems roughly to mean 'theories on which emotion is constituted at least partly by belief or something rather like belief'. The other term that one frequently encounters in the belief theory literature is 'judgement'. 'Judgement' has been used in shifting ways in the literature on emotion. As we saw in the last chapter, Brady uses it very broadly, such that it encompasses perception theories as well as belief theories.[70] On the other hand, Roberts, who rejects what he calls 'judgement theories', uses 'judgement' to denote a very specific mental state, listing six characteristics and clearly distinguishing them from both beliefs and perceptual states.[71] Martha Nussbaum uses the term 'judgement' to cover beliefs, as well as certain 'seeing as' states in infants and animals that she thinks might or might not be correctly counted as beliefs.[72] Robert Solomon characterises judgements as 'mental acts' of assenting to propositions, distinguishing them explicitly from beliefs; in later work, however, he speaks of emotions as 'judgements of the body', suggesting that his concept of 'judgement' may have grown to extend beyond the mental.[73] Other conflicting definitions, implicit or otherwise, could be adduced.[74] The category of 'judgement theories' is

---

[70] Brady, 'The Irrationality of Recalcitrant Emotions', esp. pp. 414-416.

[71] Roberts, *Emotions*, pp. 84-87.

[72] Nussbaum, *Upheavals of Thought*, esp. pp. 33-49 and 125-129.

[73] See e.g. Robert C Solomon, 'On Emotions as Judgements'; for 'judgments of the body' see his 'Emotions, Thoughts, and Feelings: Emotions as Engagements with the World' in Robert C Solomon (ed.), *Thinking About Feeling: Contemporary Philosophers on the Emotions* (Oxford: Oxford University Press, 2001).

[74] Not least of course the Neo-Kantian use we saw in the last chapter, on which judgements can constitute experiences as well as beliefs.

correspondingly unstable: what counts as a judgement theory will depend on what sense of 'judgement' is in play, and there seem to have been quite a few such senses available.

This terminological mess might lead us to suspect that the old orthodoxy was correspondingly messy in substantive terms. But in fact, on the whole, that is not true: the so-called 'cognitivists' really did share a view that distinguishes them clearly from both Classical Jamesians and perceptualists.[75] This view is what I shall call the belief theory, the view that emotions are comprised in part or whole of beliefs. Martha Nussbaum comes closest to the 'pure' belief theory, claiming that a certain kind of judgement is necessary and sufficient for emotion, and characterising judgements, as we have seen, either as simply beliefs, or as aspectual experiences that are on the conceptual borderline of our category of belief.[76] The pure belief theory was however rare: most belief theorists held that emotions are conjunctions of belief with other elements. Joel Marks proposed that emotions are conjunctions of beliefs and desires.[77] Justin Oakley characterised emotions as beliefs, desires and feelings.[78] Jerome Neu characterised them as involving 'sensation, desire, behaviour and belief'.[79] John Searle suggested that the intentionality of emotions is explicable in terms of beliefs and desires that partly comprise them, although he suspected that emotions might have some further necessary elements too.[80] Antonio Damasio, often characterised as a supporter of the bodily feelings theory, actually says that emotions involve elements both of feeling and belief.[81]

All belief theorists share the view that the element of emotion that is representational and that has a mind-to-world direction of fit is belief. The content of the belief in question is generally thought

---

[75] An exception may be Robert Solomon, who identifies emotion with judgement but does not identify judgement with belief, characterising it instead, as I have mentioned, as the mental act of assenting to a proposition. See generally his *The Passions*. The points I raise about constitutive reasons might however still apply *mutatis mutandis*, since these mental acts seem to be appropriate under the same circumstances that corresponding beliefs would be.

[76] Martha Nussbaum, *Upheavals of Thought*, esp. ch. 1. Nussbuam denies that emotion has necessary elements other than belief (or 'judgement') only with some caution (*ibid.*, pp. 56-67). Nothing of importance will turn on this. A similarly unitary account of moral emotions seems to be given in John Rawls, *A Theory of Justice* (Cambridge Mass.: Harvard University Press, 1971), pp. 479-485.

[77] Joel Marks, 'A Theory of Emotion', *Philosophical Studies* 42:1 (1982): 227-242, esp. pp. 233-5. The account that Donald Davidson develops in his *Essays on Actions and Events* (Oxford: Clarendon, 1980), esp. chs. 1-5, 14 and 15 is probably correctly categorised as a belief-desire theory, although there are complexities that I cannot enter into here.

[78] Justin Oakley, *Morality and the Emotions*, ch. 1

[79] Jerome Neu, *A Tear is an Intellectual Thing* (Oxford: Oxford University Press, 2002), pp. 11-12.

[80] John Searle, *Intentionality: An Essay in the Philosophy of Mind* (Cambridge: Cambridge University Press, 1983), pp. 29-36.

[81] Damasio, *Descartes' Error*, ch. 7, esp. pp. 143-7.

to be a proposition that the object has a certain evaluative property: gratitude is partly or wholly constituted by a belief that the object has done something good to one that one was not owed; fear by a belief that the object is dangerous to oneself or to that with which one identifies; admiration by a belief that the object has a preponderance of valuable qualities; and so on. Some belief theorists have claimed that emotions must involve some specific sort of evaluative content: a relatively extreme example of this is Nussbaum, who requires that the object must be significant for the subject's own flourishing.[82] These distinctions among belief theories are important for some purposes, but on the whole they will not be important for ours, so I set them aside and mention them only when necessary.

As I have mentioned, the belief theory is relatively uncommon today. It is held to be unable to accommodate the emotions of animals, and it is held to yield an implausible characterisation of recalcitrant emotions, implying that they involve the subject's literally believing a proposition to be true and not true. More generally, it is charged with excessive 'intellectualism'. We shall return to these doubts in the next chapter, but I ask readers to set them to one side for the time being. Chapter One has shown that perception theories get something fundamentally wrong about emotion, something that belief theories seem well-placed to get right: this is enough to warrant revisiting them. What I shall argue, however, is that belief theories do not get the constitutive aim of emotion right either. The way in which they fail to do so will help us to identify the sort of account that we need.

The structure of this chapter is as follows. In Section 2.1, I outline a basic problem for the belief theory. In Sections 2.2, I look at a way in which the belief theorist might try to accommodate this problem, arguing that it is ultimately unsuccessful. The basic claim is that although emotion does, like belief, have a constitutive aim of accuracy, constitutively rational emotion also involves grasping what grounds the evaluative property in question. Because of this, testimonial evidence is often insufficient to make emotions constitutively rational: justified emotion involves grasping something about *why* an object matters in some way, not just believing justifiedly that it does. This is untrue of beliefs, and so the identification of belief and emotion fails.

2.1 The Basic Problem

It might initially seem that the belief theory is ideally placed to accommodate the observations about constitutive reasons that we made in Chapter One. Beliefs are the paradigm case of a mental state with a constitutive aim: on any standard view, for them to fulfil this aim, they must be accurate. This means

[82] For Nussbaum's defence of this proposal, see her *Upheavals of Thought*, pp. 30-33 and 49, and her *Therapy of Desire* (Oxford: Oxford University Press, 1990), pp. 90-91. For criticism in the case of compassion, see Roger Crisp, 'Compassion and Beyond'. For Nussbaum's own apparent acceptance of at least one exception to this, namely the emotion of wonder, see *Upheavals of Thought*, pp. 53-55. This qualification would have no implications for my objection to the belief theory, and I set it to one side only for simplicity's sake.

that the belief theory could accommodate with ease every counterexample to the perception theory raised in Chapter One. If the billionaire promises to save thousands of lives if I believe that a teacup has wronged me, then the belief remains constitutively irrational, though I plausibly have excellent reason to acquire it. As we have seen, the same is true of anger at the teacup. So if anger just is the belief that the teacup has wronged me (together perhaps with some feelings, desires and the like), we have a ready explanation of why this is so, an explanation that the perceptualist proved unable to supply.

A couple of snags should be mentioned in passing. First, I argued in the preceding chapter that we probably do not have constitutive reasons for moods, or for certain quasi-emotional reflexive mental states. If this is so, then it follows that they are not beliefs. If that is right, then the present argument would support at most a slightly qualified victory for the belief theory: feeling gloomy or being started would remain beyond it. But that would still be quite a victory.[83] Second, beliefs have a constitutive aim that is or entails truth, whereas it seems more natural to say that emotions have a constitutive aim of accuracy. The difference here is that propositions can be true or false, whereas anything representational can be accurate or inaccurate: thus, say, a portrait can be inaccurate but not untrue. So if belief theorists want to identify emotion with belief, they are going to have to claim that appearances are misleading, and that the content of emotion is actually propositional, such that emotions really have a constitutive rationality condition of truth too. Some belief theorists do indeed claim this, saying for instance that we avoid the language of 'falseness' here only because it is troublesomely ambiguous with 'falseness' in the sense of insincerity.[84] Although this is contentious, I am going to let it pass for the sake of argument.[85]

I am interested in a problem for belief theories that arises where they might be expected to be strongest, namely their ability to accommodate the conditions under which emotion is constitutively rational. The most clearly problematic cases involve evaluative beliefs acquired through testimony.[86] Suppose I am reliably informed that Grey is a character of immense moral nobility, but I know nothing at all about the features in virtue of which this is supposed to be true. Indeed, the impression Grey makes on me is that of being rather shallow and nasty. Should I feel admiration for Grey? The answer seems to be no. We want to say, I suggest, that admiration here 'has nothing to go on', and that although one should surely *believe* Grey to have features that would make her worthy of admiration if one had

---

[83] Indeed, belief theorists have tended to classify moods as a distinct kind of objectless mental state, and to say little about startle and its ilk. See e.g. Nussbaum, *Upheavals of Thought*, pp. 132-133.

[84] e.g. Nussbaum, *Upheavals*, pp. 46-47.

[85] For discussion of these distinctions generally, see Tim Crane, 'Is Perception a Propositional Attitude?', *The Philosophical Quarterly* 59:236 (2009): 452-469.

[86] So far as I know, this objection has not been raised before. It resonates in a vague way with the argument in Mark Johnston, 'The Authority of Affect', *Philosophy and Phenomenological Research* 63 (2001): 181–214, but my main inspiration is the aesthetics literature on the Acquaintance Principle (see below).

more information about them, admiration itself needs to 'emerge organically' from an appreciation of what is actually admirable about its object. And so it appears that a belief that Grey is admirable is insufficient to make admiration constitutively rational. Another example: I know by testimony that Silver's life is wretched through no fault of his own. But so far as I can see, Silver is a confident, successful and indeed slightly obnoxious character, who seems to be hugely enjoying himself. In such a case, it seems to me that although there are certain emotions that this situation would warrant—humility about one's knowledge of other's lives, diminished irritation about Silver's overbearingly confident manner, and so on—to actually pity Silver seems to require at least some grasp on what there is to pity about him. A final example: suppose I learn from an older mentor whose judgement I rightly respect that Black, whom I find gracious and charming, is in fact highly annoying. I lack, however, some element of perceptiveness and moral maturity with which I could discern this. No doubt this has implications for my emotions—I should greatly qualify my admiration for Black, for instance—but for me to feel annoyance at Black, when I have no idea of what there is to be annoyed about, seems clearly inappropriate.

What is going on here? The answer seems to be that, to be constitutively rational, these emotions must not only involve an accurate representation of some evaluative feature of their objects; they must also involve some kind of grasp on what *grounds* that evaluative feature. That claim leaves much unclear, notably what this 'grasp' involves, and what about the ground of the evaluative property we need to grasp: we shall explore these areas in Section 2.2, as well as in Chapter Three. But for the time being, the upshot of this is a problem for the belief theory, which, as standardly stated, only requires the accurate representation of an evaluative feature. Interestingly, then, the belief theory now faces a problem of a very different kind to those that are familiar in its connection: the problem is that emotion seems to have more demanding conditions for constitutive rationality than those that the belief theory implies.

In the next section, we are going to look at one defence of the belief theory in some detail. In the meantime, however, it will be worth clearing away some more minor points. First, we might have concerns about the objection's generality. I have given three counterexamples, but there are more than three emotions. What of the others? Let us return to our standard list: aesthetic experience, anger, anxiety, compassion, depression, envy, fear, gratitude, guilt, happiness, hope, jealousy, love, pride, relief, sadness, and shame. Excluding moods and reflexive emotions, it seems to me that the objection works in almost every case. It is unjustified to feel, say, pride, anger or guilt when one has no idea what one is supposed to be proud, angry or guilty about, even as it is unjustified to feel admiration or pity in like circumstances. So far as I can see, there is only one exception to this, which is fear. If one knows one is in danger, but does not know why, fear seems justified: knowledge of the danger might make one more afraid, but it is not a necessary condition for that fear's constitutive rationality.

We will return to this point in Chapter Three, but a basic response will suffice for now. It goes like this:

1. Other emotions are not the sort of thing the belief theory says they are.
2. Fear is the same sort of thing as other emotions.
3. Fear is not the sort of thing the belief theory says it is.

(1) follows from my objection, if indeed it is successful. The justification for (2) is simply that fear pretheoretically seems this way. I admit this is not a decisive justification: sometimes the way things pretheoretically seem to us turns out to be wrong. But I assume we should give pretheoretical appearances some weight. It is worth stressing that this case is quite different to that of moods and reflex emotions. There is a clear pretheoretical difference between moods and other emotions, such that it is actually a desideratum of a theory that it explain what that difference is. There is also a long tradition of distinguishing between more brute and more reflective emotions. But claiming that fear is one sort of thing and all other emotions are another is unprecedented and would seem bizarre.

For the time being, that is all I am going to say about the generality problem. Another potential difficulty is that it might seem that the present objection works only against one version of the belief theory, namely the 'pure' one, on which belief is necessary and sufficient for emotion. As we saw, however, most belief theorists held 'impure' belief theories, according to which emotion has at least one other necessary component, typically bodily feeling or desire. It might be thought that the belief theorist could explain the above counterexamples by appeal to these features. The reason that I do not feel admiration for Grey upon learning her to be admirable is that, although the belief is in place, the feelings or desires are not: they require some exposure to Grey's admirable qualities, and so testimony is insufficient for them.

How exactly we should reply to this depends on which other mental states are involved, and how those mental states are understood. Take for instance a view on which emotion is comprised of belief and bodily feelings. The problem for this sort of view would be that, as we have seen, bodily feelings are not candidates for constitutive rationality, even if they do represent core relational themes. The hybrid theory might be able to explain why emotion does not normally occur in the absence of testimony, but it will not be able to explain why, if it *did* occur, it would be constitutively irrational - why there would be, for example, something irrational about being annoyed by someone when one knows nothing at all about why that person is annoying. This is why this objection is much more problematic for the belief theory than the familiar objection that the belief theorist cannot distinguish emotions from evaluative beliefs that are not emotions.[87] The belief theorist can easily reply to the latter objection by adding some kind of bodily feeling to their account. But this will not help with the present difficulty.

Matters are a bit more fiddly when we consider desire. This is because it often seems that we can give reasons for and against desires. If I learn that my boiled egg is bad, then, plausibly, my desire to eat it becomes irrational. Now, many philosophers who work on desire think that this is just because

[87] For the familiar objection, see e.g. David Pugmire, *Rediscovering Emotion* (Edinburgh: Edinburgh University Press, 1999), pp. 31-32 and 89-90; Roberts, *Emotions*, pp. 90-1.

what we normally call my desire to eat this boiled egg is in fact a compound mental state, comprised of something like a belief that this egg is delicious and a reason-independent desire to eat delicious things.[88] We can follow other writers in calling this a distinction between 'instrumental' and 'intrinsic' desires, although this is something of a simplification.[89] If that is right, then clearly my objection is unaffected: emotion, on the belief-desire theory, is still ultimately comprised entirely of reason-dependent belief and reason-independent desire, even if there is some more superficial sense in which the 'desire' has a reason-dependent component. Schematically, this sort of belief-desire theory would look like this, with the mental states listed in the lower rows constituting those listed in the upper ones:

<table>
<tr><td colspan="3">Emotion</td></tr>
<tr><td>Reason-dependent belief</td><td colspan="2">Reason-dependent instrumental desire</td></tr>
<tr><td></td><td>Reason-dependent belief</td><td>Reason-independent intrinsic desire</td></tr>
</table>

Some philosophers, however, think that even our fundamental desires are, at least in some cases, candidates for constitutive rationality. These philosophers want to say that if one had a fundamental desire for suffering or for evil - not as a means to something else, but just as an end in itself - that would still be irrational. That may well be plausible. Most of the philosophers who have argued for this, however, have done so because they think that fundamental desires are themselves beliefs: Socrates, for example, thought that for an organism to desire *p* is for it to believe that *p* is good, while Thomas Scanlon has defended the view that for an organism to have a motivating desire to do *x* is to judge that it has reason to do *x*.[90] So theories of this kind pose no problem for my objection, because, if they are true, the belief-desire theory collapses back into the belief theory. Schematically:

[88] For two of many examples, see George F. Schueler, *Desire: Its Role in Practical Reason and the Explanation of Action* (Cambridge MA: MIT Press, 1995) and Carolyn Morillo, 'The Reward Event and Motivation', *The Journal of Philosophy*, 87 (1990): 169–86.

[89] Relevant distinctions are explored in Rae Langton, 'Objective and Unconditioned Value', *Philosophical Review*, 116:2 (2007): 157-185.

[90] See Charles H. Kahn, 'Plato's Theory of Desire', *The Review of Metaphysics* 41:1 (1987): 77-103 and T.M. Scanlon, *What We Owe to Each Other* (Cambridge MA: Harvard University Press, 1998). This simplifies both Socrates's and Scanlon's views, but they are of only illustrative importance here.

<table>
<tr><td colspan="3">Emotion</td></tr>
<tr><td>Reason-dependent belief</td><td colspan="2">Reason-dependent instrumental desire</td></tr>
<tr><td></td><td>Reason-dependent belief</td><td>Reason-dependent intrinsic desire</td></tr>
<tr><td></td><td></td><td>Reason-dependent belief</td></tr>
</table>

The belief-desire theorist has one last option here: they might say that the judgements that ground the desires we have in emotion are *not* solely beliefs, but involve grasping the properties that ground those beliefs. This claim produces a convoluted view according to which an emotion consists of a belief and a desire, while that desire is further subdivisible into beliefs about means, and fundamental non-doxastic judgements about ends. Schematically:

<table>
<tr><td colspan="3">Emotion</td></tr>
<tr><td>Reason-dependent belief</td><td colspan="2">Reason-dependent instrumental desire</td></tr>
<tr><td></td><td>Reason-dependent belief</td><td>Reason-dependent intrinsic desire</td></tr>
<tr><td></td><td></td><td>Reason-dependent non-doxastic mental state</td></tr>
</table>

Now, this move would allow the belief-desire theory to technically evade my objection, albeit in rather an inelegant fashion. But the victory would be hollow, because it would have been attained only by conceding the existence of precisely the kind of mental state I am arguing for. The key point would still have been established, which is that one of the fundamental constituents of emotion is a state that is not a belief but that is nonetheless a candidate for constitutive rationality. So the resulting theory would still be, in part, the kind of theory that I am advocating. Indeed, its remaining elements would run the risk of redundancy: if it is agreed that there is a representational, constitutively rational, but non-doxastic element of emotion, then it is not clear what work there is left for belief to do.

The same basic logic will apply to any other attempt to accommodate rationality conditions through hybridisation of the belief theory. Suppose, for instance, that one claimed that emotion includes some kind of representational 'feeling' that is not merely bodily, and which can be constitutively rational or irrational. If one's account of this feeling ultimately reduced its rationally assessable component to belief, then it would fail to accommodate the counterexamples we have considered. To accommodate those, it is necessary to introduce some kind of mental state that is a candidate for constitutive rationality

but that is not a belief. And that is to concede the key point at issue: the representational, constitutively rational component of emotion is not exhausted by belief.

In what follows, I am largely going to set aside hybrid belief theories, because including reference to them makes for tortuous exposition. But this response to them is presupposed throughout, and should be borne in mind.

2.2 A Sophisticated Belief Theory?

I now pass on to what seems to me to be the most promising way for the belief theorist to respond to my objection. The objection, to recall, runs as follows: the belief theorist says that emotion is a belief about an evaluative property of the object; this implies that testimonial evidence for such a belief should be sufficient to make an emotion constitutively rational; but, in many cases, it is not. A good response for belief theorists would be to claim that emotion is constituted not simply by a belief about the evaluative property, but by a belief about the evaluative property *and* beliefs about its grounding properties. They would then have a ready response to my counterexamples: one should not admire someone if one has evidence about why they are admirable because admiration is partly constituted by beliefs about why they are admirable; one should not pity someone without knowing anything about why they are pitiable because the same is true of pity; and so on. Another way of putting the same thought: I have said that emotion cannot be just a belief about an evaluative property, but must involve 'grasping' what grounds that property. The sophisticated belief theorist might agree, but say that this 'grasp' is itself reducible to belief, meaning that beliefs end up being emotion's sole constituents after all.

This response can accommodate the counterexamples discussed in the previous section. I believe, however, that there are related cases with which it has more difficulty. We shall start by considering a central group of emotions that we have for artworks, including many instances of admiration, distaste and wonder. I submit that to occurrently have these emotions in a constitutively rational way, one needs to be occurrently experiencing the works in question. To admire, say, the Selimiye Mosque, it is not enough to know a great deal about it. It may not even be enough to have seen and admired the Mosque itself at some point in the past. I propose that there is a crucial sense in which, to be admiring the Mosque, one must actually be experiencing it. This suggests that the 'grasp' involved in emotion is not constituted belief, because beliefs our beliefs about things outlast our experiences of them. Some other kind of grasping is going on.[91]

---

[91] Here and in what follows I am inspired broadly by the account of aesthetic understanding developed by Roger Scruton in *The Aesthetics of Music* (Oxford: Oxford University Press, 1999), esp. ch. 8 and *The Aesthetics of Architecture*, 2nd edn (Princeton: Princeton University Press, 2013 [1979]), esp. chs. 4 and 5. But Scruton does not apply his account to emotion, about which he apparently favoured a belief-desire theory. See his *Art and Imagination: A Study*

This is a complex and unfamiliar argument, and I shall now develop it slowly. I begin with three clarificatory points regarding the sense in which we sometimes need to be experiencing artworks in order to have certain emotions about them. First, I do not deny that there are emotions that one can rationally have about artworks without experiencing the artworks in question. Of course there are many. I might be impressed by the feats of engineering involved in the domes of the Selimiye Mosque, or surprised by the features that distinguish its plan from those of other Ottoman mosques: to have these emotions rationally I need only know the relevant facts. My claim is only that there is a central group of emotions that depend upon the experience itself: that is enough for the purposes of the counterexample. An example of such an emotion is the admiration I feel for the Selimiye Mosque upon entering and gazing up at the dome: this, I am claiming, is an emotion that I cannot have without occurrently experiencing the building at the same time.

Second, and crucially, 'occurrently experiencing' must be understood quite broadly: it does not imply that one is necessarily looking at the object itself. The most obvious exception involves pictures or other reproductions of the artwork: plausibly, one can rationally have the central kind of admiration for the Selimiye Mosque by looking at photographs of it. Naturally there is much that reproductions are likely to miss, and one's admiration will be correspondingly attenuated, but it is plausible that it may still exist, and exist rationally.[92] The more fiddly group of exceptions involve 'imaginatively experiencing' the object. It is plausible that if I know enough about the Selimiye Mosque, I may be able to conjure up an image of it 'before my mind's eye'.[93] It is certainly true that I have once seen the Selimiye Mosque, I may be able to conjure up such an image from memory. Such images are typically shadowy and lacking in detail: if I am asked, say, exactly what the colours of the intrados of the Mosque dome are, I may discover that I do not know, despite being able to visualise the dome in some sense. But such 'inner experiences' certainly may have enough content to justify some emotion, though normally emotion of an attenuated sort. My claim is that rational emotion about the Mosque becomes impossible once one is unable to imaginatively visualise it in this way - even if one still knows a lot of information about it.

---

in the Philosophy of Mind (London: Routledge, 1973), pp. 128-9 and his 'Emotion, Practical Knowledge and Common Culture' in Amélie Oksenberg Rorty, *Explaining Emotion* (Berkeley: University of California Press, 1980), p. 524.

[92] Someone might deny this and claim that such admiration is really just admiration for the photograph of the Mosque, not for the Mosque itself. I doubt this is correct, but it would be no problem for my argument if it were.

[93] What exactly this involves is a matter of controversy into which I do not enter here. For useful conceptual distinctions, see Leslie Stevenson, 'Twelve Conceptions of Imagination', *British Journal of Aesthetics* 43:3 (2003): 238-259. For discussion, see e.g. Gregory Currie and Ian Ravenscroft, *Recreative Minds: Imagination in Philosophy and Psychology* (Oxford: Oxford University Press, 2002) and the essays collected in Amy Kind (ed.), *The Routledge Handbook of Philosophy of Imagination* (London: Routledge, 2016).

Third, and more straightforwardly: I do not, of course, deny that we can have 'dispositional' emotions for works of art that we are not experiencing in any way. Suppose I saw the Selimiye Mosque several times many years ago and was deeply impressed by it. I remember feeling impressed, and I am rightly confident that I would be impressed again if I visited now. But I cannot visualise the Mosque: perhaps I cannot even recall basic facts about, say, the interior colour scheme. In such a case I might truly say that I admire the Selimiye Mosque, and my admiration could be constitutively rational. But the sense of 'admire' here is not that with which I am concerned in this argument.[94]

The claim, then, is that there are some occurrent emotions that we have towards artworks that are constitutively rational only if we are experiencing those artworks, in a broad sense of 'experiencing' that encompasses imaginative recollection. Expressed impressionistically: admiring beauty requires experiencing beauty, not just knowing various propositions about the beautiful object. I believe this claim is intuitively compelling. What does it show us about emotion? The sophisticated belief theorist says that admiring the Selimiye Mosque is just a matter of having various beliefs about its beauty and that beauty's ground. But if this were true, it would be unclear why occurrent experience is required for such admiration to be constitutively rational. Beliefs about what grounds the Mosque's beauty do not become irrational when one ceases to experience it: beliefs about objects endure beyond our experiences of them, which is one reason they are useful to us. And so it seems that the grasp of the ground of beauty involved in rational admiration for it cannot be constituted wholly by belief. Emotions are, so to speak, sometimes 'fused' or 'blended' with experiences in a way that belief is not.

I should stress that I am not just claiming that admiring the Selimiye Mosque requires experience of the Selimiye Mosque in the way that knowing what it is like to experience the colour red arguably involves having experienced redness.[95] Once one has experienced the colour red once, one knows for good what it is like. My claim about the Mosque is that occurrent experience of it is required for certain rational occurrent emotions about it: once the experience stops, the rational occurrent emotion stops too (with due caveats about imaginative recollection). The form of experience-dependence involved is thus notably stronger than that involved in knowledge of what it is like to experience red. To my knowledge, nobody has previously claimed that there are emotions that have this kind of experience-dependence. But it seems to me that this is nonetheless true.

One possible response for the belief theorist would be to defuse the counterexamples I have raised by invoking some version of what aestheticians call the 'Acquaintance Principle'.[96] According to

---

[94] Cf. my general remarks about 'dispositional' and 'occurrent' emotions in the introduction.

[95] Cf. Frank Jackson, 'Epiphenomenal Qualia', *The Philosophical Quarterly* 32:127 (1982): 127-136 and the ensuing literature.

[96] Versions of this can be found in e.g. Roger Scruton, *Art and Imagination*, ch. 4; Richard Wollheim, *Art and Its Objects*, 2nd edn (New York: Cambridge, 1980 [1968]); Michael Tanner, 'Ethics and Aesthetics Are - ?' in José Luis Bermúdez and Sebastian Gardner (eds), *Art and Morality* (New York: Routledge, 2003).

the Acquaintance Principle, in its classical version, aesthetic knowledge requires experience of its object: one cannot know that the Selimiye Mosque is beautiful until one has seen it, and one certainly cannot know that it is beautiful through being told so by others. The belief theorist might claim that, given the Acquaintance Principle, the divergence between justified belief and justified emotion that I claim to obtain in the case of admiration for the Selimiye Mosque does not actually obtain: I cannot have the relevant justified beliefs any more than I can have the relevant justified emotions. Indeed, they might say, the fact that testimony is inadequate for aesthetic knowledge as well as for aesthetic emotion actually offers abductive support for the belief theory.

I think that, given the truth of this version of the Acquaintance Principle, this response would probably be successful in defusing my counterexample. I am therefore committed to the claim that the Acquaintance Principle, as stated, is false. Fortunately, however, this is now widely held in the literature. The Acquaintance Principle was repeatedly criticised in the early 2000s by philosophers who argued that no clear reasons had ever been given for it, and that its apparent plausibility derives precisely from the fact that aesthetic *appreciation* presupposes acquaintance.[97] Some philosophers have since defended versions of the Acquaintance Principle, but not in its classical form. The best-known defence runs that it *is* possible to acquire aesthetic knowledge through testimony, but that we have what I am calling non-constitutive reasons against doing so in most circumstances: roughly, we ought normally to act as experts on aesthetic matters, with a responsibility to make up our minds for ourselves rather than accepting the testimony of others.[98] It will readily be seen that the truth of this revised Acquaintance Principle would be no problem for me. My claim is only that there is a divergence between the conditions under which some beliefs and emotions about artworks are *constitutively* rational: I am neutral on what non-constitutive reasons we may have for them. The other well-known modified version of the Acquaintance Principle runs that although there is no general reason why we cannot acquire aesthetic knowledge by testimony, testimony is in practice frequently unreliable, at least in judgements of contemporary art.[99] Again, this is no problem for me: the point is only that where knowledge by

[97] Malcolm Budd, 'The Acquaintance Principle', *British Journal of Aesthetics*, 43:4 (2003): 386-392, Paisley Livingstone, 'On an Apparent Truism in Aesthetics', *British Journal of Aesthetics* 43:3 (2003): 260-278. See also Brian Laetz, 'A Modest Defense of Aesthetic Testimony', *Journal of Aesthetics and Art Criticism* 66:4 (2008): 355-363.

[98] Robert Hopkins, 'How to be a Pessimist About Aesthetic Testimony', *The Journal of Philosophy* 108:3 (2011): 138-157.

[99] Aaron Meskin, 'Aesthetic Testimony: Can We Learn from Others about Beauty in Art?', *Philosophy and Phenomenological Research* 69:1 (2004): 65-91 and 'Solving the Puzzle of Aesthetic Testimony' in *Knowing Art: Essays in Aesthetics and Epistemology*, ed. Dominic McIver Lopes and Matthew Kieran (Dordrecht: Springer, 2006).

aesthetic testimony is possible, aesthetic emotion is not.[100] No mainstream contemporary version of the Acquaintance Principle, then, would pose a problem for my argument.

A second possible answer for the sophisticated belief theorist is that my objection fails to generalise. Emotions about the surface ornamentation of mosques are surely distinctive in many ways: perhaps the sophisticated belief theorist can hope to offer some special account of them and their ilk. They might say that they are actually 'aesthetic quasi-emotions' of some kind, and not really emotions strictly speaking at all. Or they might say that they are emotions, but that they are constituted differently from other emotions, and that the class of 'emotions' is actually made up of two rather different sorts of things. We should be clear that, on either version, this is far from an ideal response for the sophisticated belief theorist. Both versions of the argument involve offering a fractured account of the class of things that we take pretheoretically to be emotions. Still, as I have said elsewhere in this essay, the badness of a theory's being fractured depends very much on where the fracture lies. Excluding counterexamples from the *analysandum* on an *ad hoc* basis is obviously very implausible. But to the extent that the fracture in the theory lies somewhere where it is plausible that there could be a real fracture in the phenomenon that the theory is describing, then positing it is less implausible. And in this case, perhaps the sophisticated belief theorist could claim this is so: art is strange, aesthetic experience is strange, and so we should not be that surprised if our general theory of emotion cannot cover some of the emotions that involve them.

The problem with this answer is that this quality of being 'fused with experience' applies far beyond such rarefied cases. Suppose I meet Grey, and I find his manner amusing. It seems to me that we can run through the same moves concerning Grey's amusing manner as we did with the dome of the Selimiye Mosque. Later that evening, I will be able to remember that I was amused by Grey's manner; it might be true in the dispositional sense that I find Grey's manner amusing. But I will not actually be occurrently amused, except to the limited extent that I am able to experience Grey's manner anew through some kind of imaginative recall. The fleetingness of amusement is one reason why impressions of people have such a striking effect on us: once someone's manner is reproduced in experience, we suddenly recognise it ('Yes, he talks just like that!'), and grasp again what it was that made it important to us. Nor is this sort of counterexample restricted to amusement. Grey could just as easily have a manner that was rationally seen with solicitude, irritation, unease, frustration, distaste, pity or respect, and these emotions would be similarly difficult to characterise in terms of the belief theory.[101]

---

[100] Cf. Laetz, 'A Modest Defence of Aesthetic Testimony', pp. 358-360 for a persuasive argument that, even if Meskin is right about contemporary art, he would probably be wrong about artworks like the Selimiye Mosque.

[101] So it is not important for my argument if, as some argue, amusement is not an emotion. I think however that it has been convincingly shown that emotion is either an emotion, or is not an emotion only in virtue of features that are unimportant to the present argument. See esp. Robert C. Roberts, 'Is Amusement an Emotion?', *American Philosophical Quarterly* 25 (1988): 269-274 and Roberts, *Emotions,* pp. 300-308.

I cannot be rationally irritated by someone's manner when that person is not there, unless I can imaginatively recall it, or unless someone else can do a good impression of it. There seem to be quite a few emotions, then, that are sometimes 'fused with experience', requiring us to represent their objects in experience, memory or imagination for us to have them rationally.

Now, the sophisticated belief theorist may retreat again, and say that all these apparent emotions are either not emotions or are emotions of a different kind. But this is surely a desperate move. The class of emotions that their theory cannot cover has now spread out to encompass a huge range of everyday interpersonal emotion, far beyond a few rarefied sentiments about paintings and mosque architecture. For any theory of emotion, that is a big disadvantage, both because of the sheer fact that it is no longer even claiming to be a theory of many of our pretheoretical emotions, and because the line it has drawn down the middle of the class of pretheoretical emotions looks strikingly like an *ad hoc* attempt to exclude the numerous counterexamples. I think this gives us a strong reason to be doubtful about the success of the sophisticated belief theory.

We have covered quite a lot of ground, and it might be helpful to review the argument that has got us here. The simple belief theory told us that emotion is a belief with certain evaluative content. Such a belief could be justified by testimony, but testimony that an object has a certain evaluative feature is normally insufficient for rational emotion. Rational emotion seems to require some grasp of the grounding properties of the evaluative feature in question. We then turned to look at what this 'grasp' might be. The sophisticated belief theorist claimed that it too was constituted by belief: the sophisticated belief theory thus runs that emotion is a belief about some evaluative feature of the object together with various beliefs about what grounds that evaluative feature. A problem for this theory is that emotion often seems to be 'fused with experience' in a way that belief is not. To have many emotions about artworks rationally, we not not only need to have some beliefs about what grounds the evaluative feature in question, but we actually need to be experiencing that ground for ourselves while we are having the emotion. We looked at two ways that the sophisticated belief theorist could respond to this, by appealing to the Acquaintance Principle or by offering a separate account of emotions about artworks from emotions about other objects: neither response was successful.

2.3 Concluding Remarks

What has been given here may not be a knock-down argument against the sophisticated belief theorist: they might bite the bullet and insist that their theory is still the true account of some pretheoretical emotions, even though it cannot be the true account of all of them. But to the extent that this conclusion is implausible, the sophisticated belief theory is weakened. In the next chapter I will turn to exploring the kind of theory this points us towards, in the course of which I shall highlight some further advantages that a non-doxastic judgement theory has over the sophisticated belief theory.

Chapter Three
Towards a Theory of Emotion

We are now in a position to move towards a positive account. It will be helpful to begin with some recapitulation. In Chapter One, we saw that emotions have constitutive aims, including an aim of truth or accuracy: the fact that the billionaire's teacup could not possibly wrong me is a reason not to be angry with it in the sense that my judgement that this is the case will normally be constitutive of my not being angry. This led us to reject perception theories, which struggled to accommodate this constitutive aim. In Chapter Two we saw that constitutively rational emotion not only involves representing the object as mattering in some way, but also grasping *why* it matters in this way. At least in some cases, this grasp seems to require occurrent experience. This was a problem for belief theories, because constitutively rational belief does not require occurrent experience in the same way.

We know, then, that the true theory must accommodate the fact that emotion has a constitutive aim of accuracy, and therefore that emotion is not a perceptual state. We also know that it must accommodate the fact that some emotions have this peculiarly close tie to occurrent experience, and that emotion therefore cannot be belief. For ease of reference, I am now going to introduce a bit of terminology. In Chapter Two, I mentioned the term 'cognitivism', which I characterised as a vague term for theories on which emotion is constituted in part of whole either by belief, or by something like belief. I am now going to precisify 'cognitivism' by stipulation as any view on which emotion is constituted in part or whole by a mental state that is a candidate for constitutive rationality. Since belief is the paradigmatic such state, this seems faithful to the term's vague existing sense: all the theories that were originally called 'cognitivist' will continue to count as such on my stipulative definition. But my view can be contrasted from most of those accounts as 'non-doxastic cognitivism', that is, a view on which emotion is a candidate for constitutive rationality, but is not a belief.[102]

There are potentially many forms of non-doxastic cognitivism, and this chapter is dedicated to working out how we should situate emotion in this landscape. Most existing theories of emotion have sought to characterise emotion as an instance of one or more other mental states: emotions are perceptions of value, or they are evaluative judgements, or they are combinations of belief and desire. One of the results of the argument so far is that we might have to renounce these reductionist aspirations. The reason for this is simply that there may not be any more general sort of mental state whose constitutive aim has the same structure as the one we have found emotion to have: we might find that there are no non-doxastic cognitive states available to which emotion can be reduced. This might mean we need to accept an 'emotion theory of emotion', a possibility which I discuss in Section 3.1. Some hope of a reduction remains, however, namely to a kind of *understanding*, a characterisation

[102] I will occasionally also use 'cognition' to refer to mental states that are candidates for constitutive rationality, and 'cognitive' as the corresponding adjective.

that looks plausible on some theories of what understanding is. I discuss this in Section 3.2. In Section 3.3 I look at a question of detail, namely how much we need to grasp about what grounds the evaluative property in question for a given emotion to be constitutively rational.

Although my account differs importantly from the belief theory ('doxastic cognitivism'), it nonetheless has an obvious affinity with them: emotions as I characterise them are representations for whose accuracy we are responsible: beliefs share this feature, while perceptions either do not, or do so only in a very circumscribed way. This raises the question of how my account fares against the objections that have been levelled against doxastic cognitivist theories. In Sections 3.4 and 3.5 I revisit two of those objections, namely that cognitivist accounts are inconsistent with the existence of animal emotion, and that they cannot deal with the prevalence and nature of emotional recalcitrance. In the case of animal emotion, I argue that there is no inconsistency between its existence and the claim that emotions have a constitutive aim; in the case of recalcitrance, I argue that though it may be problematic for belief theories, these problems do not extend to my account. Section 3.6 is a sort of appendix to the discussion of the preceding three chapters, in which I discuss the account of Bennett Helm. Helm has got a great deal right about what is wrong with other theories, making it awkward to deal with him earlier, but I argue that he ultimately succumbs to a version of the very problems that he identifies.

A few general points before we begin in earnest. For the first three sections I am going to be concentrating entirely on the cognitive component of emotion. As I said in Chapter One, however, I am open to the possibility that emotion has other components, even other necessary ones. I have already mentioned one reason why we should be sympathetic to the possibility that emotion is not completely constituted by judgement in Section 1.3, namely that positing other components helps us to explain the continuities that exist between emotions and moods. I will mention another reason for this in Section 3.4, namely that it is one way (though not the only way) of explaining the continuities that exist between human and animal emotion. In Chapter Four, I shall give a direct argument that not only is this rational component not the whole of emotion, but that it is not even the only necessary component. Although in Chapter Four I will say some more about one other necessary component, neither here nor there shall I make any attempt to give a general account of the other elements that emotion may include. So what I am giving here falls short of being a complete theory of emotion: it is an account of a central and necessary feature of emotion, but not of everything that may make it up. My account is still inconsistent with most major contemporary theories of emotion, and that is ambitious enough for the time being.

Second, I am not going to make any serious attempt to characterise the content of emotion. So far, I have just spoken of 'evaluative properties of the object'. It might be asked: is just any evaluative property enough, or does emotion involve evaluations only of certain kinds? This has not been one of the traditional focusses of the emotions literature, but as in some cases we have already seen, philosophers have explored it to some extent: Nussbaum argues that emotions must involve evaluating the object in terms of its ramifications for the agent's own flourishing; Roberts suggests that it involves

construing the object in terms of the agent's 'concerns'; Helm gives an unusually complex account of the 'import' in terms of which, on his view, emotions represent the object.[103] My own sympathies lie with something like the kind of account that Roberts and Helm are outlining: I think there is a difference between judging something it to matter, and judging it to matter *to me*, and that the latter is necessary for the cognition to partly constitute an emotion. I shall occasionally help myself to this way of speaking in what follows, without making a sustained enquiry into what this distinction consists in, or whether the 'to one' clause really is necessary for emotion. But nothing in my wider account rests on this point, and readers who prefer another account of the content of emotion are free to substitute it. Settling this interesting area properly is another task that must be left to the future.

3.1 An Emotion Theory of Emotion?

Here is an important distinction for what follows: on the one hand we have 'reductionist' theories, according to which it is possible to give a complete account of emotion in terms of some other mental state or group of mental states; on the other we have 'non-reductionist' theories, on which we can say various things about the nature of emotion, but must ultimately accept that it is one of the fundamental kinds of mental state. Most existing accounts of emotion are reductionist: examples include the theories that emotions are beliefs, or bodily feelings, or non-bodily feelings, or attentional focusses, or belief-desire hybrids. There are also exceptions to this. The most numerous are the theories that introduce an 'analogous to' clause. As we have discussed, for x to be 'analogous to' y does not mean that x is y: indeed, it effectively means that x is *not* y, but that one can get a better grip on the nature of x by starting from y and noting the differences. Theories that emotion is something 'analogous to perception' are thus really, in a very real sense, not perception theories of emotion but 'emotion theories of emotion': they may describe emotion partly 'in terms of' perception, but ultimately they say that emotion is its own thing, one of the fundamental constituents of our mental life. We might think of them as non-reductionist theories in reductionist camouflage

There is something pleasing about reductionism, which is that it results in a more parsimonious account of mind. According to one extreme and much-criticised reductionist theory, all our intentional mental states can be characterised in terms of two fundamental mental states, namely belief and desire, these being distinguished from one another by their 'direction of fit'. This account has few adherents today, but whatever else may be said about it, it is surely true that it is extremely parsimonious. Nevertheless, I do not think we should be greatly worried by the prospect of reductionism's failure in itself. Some mental states are going to have to be basic, and there is no obvious reason why we should

---

[103] See esp. Helm, *Emotional Reason*, ch. 4. Some philosophers have also given accounts of the distinctive content of individual types of emotion, a fascinating area that, alas, I shall not touch on at all. See esp. Roberts's brilliant discussions in *Emotions*, part 3.

suppose emotion must not be one of them. There is nothing either in folk psychology or empirical psychology that obviously suggests that emotions must be special cases or composites of some other, more fundamental, kind of mental state.

Non-doxastic cognitivism could in principle be either reductionist or non-reductionist, depending on whether emotion is reduced to some other non-doxastic cognitive state. The difficulty for reductionism, however, is that there are not many other mental states in the non-doxastic cognitivist category, if indeed there are any at all. Belief is not only a paradigm of a mental state with rational standards: on some theories of mind, it is the only one. Tentative arguments have been advanced in favour of characterising desire and pleasure as candidates for constitutive rationality, but it is far from clear (a) whether they are successful, (b) whether, if these arguments were successful, belief or desire would not collapse into a kind of belief, and (c) whether, even if they did not, it would be plausible to identify emotion with them.[104] The reduction that *might* be plausible, I believe, is to understanding, and I discuss this possibility in the next chapter. If it is not successful, however, then we are left with non-reductionism, that is, with an emotion theory of emotion. And as I have said, I do not think that is anything to worry about.

Before we pass on to look at understanding, however, I have a promissory note to deliver upon. In Chapter One, I invented a theory that I called 'Ultra Neo-Robertsianism'. Recall that Roberts proposed emotion is a kind of 'construal', a mental state of which a very simple illustration is given by the experience we have of the duck-rabbit or the Rubin Vase. Roberts thinks that construals are not subject to rational standards; I suggested that there is at least one view of perception on which they might be, but concluded that even if that view was true, it would not be enough to accommodate the rational standards of emotion. In the course of this discussion, though, we experimented with a more radical view, according to which emotion is something like a construal, except that it has constitutive rational standards of both verisimilitude and accuracy. I called this view 'Ultra-Neo-Robertsianism'. At the end of Chapter One, I concluded that Ultra-Neo-Robertsianism is an attractive theory, and one that survived my objections to perception theories, but that it could do this only because was *not* a perception theory under any plausible interpretation of that term.

According to Ultra-Neo-Robertsianism, emotion has constitutive rational standards, but it is not a belief. Ultra-Neo-Robertsianism is thus a form of non-doxastic cognitivism about emotion. What is more, it is plausible that it entails the right constitutive standards too. Ultra-Neo-Robertsian construals must of course be accurate, because I stipulated as much. But it seems to be true more generally that when one construes an object in terms of a property, this involves a kind of non-doxastic grasping of what grounds the property in question. Seeing the avant-garde chair as a chair involves grasping various features in virtue of which it is a chair as such: I cannot see a chair as a chair without occurrently experiencing some of the features that ground its chairhood. So Ultra-Neo-Robertsian

[104] Cf. e.g. Moran, 'Frankfurt on Understanding' and Scanlon, *What We Owe to Each Other*.

construals involve representing an object as having certain evaluative property, where the representation in question is not a belief but must nonetheless be accurate in order to be constitutively rational, and where, to be constitutively rational, one must also have a certain sort of non-doxastic grasp on the grounding properties of this evaluative feature. In other words, Ultra-Neo-Robertsianism seems to be exactly the sort of account that I am proposing.

I stress again that this does not mean that I am advancing a perceptual theory. Ultra-Neo-Robertsianism is not a perceptual theory: it is in fact an 'emotion theory of emotion', characterising emotion as having features that no perceptual state could plausibly have. But it does mean that perception might be a useful analogy on which to think of emotion.

3.2 Emotion as Understanding?

As I have mentioned, there *is* a mental state to which emotions as I have characterised them might turn out to be reducible, namely understanding. Specifically, it might prove reducible to *occurrent subjective grounding understanding of evaluative properties*. Some unpacking is due. First: we can distinguish between occurrent and dispositional understanding, just as we do with belief: occurrent understanding is roughly the 'clockable' state of grasping some explanation, and dispositional understanding is the ability to act and reason in ways that presuppose a grasp of the explanation in question. As I have said, I am primarily interested in occurrent emotion, so it is occurrent understanding that is primarily relevant here. In normal English the word 'understanding' is used more often for dispositional than for occurrent understanding, so this makes my terminology a little awkward, but we shall have to live with this. Second: a distinction can be drawn between 'objective' and 'subjective' senses of 'understanding': 'objective' understanding is factive, being thus analogous to knowledge, and 'subjective' understanding is not, being thus analogous to belief. We use both of these in ordinary speech: we may describe the same condition as 'a mistaken understanding of x' or as 'not understanding x'. Since emotions can misrepresent their objects, it is understanding in the subjective sense that is the plausible candidate for identification. Third: it is generally thought that understanding something involves grasping an explanation of it, but these explanations can come in two kinds: causal and grounding. The relevant kind here is grounding understanding. In admiring the Selimiye Mosque, I grasp what grounds its beauty, but I need not have any grasp on what it was about the personality of Mimar Sinan or the structure of sixteenth-century Turkish architectural patronage that caused this beauty to exist. Finally, the object must be an evaluative property. As I have said, 'evaluative property' is something of a placeholder, and it is plausible that in fact the object must be a certain sort of evaluative property. But 'evaluative property' will do for now.

Whether it is plausible to identify emotion with this kind of understanding depends, however, on how such understanding is characterised. Theories of understanding are distinguishable into those

that reduce it to belief, and those that do not.[105] On a 'reductionist' theory, the kind of understanding that interests us a network of occurrent beliefs about what grounds evaluative properties.[106] That account ought to sound familiar, for it says that understanding is just what emotion was claimed to be under the sophisticated belief theory that we discussed in Section 2.2. So if both the sophisticated belief theory of emotion and the reductionist theory of understanding are true, then emotion is indeed a kind of understanding, though not in the way I am proposing here. In Section 2.2, however, I argued that the sophisticated belief theory is false, and that emotion is irreducible to belief; I will give further reasons to think this in the later sections of this chapter. This means that if the reduction of understanding to belief is right, then emotion and understanding are not the same thing.[107]

The alternative to doxastic reductionism about understanding is of course 'non-reductionism', the view that understanding is not reducible to belief.[108] Non-reductionists do not agree on the nature of understanding, and some non-reductionist theories are probably inconsistent with an identification of emotion and understanding. Investigating exactly how different non-reductionist theories of understanding line up with my account of emotion is beyond the scope of the present enquiry. But it will be clear that there is a strong *prima facie* case for identifying the non-reductionist's understanding with my emotion. Understanding involves representing evaluative properties and grasping what grounds them. It is a candidate for constitutive rationality, and its conditions of rationality seem to be the same as those of emotion. Furthermore, significantly, it comes in degrees: one can understand how

---

[105] In the literature it is more normal to speak of reducing objective understanding to knowledge than subjective understanding to belief, but since the latter is more relevant in connection with emotion, I shall normally refer to it instead.

[106] See esp. Peter Grimm, 'Is Understanding a Species of Knowledge?' *British Journal for the Philosophy of Science* 57:3 (2006): 515-535; Peter Lipton, 'Understanding Without Explanation' in Henk de Regt, Sabina Leonelli and Kai Eigner (eds), *Scientific Understanding: Philosophical Perspectives* (Pittsburgh: University of Pittsburgh Press, 2009); Paulina Sliwa, 'Understanding and Knowing', *Proceedings of the Aristotelian Society* 65:1 (2015): 57-74 and 'Moral Understanding as Knowing Right from Wrong', *Ethics* 127:3 (2017): 521-552; Amber Riaz, 'Moral Understanding and Knowledge', *Philosophical Studies* 172 (2015): 113-128.

[107] Of course, *if* the sophisticated belief theory of emotion were true, *and* reductionism about understanding is true, then identifying emotion and understanding might be plausible after all.

[108] Linda Zagzebski, 'Recovering Understanding,' in Matthias Steup (ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue* (New York: Oxford University Press, 2001); Jonathan Kvanvig, *The Value of Knowledge and the Pursuit of Understanding* (Cambridge: Cambridge University Press, 2003); Catherine Elgin, 'True Enough', *Philosophical Issues* 14 (2004): 113-131; Duncan Prichard, 'Knowing the Answer, Understanding and Epistemic Value', *Grazer Philosophische Studien* 77:1 (2008): 325-339; Michael Strevens, 'No Understanding Without Explanation', *Studies in History and Philosophy of Science: Part A*, 44:3 (2013): 510–15; Alison Hills, 'Moral Testimony and Moral Epistemology', *Ethics* 120:1 (2009): 94-127 and 'Understanding Why', *Noûs* 50:4 (2016): 661-688.

something matters more or less 'deeply', depending on how much about the ground of the evaluative feature in question has been grasped. As we shall see in Section 3.3, this is true of emotion too.

It is beyond the scope of my project to settle this debate about understanding, so I remain agnostic about whether an understanding theory of emotion could be the right one. It is worth noting in passing, though, that the truth of my view would offer some support to the cause of the non-reductionists. This is because it would offer them a companion-in-guilt in positing the existence of a mental state that has standards of constitutive rationality but that is not belief. If I am right, then such mental states exist: positing that understanding is constituted by one of these is not less parsimonious than positing that it is constituted by belief. This by no means entails that non-reductionism about understanding is true, but it does remove one reason we might otherwise think it false.

3.3 Emotional Complexity

I have said that emotions involve the subject's grasping what grounds the object's evaluative property. An obvious question is: how much do they have to grasp? Take, say, the state of being appalled by Stalin's Purges. Does one just need to know that millions of people were murdered and that millions more were imprisoned? Or does one need to know everything there is to know about the Stalinist Purges in virtue of which they are appalling? I believe that we should resist the temptation to give a unitary answer to this question. What I wish to say is this: emotions come in degrees of what I shall call 'complexity', where a more complex emotion involves a greater grasp on what grounds the evaluation in question. This is at least analogous to, and might be identical with, the feature of understanding noted above, namely that it comes in degrees.

This distinction sounds unfamiliar, and indeed I do not claim that the term 'complexity' is ordinarily used this way in English. I think we are more likely to use the term 'depth' to describe the feature I am interested in, but I shall give some reasons below why it might be best to keep that word for another purpose. In any case, however, it seems to me that the claim that emotions come in degrees in this way is actually obviously true. It is possible to feel a thin sort of horror at the Purges knowing only a few basic facts about them; it is possible to feel a vastly more complex sort of horror, if, say, one had witnessed them, or if one has studied a great many first-hand narratives about them. This is no less true in happier cases. Suppose I see the facade of Vignola's Church of Il Gesù. I think 'Quite a nice old church' and feel a very thin, insubstantial sort of admiration for it. But then I stop and study it carefully. I read Wölfflin's famous study of its composition. I sketch some of the details. I come to grasp what grounds its beauty far better. My emotion does not change type - it remains admiration - but it grows in what I am calling its complexity, which in normal usage we might equally call its depth

or richness or sophistication. It seems to me to be an advantage of my theory that it can explain what is going on here.[109]

How does this relate to the constitutive rationality of emotion? A useful parallel here is understanding, regardless of what account of understanding we endorse, or of whether a reduction of emotion to understanding ultimately proves successful. We may distinguish between two ways in which understanding varies. One is degree, which, as we have said, is a matter about how much grounding information one has grasped. The other is whether it is correct or mistaken. It is possible not to understand much about something, but for one's understanding to be correct as far as it goes. Or it is possible for one to have an extremely sophisticated (subjective) understanding of something, but for this understanding to be totally mistaken. The ideal situation is to understand a lot correctly, but one is making no mistake if one just understands a little correctly; whereas having a complex wrong understanding is probably even worse than having a simple one. This sounds very awkward, because it is more normal to use 'understanding' in its objective (factive) sense, and so a 'sophisticated wrong understanding' sounds like an oxymoron. But if one sets aside the linguistic strangeness, it seems to me to be an unproblematic distinction that any account of understanding will make. Indeed, given the truth of doxastic reductionism about understanding, it would simply be the commonplace distinction between having lots of true beliefs, a few true beliefs, a few false beliefs, and lots of false beliefs.

I think we can extend the same model to emotion. A complex emotion has a lot of content about the ground of the evaluative property that needs to be correct; a simple emotion has correspondingly less. Again, although this may sound like a convoluted when described abstractly, I think it is actually a familiar enough distinction. Ruskin wrote an enormous and detailed book about why he hated the way the Old Masters painted landscapes and why he admired the way Turner did so.[110] Ruskin's emotions on this subject were clearly enormously complex, and took in a huge amount of information about European painting over several centuries. For Ruskin's hatred of the Old Masters' technique to be constitutively irrational would be for a great deal of mental content to be inaccurate. By contrast, for my annoyance at a noisy infant sitting near me on a train to be irrational would be for only a small amount of mental content to be inaccurate - there just would not be much mental content there to be wrong.

---

[109] Although this theme has not been much discussed by analytical philosophers, it seems to me that it has been by others. See esp. F.R. Leavis, 'Reality and Sincerity', *Scrutiny* 19:2 (1952): 90-8 and '"Thought" and Emotional Quality', *Scrutiny* 8:1 (1945): 53-71. It should be emphasised here that to grasp more is not necessarily to know more: as I have argued, the grasping involved in emotion is not constituted by belief. Someone who is naturally intensely musical but who has never studied music formally might have a very complex emotion upon listening to *The Art of Fugue* for the first time, even though they could tell us almost nothing about what it was that makes the piece important.

[110] I refer to John Ruskin, *Modern Painters* (London: Smith, Elder and Co, 1843-1860), 5 vols.

Return now to understanding. We do sometimes ask the binary question: 'Do you understand x'? What does this mean, if understanding is a matter of degree? The answer, I take it, is that in some contexts we are interested in whether people have certain degrees of understanding, and in those contexts we ask the binary question as a shorthand for 'Do you understand x to such-and-such degree?'. Hence 'Do you understand how the lawnmower works?' surely often serves as a shorthand for 'Do you understand how the lawnmower works well enough to operate it?'. In some cases, of course, we feel that the binary question presupposes a clearer norm about what degree of correct understanding is sufficient, and so we may respond to the question 'Do you understand French?' with the answer 'To some extent'.

In a like fashion, we may sometimes impose a 'cut-off' point on extremely simple emotions where we say that there is no longer enough grasp on the grounding of the evaluative property present for it to count as a real emotion. Return to horror at the Purges. As I have said, we can distinguish a complex state of being horrified by the Purges, such as say Koestler felt, to the very simple sort of horror that most of us may occasionally feel. But once we reach a certain level of simplicity - once the grasp on what made these events horrible becomes sufficiently cursory - we may wish to say that it stops counting as horror, becoming something more like the mere belief that something was horrible, without any significant grasp of what made it so. I do not think we can make any general answer to the question 'How much complexity is needed for it to count as emotion?', any more than we can say what degree of understanding is necessary for it to count as understanding an object: it depends on the emotion and on the context. Hatred surely has a higher cut-off point than annoyance, and love surely a higher one than liking. Hating a person will have a higher cut-off point than hating a tune, and likewise for love. We saw in Section 2.1 that, in the apparently unique case of fear, it is possible to have the emotion rationally with no information at all except that one is in danger. But in the case of all or nearly all other emotions, it is necessary, as I have insisted since Section 1.2, to have at least some level of grasp on the ground of the evaluative property.

'Complexity' should be distinguished from what I shall call 'depth'. As I use the term, an emotion is deep in proportion to how far the subject cares about the object (in ordinary English, I think either 'deep' or 'intense' can both be used to denote this feature). Complexity and depth are independent characteristics, but they tend to correlate. If one cares little about an object, one is unlikely to take the trouble to study it deeply, so one's emotion will usually remain simple. If one cares a great deal about an object, one is much more likely to study it deeply, and one's emotion will become more complex. But there are surely refined and underemployed people who have complex emotions about objects for which they care little; and there are surely intense and naïve people who have many deep emotions about objects of which they have very little grasp.

3.4 Recalcitrance

One area of difficulty for doxastic cognitivism about emotion is 'recalcitrant' emotion.[111] Recalcitrant emotions are emotions that persist in spite of one's being aware of reasons that they are inaccurate. Two related but distinct objections to belief theories have been developed on the basis of the existence of such emotions: one focusses on the quantity of recalcitrance, the other on the mere fact of it. In this section, I ask whether non-doxastic cognitivism succumbs to these problems. I argue, in some cases tentatively, that it has promising resources for avoiding them.

The first version of the recalcitrance objection, in its original form, ran that emotion cannot be belief because emotion is a great deal more prone to recalcitrance than belief is. Although there are recalcitrant beliefs, like delusions, they are fairly rare. Recalcitrant emotions, by contrast, are endemic: this fact, it is said, is evident everywhere, from the clash between passion and reason portrayed in so many literatures, to our everyday experience of enduring troublesome emotions that we do not wish to have. My view obviously avoids this objection in its original form, because I do not claim emotions to be beliefs. But it remains possible that it succumbs to a revised version of this objection. This revised version runs that:

(1) Emotional recalcitrance is far more common than doxastic recalcitrance;

(2) If emotional recalcitrance is far more common than doxastic recalcitrance, it is implausible that emotions have a constitutive aim of accuracy. The idea here is that beliefs are the paradigm cognitive mental state, and should give us at least a rough idea of how much recalcitrance cognitive mental states admit of. A state which pervasively fails to live up to its alleged constitutive probably does not, it is claimed, actually have that aim;

(3) Therefore, it is implausible that emotions have a constitutive aim of accuracy.

In responding to this, one possibility would obviously be to attack premiss (2). How implausible really is it a given mental state is subject to a very demanding norm which it consistently fails to live up to? But I am not going to press this line of argument. Instead, I am going to argue against premiss (1), suggesting that emotions are far less prone to recalcitrance than they appear to be. What I propose to do is to highlight three ways in which emotions may *seem* to be recalcitrant without actually being so. The general theme of the answer is that, while we no doubt have many rationally problematic emotions, true recalcitrance is not obviously more common than we would expect of a mental state that is subject to rational norms (assuming for the sake of argument, in line with premiss (2), that we do have some

[111] See originally Patricia Greenspan, 'Emotions as Evaluations', *Pacific Philosophical Quarterly* 62 (1981): 158-169 at pp. 161-162; cf. her 'Emotions, Reasons and "Self-Involvement"', *Philosophical Studies* 38 (1980): 161-168 and her *Emotions and Reasons*, pp. 17-20. See also Bennett Helm, *Emotional Reason*, pp. 41-46; Griffiths, 'The Degeneration of the Cognitive Theory of Emotion', p. 300; Roberts, *Emotions*, pp. 89-91.

such expectations). Throughout I refer to apparent recalcitrance that is actually not recalcitrance as 'pseudo-recalcitrance'.

The first kind of pseudo-recalcitrance is best understood by starting with some examples. I might be pleased by my friend's losing his job because I found his erstwhile success threatening; or I might resent his domestic happiness because of my lack of it; I might despise someone for their failure or vulnerability; I might glory the suffering of my rival; and so on. In all these cases, it is perfectly possible to be aware that one's emotion is profoundly rationally problematic without the emotion's going away. So it may look as though these are all cases of recalcitrant emotion. But this is mistaken: they are a distinct phenomenon. None of these emotions need actually be *inaccurate*, in the way that fear of Fido is inaccurate: the world really may matter to me in the way that I represent it as mattering, and there may be no fact about the object of the emotion that could be adduced to persuade me that my emotion was mistaken. If they are not inaccurate, these emotions cannot be constitutively irrational. I have no complete account of what really is wrong with them, but at a rough first stab, it is the fact that the world matters to me in these ways in the first place. Perhaps some emotions are *always* like this, like envy or hatred: there might be something about envy in virtue of which it is always wrong, but it is not the case that envy is in every case analogous to phobic fear. Cases like these are puzzling in many ways, but they are not true recalcitrance, and so they do not weigh in favour of the recalcitrance objection against my account.

The second kind of pseudo-recalcitrance consists in cases in which we have practical reasons not to have emotions. As we discussed in connection with Michael Brady's account in Section 1.5, emotions 'capture and consume' attention.[112] The presence of a given emotion may make us highly sensitive to its object, but problematically insensitive to others. Suppose I have to rescue a child from a burning building. Fear is a perfectly rational emotion to feel in this situation, but focussing solely on the danger to myself will capture and consume attention needed elsewhere. As we shall discuss in Chapter Four, emotions can also be painful in problematic ways. Medical professionals often work with people for whom compassion is constitutively rational. If they feel all the rational compassion they can, however, they are wont to suffer from 'compassion burnout', a kind of mental exhaustion on account of the painfulness and tension that compassion involves. Aside from other disadvantages, compassion burnout will render them unable to help their patients, and so they often have good reason to avoid feeling it. In cases like these, then, we may have emotions in spite of having strong reasons not to. This could be taken to be recalcitrance, but it is not, for the emotions are not inaccurate. So again, no support for the recalcitrance objection can be derived from such cases.

---

[112] See esp. Brady, *Emotional Insight*, ch. 4.

A third case of pseudo-recalcitrance arises when we have emotions that *are* accurate, but for which we cannot articulate the reasons.[113] This is best illustrated with the results of two well known studies. One examined the ability of neonatal nurses to diagnose infants with sepsis.[114] It found that nurses were able to make these diagnoses at a high level of success. When they were asked to list the diagnostic criteria for sepsis, however, they were unable to do so accurately. Further investigation seems to have revealed that the textbooks from which the nurses had been taught did not actually include the diagnostic criteria of sepsis in children. The researchers concluded that the nurses made a reliable 'intuitive judgement', despite being unable to support it with 'explicit knowledge'.

In another interesting study, participants were presented with five posters: a reproduction of Monet's *Nymphéas*, a reproduction of Van Gogh's *Irises*, a cartoon depiction of animals in a hot-air balloon, a picture of a cat balancing on a tightrope with the caption 'Gimme a break', and a picture of a cat standing by a fence with the caption 'One step at a time'.[115] The control group were simply invited to give the pictures a score, and then to choose one to take home. The other participants were told to list reasons for liking or disliking each poster, after which they rated them and took one home. Almost all members of the 'no reasons' group ranked the art posters higher and took one of them home, while members of the 'reasons' group ranked the frivolous posters relatively higher and were more likely to take one of them. Participants were then contacted several weeks later and asked whether they had kept the poster, whether they had put it up, how much they liked it, and how much they would be prepared to sell it for. The no reasons group were more satisfied with their posters, especially those who had chosen the art ones. The researchers concluded that the participants' explicit reasoning was actually worse than their intuitive judgements at telling them what they would like in this case. Participants found it harder to articulate what they like about the art posters than it was to articulate what they liked about the humorous posters, so if they were encouraged to make their decision on the basis of the reasons they could articulate, they actually tended to make a worse one than if they relied on their initial hunch.

The feelings that the subjects were relying on in these cases may not themselves be emotions. The 'intuitive judgement' of the nurses was not in first instance probably was not, because its content was not an evaluative property; the likes and dislikes of the participants in the second study might have been emotions of a kind, but not exactly a central case. But this is not important for our purposes. The

---

[113] The following discussion is indebted to Kelly Epley's recent paper 'Emotions, Attitudes, and Reasons', pp. 262-269.

[114] B. Crandall and R. Calderwood, 'Clinical Assessment Skills of Experienced Neonatal Intensive Care Nurses' (Final Report). Prepared Under Contract 1 R43 NR0191101 for the National Center for Nursing, NIH (Yellow Springs, OH: Klein Associates Inc., 1989).

[115] T. Wilson et al, 'Introspecting about Reasons can Reduce Post-Choice Satisfaction,' *Personality and Social Psychology Bulletin* 19:3 (1998): pp. 331–339.

point of these examples is to present us with a model: on a plausible characterisation, what happens here is that (a) the reasons that subjects are able to articulate appear to support one action but (b) subjects make a non-doxastic judgement in favour of another action and (c) the non-doxastic judgement is in fact constitutively rational. In such cases the non-doxastic judgements might *look* irrational or non-rational: imagine here the patronising doctor reprimanding the nurses for relying on fuzzy intuition rather than proper reasoning. But this appearance is misleading.

This model is readily applied to paradigmatic emotions. Kelly Epley has an excellent discussion of Susan, who is engaged to Todd, but who has misgivings:

> We can imagine Susan persisting in her judgment, but being unhappy and joyless as she prepares for the wedding. She anticipates feelings of regret for following through with it, and she is hoping for a convenient way out of the situation. Eventually, she acts on her feelings and calls off the wedding.[116]

Susan's emotions are in tension with her beliefs about what she should do, and it may be that she cannot articulate reasons to support them. But it is also possible that her emotions are sensitive to a much wider range of reasons than she is able to articulate, and that she is right in acting on them. It seems to me that this is an experience we have fairly often, and one that may explain away many cases of apparent recalcitrance.

This is not a decisive refutation of the 'too much recalcitrance' argument: because we do not have quantitative evidence on how often these kinds of pseudo-recalcitrance occur, and because the objection does not set a precise threshold at which the quantity would become acceptable, a decisive refutation is probably impossible. But it shows we have good general ways of explaining away many cases of apparent recalcitrance, and thus reason to be sceptical that levels of true recalcitrance are problematically high.

There is also a second recalcitrance objection. This objection does not rest on the quantity of recalcitrant emotion, but rather on the way in which recalcitrant emotion is characterised when it arises, however rare it may be. I cannot stand snakes; I see the pet snake, which I know to be harmless, slithering towards me; and despite my knowledge, I feel fear. According to the belief theory, on its simplest version, emotions are beliefs: to fear the snake is, roughly, to believe it to be dangerous. So if I fear the snake while believing it to be harmless, I thereby believe it to be dangerous and believe it not to be dangerous. The problem for the belief theory, according to this objection, is that this seems to attribute too high a level of irrationality to me. While recalcitrant emotions are surely irrational, it seems implausible that they are as irrational as consciously believing a contradiction. Nor is it obvious how any of the more complex and sophisticated belief theories discussed in Chapter Two can get out of this

---

[116] Epley, 'Emotions, Attitudes and Reasons', p. 256. The example is in turn based on one in Nomy Arpaly, 'On Acting Rationally Against One's Best Judgment,' *Ethics* 110:3 (2000): 488–513.

problem: on all of them, emotions include evaluative beliefs, and these beliefs are inconsistent with the other beliefs that one has in cases of recalcitrant emotion.

This objection may seem to carry over to non-doxastic cognitivism. On my view, although belief and emotion are different states of mind, they both have a constitutive aim that includes accuracy: I am responsible for the accuracy of my emotions in the same way that I am responsible for the accuracy of my beliefs. It might seem, then, that I must predicate the same sort of irrationality of people who have recalcitrant emotions as I must of those who consciously have inconsistent beliefs: both have mental states of which it is true that (a) their content has the constitutive aim of accuracy, (b) their content is mutually inconsistent, and (c) I am aware of the inconsistency of their content. The question, then, is whether my account implies a similarly high level of irrationality on the part of the subject.

This objection is more powerful than its predecessor, and my response to it is somewhat tentative. The core thought of my response is that it is simply not obvious that I have to attribute the same level of irrationality to the subject of recalcitrant emotion as the belief theorist does. It all depends on what it is about consciously holding inconsistent beliefs that makes it so irrational. If it is the fact that they are mental states with a constitutive aim of accuracy which we know to represent the world in inconsistent ways, then clearly the problem carries across to my theory, since this property is shared by the subject of recalcitrant emotion on my view. But it is not obvious that this is correct. It might be the sheer fact that they are both beliefs: that is, it might not be so irrational to have two mental states with inconsistent content if those are mental states of different kinds. On my account, emotion is in some ways quite different to belief: often, as we saw in Chapter Two, it is fused with experience, such that one might impressionistically say that to have the emotion just is to experience the object in a certain way, as infused with some kind of importance. Emotions are also plausibly harder to control than beliefs. Maybe it is just less irrational to have mental states with inconsistent content when the content is presented in such different formats.

I am inclined to think this is true, but I find this a hard claim to prosecute. The ideal way to support it would be to look at a parallel case in which we consciously have mental states that are (a) of different kinds but (b) share a constitutive aim of accuracy and (c) have inconsistent content. If we found that, in those cases, the subject's level of irrationality seemed to be lower than it is in the case of consciously having inconsistent beliefs, then we would have good reason to believe that my account of emotion involves the second recalcitrance objection (a 'companions in guilt' argument). The difficulty here is such parallel cases are quite hard to find - it is effectively the lack of mental states that share this characteristic that pushed us towards a non-reductionist account, as I argued in section 3.1. As I have mentioned, philosophers have occasionally argued that some or all kinds of desire or pleasure are mental states of this kind. If that is the right characterisation of desire or pleasure, I think these examples would support my account. It is clearly possible to consciously desire something while consciously believing it to be bad without being as irrational as one is in consciously believing a contradiction. If, then, desire represents its object as good, and has a constitutive aim of accuracy - if, that is, it is constitutively

irrational to desire something one knows to be bad - then desire for something known to be bad constitutes a companion in guilt for my position. Similar considerations apply to pleasure. But it is so contentious that desire or pleasure *is* a non-doxastic mental state that nonetheless has a constitutive aim of accuracy that these companions in guilt do not offer much support to my position.

A more promising companion in guilt is aesthetic experience.[117] Suppose I believe on testimonial evidence that the Adagietto of Mahler's Symphony No. 5 is kitsch, but when I listen to it, it seems to me to be full of noble yearning. I know that I am hearing the Adagietto wrongly, but I cannot work out how to hear it rightly. I think this is quite a promising companion in guilt for my account of emotion. Plausibly, I have here two mental states whose constitutive aim includes accuracy, namely belief and aesthetic experience. There is obviously something rationally unacceptable about my mental states: I know that one of them involves my making a kind of mistake. But I am not irrational in the same way or to the same extent that I am in the case of consciously endorsing two inconsistent beliefs. So plausibly I end up with a level of irrationality which is about right for recalcitrant emotion: higher than perceptual illusion, lower than consciously endorsing p and not-p.

I think this has a potential as a companion in guilt. But it is contentious, because it rests on presuppositions about aesthetic experience that are not universally accepted: not everyone accepts that the aesthetic experience itself has constitutive aims. If it does not, then aesthetic experience would be more like standard kinds of perception, and its divergence from our beliefs would be rationally problematic only in the way a standard perceptual illusion is. I cannot settle the nature of aesthetic experience here, and so this comparison must remain to some extent an area for future work. Nevertheless, it is clear that my account has more resources for responding to the recalcitrance objection than did the classical belief theory. The level of irrationality that we should attribute to subjects when they have mental states that (a) have the constitutive aim of accuracy, (b) are mutually inconsistent and (c) are of different kinds is, at worst, an open question. And that is certainly an improvement on the predicament of the belief theorist.

3.5 Animal Emotion

One argument that was influential in the decline of belief theories was the supposed inability of the latter to accommodate the emotions of animals.[118] The objection has been developed in slightly varying forms, but a useful reconstruction runs as follows:

[117] cf. e.g. Scruton, *The Aesthetics of Music*, esp. ch. 8.

[118] See esp. Deigh, 'Cognitivism and the Theory of Emotion', esp. pp. 842-852 and Roberts, 'Propositions and Animal Emotion'. Cf. Roberts, 'The Sophistication of Animal Emotions' and John Deigh, 'Primitive Emotions' in Robert C. Solomon (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotion* (Oxford: Oxford University Press, 2004).

1. Animals do not possess beliefs.
2. If animals do not possess beliefs, their emotions are not partly or wholly constituted by beliefs.
3. If animal emotions are not partly or wholly constituted by beliefs, human emotions are not partly or wholly constituted by beliefs.
4. Therefore, human emotions are not partly or wholly constituted by beliefs.

An expanded version of this argument could be offered against my account, with 'beliefs' replaced by 'cognitions'. One way for me to resist this argument would be to argue that it is more plausible that animals have non-doxastic cognitions than it is that they have beliefs. I think this line of argument would have promise. But I also think it is unnecessary for me to argue this, because the argument is unsuccessful even when used against the belief theory. It is, I shall argue, probably unsuccessful even against the *pure* belief theory; and it is certainly unsuccessful against hybrid cognitivist theories - a category into which both my account and most classic belief theories accounts fall. In other words, I think a considerably more confident rebuttal is possible here than it was in the case of the recalcitrance argument. In what follows, I am mostly going to stick to defending the original belief theory against the 'argument from animal emotions': my reason for this is that the scientific literature I am going to appeal to mostly concerns animal belief, making it confusing to refer constantly to non-doxastic cognitions. But the defence carries across easily to non-doxastic cognitivism too, as shall become clear.

Let us turn back to the above argument. It is valid, and the second premiss is clearly true. The key work is thus being done by the first and third premisses. We shall start by considering premiss three. Rejecting this premiss while accepting the others would be a sort of fall-back option, in that it would mean accepting that the belief theory is not true of animals. Let us call this a 'human-only belief theory'.[119] I am ultimately going to go on to argue that we probably do not need to accept a human-only cognitivist theory, because we can be cognitivists about animals too. But my present question is whether there is anything implausible about the human-only belief theory, supposing that holding a belief theory about animals turned out to be unviable. Roberts, describing what he takes to be a respect in which his construals theory is superior to the belief theory, argues as follows:

> The construal account of emotion admits that in being non-propositional, animal emotions differ significantly from their human counterparts. But since construals are not necessarily propositional, it leaves room for animal analogues of many human emotions. All that is needed for emotion, on this account, is a power of perception which is not merely sense-perception, but some organization of sense-perception that can impinge on some concern (perhaps instinctual, perhaps learned) of the animal. If this organization is similar enough to the

[119] Several existing belief theories are explicitly human-only. See e.g. Nussbaum, *Upheavals of Thought*, esp. ch. 2; Robert Gordon, *The Structure of Emotions: Investigations in Cognitive Psychology* (Cambridge: Cambridge University Press, 1987), esp. pp. 71-72; cf. also Ronald de Sousa, *The Rationality of Emotion* (Cambridge Mass.: MIT Press, 1987), pp. 100-101 and 181-182.

> organization that is achieved partly through propositions in the case of the animal-emotion's human analogue, we are justified in ascribing something like the standard human emotions to animals.[120]

Similarly, Deigh argues against giving different accounts of the nature of human and animal emotion as follows:

> [such a view] justifies cognitivist inattention to the emotions of beasts and babies only if it does not derive from a distinction between human beings and other animals that is as implausible as Descartes's. It could not, for instance, rest on the supposition that the human soul was ontologically distinct from the souls of animals who lacked reason. [...] Whatever differences in psychological capacities exist between humans and beasts, or grown-ups and babies, they cannot, consistent with evolutionary biology and developmental psychology, imply that mature human thought and feeling are phenomena utterly incomparable to their bestial and infantile counterparts.[121]

What both Roberts and Deigh insist upon, then, is that a theory of emotion must be consistent with the continuities that clearly exist between human and animal emotion. Animal emotions are clearly similar to human emotions in quite a few ways. They are a response of some kind to important situations, and they play a role in motivation. They come in types, and some of those types, like fear, rage and joy, are common to humans and animals. An account that characterised human and animal emotion so differently that these commonalities were inexplicable would be unsuccessful.

All this is surely correct.[122] The problem with this as an objection is simply that it seems fairly easy for human-only belief theories to do all these things. Consider an extreme sort of human-only belief theory, according to which human emotions are entirely constituted by a belief, and animal emotions are entirely constituted by bodily feelings. Even in this case, the human-only belief theorist has some resources for explaining the commonalities between human and animal emotion. They may say that there are good evolutionary reasons for both humans and animals to develop mechanisms for responding to important situations, and that there are obvious evolutionary reasons why their responses would follow similar patterns. For instance, both humans and animals are likely to have evolved mechanisms that motivate them to avoid danger in similar ways, even if those mechanisms work very differently. The human judges danger to be near and takes evasive action; the animal gets a rush of displeasure when it is near the danger, and moves away from the danger to reduce the displeasure. Thus

[120] Roberts, 'Propositions and Animal Emotions', p. 153.

[121] Deigh, 'Cognitivism and the Theory of Emotion', p. 349.

[122] Or, at least, it is surely correct for at least some of the emotions of at least some animals. The continuity between the emotions of humans and the emotions - if emotions they be - of frogs and beetles are not so pretheoretically obvious. But Deigh and Roberts's assumption is surely plausible so far as the many of the emotions of higher mammals are concerned, and that is enough for present purposes.

quite different mechanisms might produce similar responses to similar situations, such that we end up calling both of them 'fear'. It is not, after all, unknown for similar functional programmes to be implemented in very different ways in different animals. It is sometimes said that cephalopod nervous systems are so different to ours that the whole of cognitive neuroscience will have to be reinvented to describe them. Yet some of the functional programmes that are implemented through these systems have striking similarities to our own, such that scientists are prepared to ascribe pain, pleasure, learning, communication and thought to the octopus.[123]

In the case of more moderate human-only belief theories, which allow that human emotions have components other than belief, these resources are much more abundant. Even if it is true that animals do not have beliefs, it is thus perfectly possible that they share other elements of emotion with us. If human emotion is comprised of a mixture of hedonic states, perceptual states, bodily feelings, and beliefs, animal emotion may just be comprised of hedonic states, perceptual states and bodily feelings. So no radical discontinuity is posited. As noted above, it is to theories like these that my own account is parallel: I suggested that emotions may include non-cognitive perceptual elements in Chapter One, and in Chapter Four I will argue that they include non-cognitive hedonic elements.

I think, then, that it would not be greatly problematic for the belief theory (let alone to my own account) to be restricted to human emotions, thereby allowing that animal emotions are structured differently. But I am not sure it is necessary to make this concession. Roberts and Deigh assert that animals cannot form beliefs. But is this true?

There have certainly been philosophers who thought so. Many canonical figures, including Chrysippus, Aquinas, Descartes, Leibniz and Kant, argued that the kinds of mental states that animals can have are very limited, certainly excluding beliefs or judgements of any kind.[124] This view still had some influential advocates in the decades before Deigh and Roberts developed their objections to belief theories: the best-known was Donald Davidson, and others included John McDowell and Wilfrid Sellars.[125] The view that animals cannot make belief has however never been universally accepted, and

---

[123] A. Packard, 'Cephalopods and fish: The limits of convergence', *Biological Reviews* 47:2 (1972): 241–307; Graziano Fiorito and Pietro Scotto, 'Observational Learning in *Octopus vulgaris*', *Science* 256:5056 (1992): 545–547; J.K. Finn, T. Tregenza and N. Tregenza, 'Defensive tool use in a coconut-carrying octopus', *Current Biology* 19:23 (2009): 1069-1070.

[124] For a historical survey, see Richard Sorabji, *Animal Minds and Human Morals: The Origins of the Western Debate* (Ithaca, NY: Cornell University Press, 1995). The view that animals do not have rationality has never been universal among philosophers: it was rejected by, for instance, Theophrastus, Porphyry, Galen, Gassendi, Locke and Hume.

[125] Donald Davidson, 'Thought and Talk' in his *Inquiries into truth and Interpretation* (Oxford: Clarendon, 1984); 'Rational Animals' in E. Lepore and B. McLaughlin (eds), *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (New York: Basil Blackwell, 1985) and 'The Emergence of Thought', *Erkenntnis* 51 (1997): 7-17;

in recent decades there has been a very general shift away from it. Davidson's account of animal minds has been widely and cogently criticised,[126] and a range of positive arguments for animal belief have been developed. Many philosophers, usually presupposing currently widespread functionalist accounts of the mind, hold that the fact that positing beliefs obviously allows us to explain and predict animal behaviour gives us good reason for supposing such beliefs to exist.[127] To give a rough illustration: the fact that 'Fido believes there is a cat in the tree' seems to be an obviously good explanation of Fido's circling the tree and barking up at its branches; on this view, this constitutes a good reason for supposing that Fido does indeed have that mental state. Other philosophers hold that we should infer animal beliefs from the conjunction of these features of animals' behaviour and the fact that animals have perceptual organs similar to ours.[128] A third school holds that we ought to accept what our best scientific evidence says on the subject: since animal scientists now overwhelmingly do hold that animals have beliefs, they conclude that we ought to think so too.[129] Although there is lively debate between these accounts, as

---

McDowell, *Mind and World*; Wilfrid Sellars, 'Intentionality and the Mental' in H. Feigl, M. Scriven and G. Maxwell (eds), *Minnesota Studies in the Philosophy of Science* (Minneapolis, MN: University of Minnesota Press, 1957), vol. 2.

[126] See e.g. Michael Tye, 'The Problem of Simple Minds: Is There Anything it is Like to be a Honey Bee?', *Philosophical Studies* 88 (1997): 289-317; José Luis Bermúdez, *Thinking Without Words* (Oxford: Oxford University Press 2003); Peter Carruthers, 'Meta-Cognition in Animals: A Skeptical Look', *Mind and* Language 23 (2008): 58-89.

[127] e.g. Stephen Stich, 'Animals Have Beliefs?', *Australasian Journal of Philosophy* 57 (1979): 15-28; Fodor, *Psychosemantics: The Problem of Meaning in Philosophy of Mind*; Peter Carruthers 'On Being Simple Minded', *American Philosophical Quarterly* 41 (2004): 205-220. For some more recent moves in the debate, see Peter Carruthers, 'Invertebrate Concepts Confront the Generality Constraint (and win)' in Robert Lurz (ed.), *Philosophy of Animal Minds* (Cambridge: Cambridge University Press, 2009). Daniel Dennett also argues that we can infer animal beliefs from the behavioural evidence, though his presuppositions were not those of contemporary functionalism. See his *The Intentional Stance* (Cambridge, MA: MIT Press, 1987).

[128] John Searle argues for this kind of view in his *The Rediscovery of Mind* (Cambridge, Mass.: MIT Press, 1992) and 'Animal Minds', *Midwest Studies in Philosophy* 19 (1994): 206-219.

[129] See e.g. Colin Allen and Marc Bekoff, *Species of Mind* (Cambridge, MA: MIT Press 1997) and Bermúdez, *Thinking Without Words*. The work of Donald Griffin has been influential in shifting scientists' approach to animal minds; see esp. his *Animal Thinking* (Cambridge, MA: Harvard University Press, 1984) and *Animal Minds: Beyond Cognition to Consciousness* (Chicago: Chicago University Press, 2001).

well as debate concerning what sort of beliefs animals can have,[130] few philosophers still claim that animals cannot have beliefs at all.[131]

To definitely settle which of these accounts is true, if any, lies beyond the remit of the present essay, especially because it would probably require the solution of more general problems in the philosophy of mind. I will say only that the anti-doxastic assumptions about animal minds were contentious even when they were made by philosophers of emotion in the early 1990s, and they are now usually regarded as discredited: strangely enough, then, the idea that animals do not have belief has been undermining belief theories of emotion just as it fades away as a seriously held view in its own right. Once this fact is appreciated, we should surely regard the argument against belief theories from animals' supposed lack of beliefs as unpromising, at least without some justification of its key premiss.[132] And supposing that animals have beliefs, there seems no obvious reason to deny that they might have non-doxastic cognitions too.

One further note. In his paper, Deigh also makes reference to the emotions of infants, which he believes to be problematic for belief theories in the same way that the emotions of animals are. Is the same sort of response possible? The first part, I think, carries across straightforwardly. One could give a non-doxastic theory of infant emotion without undermining the plausibility of the claim that adult emotion is partially belief-constituted, explaining the continuities between infant and adult emotion in terms of the elements of adult emotion that have been conceded to be non-belief-constituted. It may be that the second part of my response could be carried over too, that is, that a parallel argument could be developed that infants are capable of belief despite their lack of language. Unfortunately, however, there is remarkably little philosophical literature on infant cognition, despite the immensely rich results of developmental psychology in recent decades.[133] A systematic philosophical examination of research

---

[130] See for instance the debate on whether animals can think about their thoughts in Bermúdez, *Thinking Without Words* and Robert Lurz, 'In Defense of Wordless Thoughts about Thoughts', *Mind and Language* 22 (2007): 270-296.

[131] For summaries of this debate, see Robert Lurz, 'The Philosophy of Animal Minds: An Introduction' in Robert Lurz (ed.), *Philosophy of Animal Minds* (Cambridge: Cambridge University Press, 2009) and Hans-Johann Glock, 'Animal Rationality and Belief' in Kristin Andrews and Jacob Beck (eds), *The Routledge Handbook of Philosophy of Animal Minds* (London: Routledge, 2017).

[132] The evidence that animals have beliefs tends to be somewhat less clear in the case of cognitively simple animals like wasps, maggots and so on. But it is similarly less clear that they have emotions that are continuous to ours, so this fact does not pose serious problems for the belief theory.

[133] An exception is Bermúdez, *Thought Without Words*, in which it is argued compellingly that infants' absence of language need not prevent them from thinking. The discussion in Gareth Matthews, *The Philosophy of Childhood* (Cambridge MA: Harvard University Press, 1994), chs. 2-4 is also promising, but inconclusive for our purposes, being concerned chiefly with a critique of Piaget. Nussbaum has an interesting discussion at *Upheavals of Thought*, ch. 4, but it does not help with our present difficulty.

on infant minds and its implications for the philosophy of emotion would be a task of immense complexity and interest, but one that, alas, lies beyond the scope of the present enquiry. I am therefore going to bracket the question of whether infant emotion is belief-constituted as an area for future enquiry, with due acknowledgement that this leaves open, for the time being, the possibility that my account will have to concede some ground in this area.

3.6 Helm's Account

I am going to close this chapter by considering the account offered by Bennett Helm.[134] Helm offers several cogent criticisms of existing accounts, which seem to set him up for just the sort of account I am proposing. In the event, however, he does not offer such an account, and the alternative he suggests seems to be unable to deal with the very problems he raises with existing theories. This means my dialectical position vis-à-vis Helm's view is rather complicated, and so I have chosen to deal with it individually, rather than trying to shoehorn it into one of the earlier critical chapters.

Helm's objection to existing theories is based on how they understand the rationality of recalcitrant emotions.[135] It is not irrational to have a perceptual illusion, Helm notes, even if one knows that an illusion is what it is. But it is irrational to have an emotion that misrepresents its object when one knows that this is what it is doing. So emotion cannot be a kind of perception. So far, this is closely related to the argument I developed in Chapter One, though Helm does not distinguish constitutive rationality from other kinds. Helm goes on to argue that although it is irrational to have recalcitrant emotions, it is much less irrational than it is to have 'recalcitrant judgements', that is, to knowingly make two judgements that are inconsistent with one another. To be afraid of something that one knows to be harmless is irrational, but it does not actually involve conscious incoherence. So fear cannot be a judgement, and judgement theories fail.

Although Helm seems to leave conceptual room for 'judgements' to be something different from beliefs, his discussion generally seems to treat them as identical, or as distinct only in some way that is unimportant to the matter at hand.[136] So although Helm's theory appears to be inconsistent with mine, this might be an illusion: it might be that Helm's view is at least consistent with, and perhaps simply is, what I call non-doxastic cognitivism. And at certain moments, Helm seems minded to go this

---

[134] Helm, *Emotional Reason*, esp. chs. 2, 3 and 6; see also his 'Emotions as Evaluative Feelings' and 'Emotions and Recalcitrance'.

[135] Ibid., pp. 36-46. Helm also argues that belief theories have difficulties saying what distinguishes emotion from other beliefs or conjunctions of other beliefs and desires.

[136] They are apparently distinguished at ibid., pp. 41-42, but for the rest of ibid., ch. 2 they seem to be treated in the same way. Fortunately, my argument will not ultimately depend on what Helm means 'judgement', because Helm does not define his own positive view in terms of this word. So we do not need to be absolutely sure how he is using it for present purposes.

way. Confronted with an unsatisfactory choice between the view that emotions are judgements and the view that emotions do not involve assenting to their content because they are not judgements, Helm reasons as follows:

> In each case, the conclusion presupposes the implicit premise that all assent is judgmental. The way out is to deny this premise: emotions must be understood as a kind of assent if we are to make sense of rational conflict with judgement at all, but not a kind that can be reduced to judgment if we are to make sense of that conflict as something other than incoherence.[137]

Provided, as I say, that 'judgement' is read as synonymous with belief, Helm's conclusion here is consistent with my own view: the non-judgemental assents (or 'disclosive assents')[138] with which he identifies emotions might, for all that has been said so far, be my non-doxastic cognitions.

In fact, however, this does not seem to be the right way of interpreting Helm's considered view. Considering a proposal according to which emotions are characterised as a kind of perceptual belief, Helm argues as follows:

> Although our initial acquisition of perceptual beliefs involves a process in which we are passive, the assent implicit in the belief thus acquired is potentially one with respect to which we are active: as further evidence comes to light, we may actively revise or confirm such beliefs. Although we can have a kind of control over our emotions and so can be responsible for them (see §6.4 for an explicit account), exercising this control for the most part involves training and habituation. Consequently, the kind of control we can exercise over a belief is very different with respect to the activity of assenting than that which we can exercise over emotions.[139]

What this means is that Helm's 'non-judgemental assent' is not my non-doxastic cognition after all. In fact, his account of how we have more rational responsibility for the accuracy of our emotions than we do for our perceptual states is remarkably similar to Christine Tappolet's, which we discussed in Chapter One. Emotions are not actively revised or confirmed in response to evidence concerning their accuracy: the reason that knowingly inaccurate emotions are more irrational than known perceptual illusions is rather that we can control emotions through training and habituation - that is, in Tappolet's language, that our emotional dispositions are more plastic.

The 'explicit account' of how to overcome emotional recalcitrance that Helm mentions seems to confirm this. Helm argues that overcoming emotional recalcitrance involves cultivating two habits. The first is the habit of attending to the situation in the way that one judges to be appropriate.[140] One acquires this habit simply through attending in the right way, and one attends in the right way through the exercise of will. In the case of one's fear of harmless old Fido, then, this means that one consciously

---

[137] Helm, *Emotional Reason*, p. 45.

[138] Ibid., p. 67.

[139] Ibid., pp. 66-67.

[140] Ibid., pp. 190-191.

makes oneself attend to the various features of Fido in virtue of which one knows him to be harmless, until such attending becomes a habit, and consciously making oneself do this becomes superfluous. The second habit one needs to acquire to get rid of emotional recalcitrance is that of 'getting oneself to respond as if one has the relevant felt evaluations, acting on behalf of its focus, when, according to one' s conception of the situation in light of one's evaluative judgement, these felt evaluations are warranted'.[141] This too one acquires through acting in the appropriate way in spite of not yet being habituated to do so, which one instead does through an exercise of will. In the case of Fido this means consciously making oneself not avoid Fido until not avoiding Fido becomes a habit. Helm thinks that once these two responses have become habitual, the recalcitrant emotion will disappear.[142]

The upshot of all this is, so far as I can see, that Helm's account is susceptible to the same objections as Tappolet's was. Let us return to the eccentric billionaire. He has promised to save thousands of lives if I become furious with a teacup, as well as developing a disposition to do so in the future. According to Helm, what distinguishes the irrationality of knowingly inaccurate emotions from that of known perceptual illusions is that we have an element of rational control over emotions, and are therefore in a certain sense responsible for this. This element of rational control turned out to be a matter of our capacity to habituate dispositions to have accurate emotions. But in this case (assuming my reasons to save the lives of thousands are more stronger than my reasons against developing a bizarre emotional disposition), I do not have reason to habituate a disposition to have accurate emotions: indeed, I have the opposite. And since Helm has denied that our rational control of emotion is also a matter of revising and confirming them in view of evidence, he has no resources left over to deal with this case. There can be no sense, on his view, in which the teacup's total lack of agency constitutes a decisive reason that my fury with it is irrational. And this is an unacceptable result. Consider also the dying person, who has no further opportunities to habituate more accurate emotional dispositions. On Helm's view, it seems that they have no rational control over their emotions, and that there is therefore no sense in which having inaccurate emotions is irrational for them. But this is not so: it is just as irrational for them to fear Fido as it is for anyone else, at least in the primary sense in which we predicate irrationality of such emotion.

I conclude, therefore, that Helm's non-judgemental assents are not my non-doxastic cognitions after all: they are, in my sense, not cognitions at all. And that is what is wrong with them. The pursuit of reason does not stop with habituating our emotional dispositions: rational control of emotion is also a matter of revising and confirming emotions in the light of evidence. Helm is right to seek a third way between belief and perception, but he has erred a little too far on the side of perception. Emotions are not beliefs, but they involve us in the pursuit of reason in a somewhat stronger sense than Helm allows.

---

[141] Helm, *Emotional Reason*, p. 191.

[142] He appears to think that instantiating these two habits 'just is' having the emotion (ibid., p. 190), although my objection would work equally well if the relationship were causal.

3.7 Concluding Remarks

Most theories of emotion have offered reductions of emotion to some other mental state. There is one obvious possibility remaining for such a reduction, namely to understanding: emotion might be constituted by occurrent subjective grounding understanding of evaluative properties. This reduction will, however, go through only if certain controversial theories of understanding are true. If not, we will be left with an 'emotion theory of emotion', according to which emotions are one of our basic mental states. In either case, we will be left with a form of non-doxastic cognitivism: emotions are candidates for constitutive rationality, but they are not beliefs. I have shown that non-doxastic cognitivism has resources to withstand the objections that undermined the belief theories, either because those objections were actually unsuccessful in the first place, or because of respects in which it differs from its doxastic counterpart.

Chapter Four

Emotion, Pleasure and Pain

In this chapter, I investigate the nature of emotional pain and pleasure. Many, perhaps all, emotions seem to have a hedonic character. Some are pleasurable, like pride, joy, relief, wonder. Many are painful, like grief, shame, guilt, humiliation, anger, fear, despair, jealousy and envy. Some, plausibly, involve elements of both pleasure and pain: maybe hope and gratitude can be like this, and perhaps many other emotions can be so in some measure too. Crucially, this hedonic character seems to be necessary to them. Painless grief or pleasureless joy seem to be impossible: grief that was not painful would not be grief, and joy that was not pleasurable would not be joy. The question that concerns us is how theories can accommodate this fact.

In Section One of this chapter I distinguish some differing ways that theorists of emotion have sought to accommodate emotion's essentially hedonic character. A key distinction is drawn between accounts that identify emotion with a mental state that we are motivated to avoid or foster *per se*, and accounts that identify emotion with a mental state of which we are only motivated to avoid or foster the object. In Section Two, I consider what I call the 'Oblivion Problem' for the latter group: I argue that we can only accommodate the pursuit of oblivion to escape painful emotions by positing that painful emotions are the kind of mental states that we are motivated to avoid *per se*. I argue that this means that many existing accounts of emotion are in need of modification. I am going to speak mostly about pain, because it throws the problem into sharper relief, but a parallel argument can be mounted in the case of pleasure, an area that I shall discuss in more detail in Chapter Five.

4.1 Existing Approaches to Emotional Pain

It may seem to be a platitude that the painfulness of some emotions motivates us to avoid them. A useful model of a theory that can easily accommodate that platitude is the James-Lange.[143] On the James-Lange theory, emotions are constituted by bodily feelings caused by one's perception of core relational themes: fear just is the feeling of the heart racing and the stomach churning that is caused by one's perception of danger. Plausibly, the feeling of one's stomach churning is essentially painful, in such a way that one is normally motivated to avoid it. So the James-Lange theory can explain the painful or pleasurable quality of emotions by appeal to the painful or pleasurable quality of the states with which their theory identifies them. This explanation is also available to the Neo-Jamesian, who holds that such

[143] See again Lange, 'The Mechanism of Emotions' and James, 'What is an Emotion?'.

physical feelings can also have representational content, a traditional example being the 'gut reaction'.[144] Nothing in the Neo-Jamesian account implies that these physical feelings lose their original essential hedonic qualities in virtue of their being representational. So the Neo-Jamesian inherits from the James-Lange theory this simple and elegant explanation of the hedonic properties of emotion. Neo-Jamesians may, of course, wish to claim that the hedonic character of emotion is not exhausted by pains and pleasures of this kind, and that there are also other senses in which, say, grief is painful. But they have this explanation at least.[145]

For many other theories, however, this kind of approach is not going to be possible. For a simple example, consider the belief theory. It is plausible that the feeling of a churning stomach is painful regardless of whether that feeling is also an emotion. But the mere having of a belief is not normally either pleasurable or painful: beliefs normally have no hedonic character. So identifying emotions with beliefs does not in itself provide an answer to the question of how emotions can have essential hedonic character: something more needs to be said about what changes in *these* beliefs to make them pleasurable or painful. The same is true of other cognitivist theories, and of many perception theories. Perceptual construals, for instance, are normally hedonically neutral: it is neither pleasurable nor painful to construe the duck-rabbit as a duck or as a rabbit. If a construal of something as dangerous is painful, we need to know what changes.

One recourse for such theories would be to 'add on' the hedonic character by saying that a painful emotion consists in, say, a belief *and* a feeling of pain which the belief occasions. The belief and the hedonic state are logically separable, in the sense that either could in principle occur without the other. But if they did, what resulted would not count as an emotion. If the occurrent belief involved in, say, grief occurred without the feeling of pain, it would be something like 'serene contemplation of loss'. What would be the case if the pain occurred without the belief depends on what sort of pain the hybrid belief theorist chooses to 'add on', but on one possible model it would be a kind of intense aching whose object was, at most, the parts of the body in which the aching takes place. Such an aching, like the stomach churning of the Jamesian, would be the kind of state that one is motivated to avoid just in virtue of its painfulness.

A few of the 'impure' belief theories discussed in Chapter Two seem to be at least open to this model.[146] But most belief and perception theorists seem to deny that emotions necessarily include hedonic states that are something over and above the beliefs or perceptions with which they identify

---

[144] See esp. Prinz, *Gut Reactions*.

[145] A possible problem for the Jamesians: are the bodily feelings that are plausibly involved in emotion really as painful as we suppose emotions to be? Are they really painful at all, as opposed to just uncomfortable? Maybe not: see Roberts, *Emotions*, pp. 155-157. But the Jamesians are only really doing illustrative service here anyway, so I shall not explore these questions.

[146] e.g. those in Oakley, *Morality and the Emotions*, ch. 1 and Neu, *A Tear is an Intellectual Thing*.

emotion.[147] Such pains and pleasures do, they concede, frequently correlate with emotion. But they are not part of its essence. Instead, they argue, the painfulness of the emotion is *constituted by* the badness of its object. Martha Nussbaum, for example, argues as follows:

> There are pains that seem to be definable in purely physiological terms, or in purely nonintentional psychological terms. But is the pain we associate with grief among them? Aristotle's definitions of pain-linked emotion always speak of the 'pain at...', suggesting that he views pain itself as an intentional state with cognitive content. I believe this is correct, in such cases. We may have nonintentional pains in connection with grief, fear and pity. These would be dull aches and bodily feelings of nerves being painfully stimulated. But these seem like the 'boiling' and the 'trembling' - frequent correlates, but not necessary to the identity of the emotion.[148]

Similarly, Robert Roberts argues that:

> The pleasure or discomfort of a paradigm case emotion is the perceived satisfaction or frustration of a concern by a situation. [...] [For example,] a student has had a demanding semester, and in the tenth week he begins to desire intensely that it be over. The pleasure in this relief, when the semester is over, is that involved in perceiving the object of a desire as attained. On the other hand, he strongly desired, and even half expected, to get an 'A' is chemistry. When the report says 'B' he is disappointed. The displeasure in this emotion is that involved in perceiving the object of his desire as failing to be attained.[149]

If I understand this view rightly, the claim is that we misunderstand the nature of emotional pain by assimilating it to the kind of pain involved in stomach aches. We do not want, say, to be humiliated, and we wish to get out of humiliating circumstances when they arise. There is thus a sense in which humiliation is painful. But what motivates us here is our concern for our dignity and standing in the eyes of others, not some intrinsic feature of the mental state that provides an additional motivation over-and-above our desire to avoid indignity. Another example: the reason we do not wish to have grief is not because grief is painful in something like the way a stomachache is painful, but because we do not wish to lose the people we love. The key feature of this view is that painfulness is entirely constituted by features of the representational content of the emotion, and is not a distinct feature of the emotion over and above this.

Let us call this the distinction between the view that emotional pain is content-constituted and the view that it is content-distinct. Note that this is not the same as the distinction between intentional

---

[147] See e.g. Nussbaum, *Upheavals of Thought*, esp. pp. 53-64; Roberts, *Emotions*, 155-157; Solomon, 'Emotions and Choice', pp. 23-24. The belief-desire theory in Marks, 'A Theory of Emotion' is likely to fit here too, although this depends on what the correct account of desire is.

[148] Nussbaum, *Upheavals of Thought*, p. 64.

[149] Roberts, *Emotions,* p. 155.

and non-intentional pain, as the passage from Nussbaum might seem to suggest. The gut feelings of the Neo-Jamesian represent core relational themes in the external environment, so they are intentional, but their painfulness is content-distinct. Nor is being content-constituted the same as having a constitutive aim of accuracy. As we have noted, perceptual construals do not have a constitutive aim of accuracy, but their painfulness is still content-constituted. Some readers may wonder whether content-constituted pain is really pain properly so-called. Surely painful mental states are necessarily the sort of things that we are motivated to avoid because of what they are like to have? I have some sympathy with this idea, but I am not going to press it here: for the sake of argument, I accept that there is a sense in which content-constituted pain is pain. My interest is in whether it could exhaust the pain that we take to be essential to emotion.

Another source of pressure on this distinction arises from recent discussions on the metaphysics of pain. Some philosophers argue that all pain can be completely described as a perception or representation of something going wrong with its object.[150] Even a headache, on such a view, is a perception or representation of something wrong with one's brain. Would the truth of such a view mean that all pain is in fact content-constituted?

This is a fiddly point. There clearly is a sense in which a perceptualist or representationalist would hold that emotional pain must be content-constituted either way. Note however the specific definition I gave of content-constitution above: for emotional pain to be content-constituted in the sense that concerns us in this chapter is for it to be entirely constituted by features of the representational content *of the emotion*. A perceptualist about pain generally may hold that, for instance, the bodily feelings that constitute emotion according to Prinz are in fact constituted by perceptions that something is wrong with one's stomach, skin, heart and so on. But they certainly will *not* hold that they are constituted by perceptions of core relational themes. So in fact the distinction I am drawing is not collapsed on such views.

It is worth illustrating this. On the Nussbaumian view, the essential pain of grief is entirely constituted by the badness of loss: the reason I do not want to have grief is just that I do not want to have loss. On the Prinzian view, this is no doubt one cause of my not wanting grief, but there is another, which is that grief is constituted by bodily states that are themselves painful in a way that is not

[150] For classic perceptual theories of pain, see David Armstrong, *A Materialist Theory of the Mind* (New York: Humanities Press, 1968) and George Pitcher, 'Pain Perception', *Philosophical Review* 79:3 (1970): 368-393. For representational theories, see e.g. Gilbert Harman, 'The Intrinsic Quality of Experience', *Philosophical Perspectives*, 4: 31–52; Fred Dretske, *Naturalizing the Mind* (Cambridge, MA: MIT Press 1995); Michael Tye, 'Another Look at Representationalism about Pain' in Murat Aydede (ed.), *Pain: New Essays on Its Nature and the Methodology of Its Study* (Cambridge, MA: MIT Press 2013); David Bain, 'What Makes Pains Unpleasant?' *Philosophical Studies* (Supplement) 166: 69–89. For some interesting objections to both views, see Murat Aydede, 'Is the Experience of Pain Transparent? Introspecting Phenomenal Qualities', *Synthese* 196 (2019): 677–708.

constituted by the content of grief. Now, the perceptualists will say that this pain is itself constituted by a perception that something is wrong with my body, and so is itself content-constituted. It is beyond the scope of the present essay to determine whether that is true. But what is clear is that it is at least not constituted by the content *of the emotion:* nobody thinks that the content of grief is that something is wrong with one's heart or throat.

Representationalists and perceptualists about pain do not, therefore, need to reject the distinction I am drawing: the kind of content-constitution that they are concerned with is more fundamental than the kind that we are concerned with here, and the distinction I am drawing can be described in terms of a representationalist or perceptualist theory of pain. To avoid the prose becoming excessively tortured, I am going to bracket them in what follows: I shall continue to say that Nussbaum and Roberts think that emotional pain is entirely content-constituted, and that the alternative view is that emotion's painfulness is distinct from its content. Readers should simply bear in mind that there are different accounts of what this painfulness consists in at a more fundamental level.

4.2 The Oblivion Problem

We have seen, then, that some theorists limit the essential painfulness of emotion to content-constituted pain. I shall now look at a phenomenon that is difficult for such views to explain, namely what I shall call the pursuit of oblivion. People seek to escape from grief, shame, guilt, despair and other painful emotions, not only by altering the world so that the content of those emotions is no longer accurate, but by trying to distract themselves or to stop themselves from thinking altogether. The most extreme and melancholy case of this is suicide, offering complete and permanent oblivion. Alcohol and other narcotics can offer a temporary version. There are also far less extreme versions of this phenomenon. One might watch television to distract oneself from a stressful project at work. Or one might avoid talking about a friend's recent sacking or the recent death of their parents because one thinks that it would be painful for one's friend to be reminded of these things.[151]

The pursuit of oblivion presents no problem for theories of the Jamesian variety. The Jamesian and Neo-Jamesian will say: of course one seeks to avoid painful emotions *per se*: one avoids bodily pain *per se* in every other context, so naturally one will do so here too, where those bodily feelings are doing double-duty as emotions. Nor is it problematic to the hybrid theorists who add on some content-distinct painfulness to a cognitivist or a perceptual theory. But it *is* a problem for standard cognitivist and perceptual theories. For it seems to show that our motive for avoiding painful emotion is *not*

[151] 'Oblivion' is of course a strained word to use what I am aiming at in the latter sort of case, where I seek to avoid occasioning a particular mental state without stopping mental life in general. But because the case of seeking oblivion illustrates my point most strikingly, I shall name the problem for it.

exhausted by our motive for avoiding its object: we seem to be motivated to avoid painful emotions *per se* too.

So far as I am aware, this problem has not been seriously engaged with in recent literature. But I have found an interesting hint at a response in a throwaway remark in Fichte's *Addresses to the German Nation*. Fichte announces that he presupposes of his readers that they have:

> overcome the common inclination to deceive oneself over one's own affairs and to hold before oneself a less unpleasant picture of them than is consistent with the truth. This inclination is a cowardly flight from one's own thoughts, a childish disposition in which one seems to believe that if one cannot see one's misery, or at least does not admit to seeing it, then the misery is also annulled in reality, even as it is annulled in thought.[152]

With a little reconstruction, we might say that Fichte's model for understanding people's desire to avoid painful emotions is the behaviour of the child who covers her eyes and stops her ears, believing that she has thereby stopped external reality from existing. It might seem as though we are trying to avoid the mental state *per se*, but what we are really engaged in is a rationally defective attempt to change reality *by* changing our mental state, illicitly presupposing the quasi-solipsistic metaphysics of the infant ('infantile metaphysics', we might say). All we really cared about was, after all, the object of the emotion: hence there is no genuinely content-distinct pain, and no genuine counterexample to the Nussbaumian theory.

I think this response has some attractions. First, it is plausible that this sort of thinking does sometimes happen. I think it is plausible, for instance, that people often fail to go to the doctor despite having worrying symptoms because they do not want it to be *true* that they have a serious condition. Perhaps this sort of regression to infantile metaphysics is actually fairly common in adult life. Second, it is plausible that there is often something rationally problematic about what we are doing when we apparently try to avoid painful emotions *per se*. There is something attractive about the idea that one ought never to compromise on looking reality full-on because of the pain involved in doing so, that one's thinking about the world ought always to be governed solely by the nature of the world, not by whether it is pleasurable or painful to think about it in certain ways. This might seem to weigh in favour of identifying these emotions with the infantile thinking that Fichte proposes.

Still, the identification will not work. There are plausibly cases in adult life of refusing to believe something in order to stop it from being true, but it is not plausible that all of the cases I have mentioned of avoiding painful emotion can be described this way. There is no obvious reason to think that most people who drink themselves to oblivion or who die by their own hand are seeking to change the fact that occasions their misery by doing so: surely they are often simply trying to avoid emotional pain. This is all the more obvious when it comes to watching television to avoid thinking about one's stressful assignment. Nor is it plausible that all pursuit of oblivion is in any way rationally defective. To avoid discussion of painful topics seems to be a basic form of consideration for others, not a problematic

---

[152] Johann Gottlieb Fichte, *Reden an die deutsche Nation* (Cologne: Atlas-Verlag, n.d. [1808]), p. 9, translation mine.

form of abetting them in irrational behaviour. Nor, finally, is it plausible that, even when the pursuit of oblivion is irrational, it is normally irrational in the same way that regressions to 'infantile metaphysics' are. Surely we want to draw a distinction between someone who announces calmly intends to drink themselves into a stupor to temporarily escape their grief, and someone who is engaged in the adult equivalent of covering their eyes to stop the world from existing. Perhaps both are rationally problematic, but surely in very different ways. I conclude, therefore, that we are at least sometimes motivated to avoid painful emotions *per se*.

Let us return to our original question with the belief theory with this conclusion in mind. The mere fact that something is the feeling of a clenched stomach explains why it would be painful, but the mere fact that something is a belief does not. If emotions are beliefs, something is needed to explain why they are painful. The answer given explicitly by Nussbaum, and implicitly by others, is their content: they are painful because they are beliefs that something is the case that is harmful to our flourishing. But this gives us at most content-constituted painfulness, whereas, we have learnt, emotions are also painful in a content-distinct way: we want painful emotions to go away just because they hurt, not only because we do not like what they have to tell us. To explain this, we must either fundamentally revise our understanding of the painfulness of belief, moving away from the idea that belief's painfulness is content-constituted; or we must add a feeling of painfulness that is logically distinct from belief, arriving at our inelegant old friend, a hybrid theory; or we must move towards something like a bodily feelings theory, on which emotion is identified with something that is clearly painful in itself.

Two closing observations. First, the distribution of content-constituted and content-distinct emotional pains and pleasures is probably rather more complicated than the above discussion may have seemed to suggest. For instance, we sometimes speak of being 'painfully happy', 'so happy it is hard to bear', and so on. What seems to be happening here is that we have content-constituted pleasure - there is nothing about the situation we would like to change - but at least some content-distinct pain - our happiness about this completely good thing is actually painful in such a way that we have some motivation to moderate it. Perhaps the phenomenon of having 'a good cry' is sometimes an example of the inverse phenomenon, in which a content-constituted pain occasions an content-distinct pleasure. Second, as I have implied, there is plausibly an inverted version of the Oblivion Problem that applies to pleasurable emotion. This would obtain where people seek pleasurable emotions for their own sake, rather than, or in addition to, seeking what the pleasurable emotions are about. The nature of such cases is part of the subject of the next chapter.

4.3 Concluding Remarks

A better understanding of the nature of emotional pain and pleasure is valuable in itself. It also has some interesting further implications. First, it establishes that a successful theory of emotion must either identify emotion with a mental state that has content-distinct hedonic character, or it must be a hybrid

theory. Although I think Prinz's Neo-Jamesian account is unsuccessful for other reasons, its success in accommodating emotion's hedonic character in the first of these ways is surely admirable. Many 'feeling theories' of emotion have similar advantages.[153] The half-forgotten hybrid belief theories of writers like Neu and Oakley also deserve reassessment.

According to my own theory, a central and necessary element of emotion is a cognition. If I claimed that such cognitions are also sufficient for emotion, I would encounter difficulties similar to those of Nussbaum and Roberts, for cognitions do not obviously have any more hedonic character than beliefs do. But as I have made clear since Section 1.3, I am open to the possibility that emotions have elements apart from this. The argument of this chapter demonstrates that emotions not only have such elements, but that they have at least one necessary further element, namely a content-distinct hedonic character. Hybrid theories of emotion are not as elegant as unitary ones. But perhaps emotions just aren't very elegant phenomena.

[153] e.g. Greenspan's account in *Emotions and Reasons*, according to which emotions are constituted by feelings of comfort or discomfort.

Chapter Five

Sentimentality

A central theme of this essay has been the rational standards to which we hold emotion. In this chapter, I explore another dimension of this, looking at a way in which emotion can be defective that is largely and perhaps completely special to it, namely that it can be sentimental. The word 'sentimental' originally just meant 'emotional', and in philosophy 'sentimentalism' has remained a neutral technical term for a view that ascribes a central metaethical role to emotion. But in the sense that concerns us here, 'sentimental' emotion is emotion that is somehow insincere or fake.[154] The most famous words written on the subject probably remain Oscar Wilde's claim that 'the sentimentalist is one who desires to have the luxury of an emotion without paying for it'.[155] There is something plausible here, but of course it is a description of a mystery in terms of a metaphor: we do not literally pay for our emotions, so in what sense does the sentimentalist desire to avoid payment?

In this chapter, I seek to answer this question, building on existing attempts to offer a new account of sentimentality. In Section 5.1 I look at the best-known account of sentimentality in analytical philosophy, that of Michael Tanner. Tanner identifies several interesting features that sentimentality typically has, but his attempts to name some sufficient conditions for sentimentality yield an excessively broad account. In Section 5.2 I turn to the kind of account that most other philosophers writing on sentimentality have defended, according to which sentimentality involves misrepresenting the object because of a desire for the emotion itself. I argue that misrepresentation, though common, is not necessary to sentimentality. In Section 5.3 I develop my own account, according to which being caused in a certain way by the subject's desire is both necessary and sufficient for an emotion's being sentimental.

---

[154] Robert Solomon argues that 'sentimentality' need not have these normative problems in his *In Defence of Sentimentality* (Oxford: Oxford University Press, 2004). But he defines 'sentimentality' as 'nothing more nor less than the 'appeal to "tender feelings"' (p. 4) and characterises his opponents as those who believe that 'any emotion [...] is excessive, for an emotion or a sentiment as such is a disruption of the life of reason' (p. 10). This suggests he is using 'sentimentality' in something like the eighteenth-century sense, in which case the inconsistency with my claims here is merely apparent. Solomon does not seem to recognise that he is using the term differently to other writers, obscuring this fact. Cf. Michael Tanner, 'Review: In Defense of Sentimentality', *British Journal of Aesthetics* 46:3 (2006): 312-313 for discussion.

[155] Oscar Wilde, 'De Profundis' in *De Profundis and Other Writings,* ed. Coim Tolbin (London: Penguin, 2013 [1897]).

5.1. Tanner's Account

The *locus classicus* on sentimentality in analytical philosophy is Michael Tanner's paper 'Sentimentality'.[156] At the start of the essay, Tanner remarks that 'I have found [this] too perplexing and difficult a subject to be able to offer more than a series of rather loosely related thoughts'.[157] So it is perhaps unjust to treat his subsequent claims as a seriously proposed theory. Nevertheless, Tanner does go on to outline what appears to be an account of sentimentality, and it is at least interesting and useful to treat it as such in order to see what we can learn from it. Here is his summary of his view:

> It is sentimental to feel in a certain way towards some objects, on account of what those objects are; or to feel in certain ways irrespective of the objects concerned; or to feel in certain ways and not to act accordingly. Sentimentality, then, is the name of several kinds of disease of the feelings in which the elements of feeling 'in the void', of unfocused emotion, and of being prepared for huge bouts of emotional response to virtually random, or alternatively direly predictable, stimuli, are all closely connected.[158]

Tanner appears here to give a disjunctive account of sentimentality, according to which sentimentality comprises three distinct things. First, if certain emotions have certain objects, they are sentimental *ipso facto*. The example Tanner gives is that of extremely strong feelings about not stepping on snails.[159] Tanner's second condition, 'to feel in certain ways irrespective of the objects concerned', is ambiguous: it could mean that certain emotions are always sentimental, regardless of their objects, or it could mean that there are certain emotions that are sentimental if one feels them without paying attention properly to what their objects are like. Tanner discusses both possibilities elsewhere in the essay, so perhaps he meant to cover both here. I will come back to the possibility that feeling in certain ways just is sentimental in section 5.3: for the time being, I will focus on the second interpretation, which in any case seems to me more immediately plausible. The third is that certain emotions are sentimental if they are not accompanied with appropriate action: an example would be compassion for the needy unaccompanied by any attempt to help them.

Each of these things, I shall argue, are insufficient for sentimentality: there are cases of emotion that conform to each description without being sentimental. Consider first the idea that it is sentimental to feel certain emotions for certain objects. This is difficult to falsify decisively, because for it to be

---

[156] Michael Tanner, 'Sentimentality' in José Luis Bermúdez and Sebastian Gardner (eds), *Art and Morality* (London: Routledge, 2003 [1976-1977]). Joel Feinberg seems sympathetic to a similar view in his 'Sentiment and Sentimentality in Practical Ethics', *Proceedings and Addresses of the American Philosophical Association* 56:1 (1982): 19-46. Tanner draws on, but also cogently criticises, the discussion in I.A. Richards, *Principles of Literary Criticism* (London: Routledge & Kegan Paul, 1926 [1924]).

[157] Tanner, 'Sentimentality', p. 95.

[158] Ibid., pp. 105-106.

[159] Ibid., p. 105.

true, there need be only one case in which a given emotion just is sentimental if one has it for a given object - and I obviously cannot consider every case of this that might possibly be proposed. But I will cast some *prima facie* doubt on the proposal nonetheless. Tanner's example is that of grieving over a snail. To my mind, although this might be sentimental, it is not obvious that it *necessarily* is: such grief is no doubt excessive, but it is not clear that it must be fake or that the subject must in the relevant sense be unwilling to pay for it. Consider Hopkins's 'Spring and Fall':

> Margaret, are you grieving,
> Over Goldengrove unleaving?
> Leaves, like the things of man, you
> With your fresh thoughts care for, can you?
> Ah! as the heart grows older
> It will come to such sights colder
> By and by, nor spare a sigh
> Though worlds of wanwood leafmeal lie; [...][160]

If grieving for a snail is necessarily sentimental, then grieving for a leaf surely must be. But I submit that this is not how we feel about Hopkins's Margaret: her grief, though in some sense inappropriate, is perfectly sincere. Indeed, the very unwarrantedness of her grief seems to highlight its ingenuousness. It is surely right that grief for trivial objects tends to be sentimental, and any good account of sentimentality must explain that. But the mismatch alone is not enough. Perhaps there are other, better examples: as I say, Tanner only needs one. But he does not offer it, and I cannot think of it. So, cautiously, I conclude that Tanner's first condition is not sufficient.

The second condition was that it is sentimental to feel certain emotions without proper regard to the nature of their objects. Tanner claims that:

> the feelings it is worth having are those which it costs an effort to have - which doesn't mean that the effort is what makes them worth having, but is a consequence of the complexity of the data with which we have to deal. To have them it is necessary to understand what they are. To understand what they are it is necessary to run the risks involved in the depth-probing which results in their emergence and growth. The process is long, slow, time-consuming, painful and demanding of a degree of commitment that it is fearsomely difficult for most people to make. In this sense most people are cynics and sentimentalists.[161]

Again, this seems too strong. It is not quite clear from this passage (or, I think, in context) what is involved in the 'long, slow, time-consuming and demanding' process that is required for unsentimental emotion, but on a plausible interpretation, it involves carefully assessing the nature of the object of one's

[160] Gerald Manley Hopkins, *Major Works*, ed. Catherine Phillips (Oxford: Oxford University Press, 1987), p. 152. In its original context this passage has a slightly different significance, but we can set that to one side here.

[161] Tanner, 'Sentimentality', p. 103.

emotions in order to determine whether those emotions are constitutively rational. But there is no obvious sense in which, say, children go through such a process before having their emotions: on the whole, they seem to manage to be sincere without careful reflection. Surely this is also true of many adults. Though they may share the quality of being unexamined, there is surely a difference between naïve sincerity and sentimentality, one that we should not abandon lightly.

Tanner's third proposal is that it is sentimental to 'feel in certain ways and not to act accordingly'. Now there is surely something right here. The canonical example, due to William James, is the Russian lady who weeps over the characters in a play while her coachman freezes to death outside.[162] Once again, however, it is questionable whether the failure to act is sufficient for sentimentality. People fail to act on their emotions for all sorts of reasons: because they do not know how to put them into action, because they are shy or uncertain or depressed, or because they are for whatever reason genuinely unable to do anything. Not all of these constitute sentimentality. It is plausible that the emotions of the sentimentalist will frequently not be accompanied by the kinds of action that such emotions are normally accompanied by, but this seems to be *because* those emotions are sentimental rather than being what constitutes that sentimentality. A modified version of Tanner's claim might be that it is sufficient for an emotion's being sentimental if the subject does not have certain dispositions to action that accompany that emotion in its non-sentimental form (which is consistent with there being cases in which those dispositions are not realised and the emotion remains non-sentimental). But even this seems to get things the wrong way round. For it is surely right to explain the lack of the disposition in terms of the emotion's being sentimental; and if that is so, the emotion's being sentimental must be something prior to the lack of the disposition.

Each of Tanner's proposals has something plausible to it, then, and any successful account is going to need to explain why that is. But none is quite sufficient for sentimentality: there is something else going on.

5.2 Sentimentality and Misrepresentation

The most common kind of account posits two necessary conditions for sentimentality: it misrepresents its object in some way, and it is in some way inappropriately motivated. Anthony Savile proposes that:

> A sentimental mode of thought is typically one that idealizes its object under the guidance of a desire for gratification and reassurance. Derivatively, emotion is sentimental which is supported by such a thought.[163]

---

[162] William James, *Psychology* (London: Macmillan and Co, 1892), p. 148.

[163] Anthony Savile, *The Test of Time: An Essay in Philosophical Aesthetics* (Oxford: Oxford University Press, 1982), p. 241.

Similarly, Mark Jefferson claims that:

> Sentimentality involves attachment to a distorted series of beliefs [...] [it involves attributing] such things as the sweetness, dearness, littleness, blamelessness, and vulnerability [to] the emotions' objects. The qualities that sentimentality imposes on its objects are the qualities of innocence.[164]

Mary Midgley offers a slightly broader account:

> Being sentimental is misrepresenting the world in order to indulge our feelings.[165]

There are some differences between these accounts: Savile requires that sentimental emotions idealise their objects and Jefferson that their objects be represented as innocent, whereas Midgley's official definition seems to allow any kind of misrepresentation; Savile speaks of 'modes of thought' that 'support' emotions, suggesting a non-cognitive account of emotion itself, while Jefferson elsewhere commits himself to a belief-desire theory. But they all say that an emotion is sentimental just if it is, or is based on, a misrepresentation of the object (perhaps of a certain kind) and it is formed due to the wrong kind of motivation.

There is surely something plausible about this sort of account: sentimentality at least often involves misrepresenting the object, and it seems to have some kind of connection to self-indulgence or otherwise flawed motivations. These accounts also avoid the problems encountered by Tanner. Hopkins's Margaret does not have the wrong motive in forming her emotion, so she is not sentimental; the carefree souls who fail Tanner's 'fearsomely difficult process' test need not have had the wrong motives either, so they are not necessarily sentimental; those who fail to act in accordance with their emotions out of uncertainty or fear or depression are not sentimental provided that the emotion itself was sentimental. They also seem well placed to explain the normative problems surrounding sentimentality.

Are these writers correct, though, that sentimentality necessarily involves misrepresenting its object? It seems to me that they are not. We are surely familiar with sentimental compassion for malnourished children in faraway countries, sentimental outrage about mining companies displacing indigenous people, and sentimental admiration for Nelson Mandela or Mother Theresa. But none of these emotions seems to involve attributing an evaluative feature to its object that the object does not possess. Malnourished children *do* suffer innocently, the displacement of indigenous peoples *is* a terrible wrong, and Nelson Mandela and Mother Theresa *were* noble people. It is surely common that sentimental emotion misrepresents its object, but it is not necessary: this must be added to the list of frequent correlates of sentimentality that a successful theory should ideally explain.

---

[164] Mark Jefferson, 'What is Wrong With Sentimentality', *Mind* 92:368 (1983): 519-29 at pp. 526-527.

[165] Mary Midgley, 'Brutality and Sentimentality', *Philosophy* 54:209 (1979): 353-9 at p. 353. A similar view is supported in Scott Alexander Howard, 'Lyrical Emotions and Sentimentality', *Philosophical Quarterly* 62:248 (2012): 546-568.

The following response might be mounted: to be sure, sentimental emotion need not misrepresent any of the intrinsic qualities of the object. But it does not misrepresent *how the object matters to the subject*. The emotion implies that the object actually matters to the subject, but this is not so: in James Baldwin's words, 'the wet eyes of the sentimentalist betray his aversion to experience, his fear of life, his arid heart'.[166] Once again, there is something right here, but it is not quite correct as it stands. It might be unusual for people to feel sentimental emotions about objects that actually matter to them in the way that the emotion represents them as doing, but I see no reason to suppose it is impossible. Suppose on some occasion I make a great show of grief for my late grandmother because I like the idea of being so rich in feeling. Surely it might be the case that I actually am genuinely bereaved, and on other occasions feel sincere grief, but that on this occasion my grief is sentimental. Genuine concern for an object does not seem to endow it with complete immunity to also being the object of sentimentality.

5.3 A Motivation-Based Account

This leaves the second clause, the idea that the subject must be somehow wrongly motivated in forming the emotion. We have two proposals on the table already: Savile's suggestion that sentimental emotion involves being motivated by a 'desire for gratification and reassurance' and Midgley's claim that it involves aiming to 'indulge our feelings'. Another proposal is that of Roger Scruton:

> For the sentimentalist it is not the object but the subject of emotion that is important. Real love focuses on the other: it is gladdened by his pleasure and grieved by his pain. The unreal love of the sentimentalist focuses on the self, and treats the pleasures and pains of its object only as an excuse for playing the role that most appeals to it.[167]

What should we make of these? Savile's requirement that we seek 'reassurance' seems too narrow. The sentimentally outraged or sentimentally horrified do not seem to be seeking reassurance in any obvious sense, but they are still sentimental: so the search for reassurance is not necessary. As Savile notes, there are also people who form their emotions under the guidance of a desire for gratification and reassurance without being sentimental, like the student who convinces himself to be scornful of Kant in order to save himself embarrassment at not understanding him. Ruling out such cases is what forces Savile to introduce the clause that the sentimentalist misrepresents objects by idealising them, a clause that we have already shown to be unviable. So the search for reassurance is not sufficient for sentimentality either.

This leaves Scruton's and Midgley's proposals. Scruton's is clear: it says that an emotion is sentimental just if it is motivated by a desire for the emotion itself in virtue of the role that the emotion

[166] James Baldwin, *Notes of a Native Son* (Boston: Beacon Press, 2012 [1966]), p. 14.

[167] Roger Scruton, *Modern Culture* (New York: Continuum, 2005), p. 59.

plays in supporting a desired self-image of the subject. Midgley's 'indulging one's feelings' is vaguer, but one thing it might refer to is a desire for emotion in virtue of the content-distinct pleasurableness of emotion that we explored in Chapter Four. It seems to me that both of the types of emotion that Scruton and Midgley describes are cases of sentimentality: an emotion is sentimental if it is somehow motivated by the role that it plays in the subject's self-image, *or* if it is motivated by the content-distinct pleasurableness of the emotion. This leads me to suspect that specifying the character of the desire is actually unnecessary. I therefore propose the following positive account:

> An emotion is sentimental just if it is caused in a certain way by the subject's desire for the emotion.

Several points should be made about this account:

1. According to this account, we can distinguish between different kinds of sentimentality on the basis of why the subject wants to have the emotion.[168] Midgley and Scruton seem to highlight the two most important of these. It is surely possible for people to have a good cry at *Lassie Come Home* on Christmas Eve, without thinking that their crying shows what compassionate people they are (per Midgley and contra Scruton). As I argued in Chapter Four, emotions essentially have content-distinct hedonic properties, so it is perfectly intelligible that people should seek out those emotions for the sake of those properties. But content-distinct pleasure is not the only thing people seek from emotion. For example, it is not unusual for clever adolescents to regard themselves as so profound and highminded that everyday life must always fill them with anguish, disgust and existential horror. There might be content-distinct pleasures in these emotions, but are not necessary for one to be motivated to form them, and I suspect they do not typically play a great role: the main attraction, surely, is the idea of oneself as profound and highminded. These emotions are surely sentimental too (per Scruton and contra Midgley, at least on my interpretation).

Scruton's and Midgley's proposals seem to cover the standard reasons why people have sentimental emotions, but my account entails that other motives are at least technically possible. Suppose the eccentric billionaire offered to donate a fortune to charity in return for my feeling a sentimental emotion. My account entails - plausibly, I think - that an emotion I form as a result could count as sentimental, despite my selfless motive in forming it, provided that I formed it in the 'certain way' that I shall explain in point 3 below.[169]

2. A central attraction of this account is that it can explain the four correlates of sentimentality that we noticed in discussing other existing accounts, without overreaching by making them into sufficient or necessary conditions. The first was that there are certain emotions that would be hard (but not impossible) to have in non-sentimental form for certain objects, like grief for snails and leaves. This is simply because those objects do not actually matter enough in the relevant way to make the emotion

[168] A similar distinction is drawn by Savile elsewhere in *The Test of Time*, though it does not enter his definition.

[169] I am grateful to my external examiner, Alex Grzankowski, for pointing this out.

in question constitutively rational. It is not the case that only way that they could have them only if they were sentimentally motivated - it could be an innocent mistake, as with Margaret. But given the attractiveness of a compassionate self-image, it would not be surprising if that were the most common cause. Second, sentimental emotions are typically accompanied by a lack of careful thought about their natures and objects. This makes sense if one's motive is not genuine concern for the object but a desire for the emotion itself in virtue of its hedonic properties or its role in supporting one's self-image. Third, sentimental emotions are typically unaccompanied by the actions that accompany them in their unsentimental forms. I have said that sentimental emotions do not involve genuine concern for their objects: this being so, it is not surprising that they would usually be unaccompanied by the kind of actions that such concern motivated.[170] Finally, sentimental emotions often involve misrepresenting their objects. Again, this is not surprising if the subject is not actually motivated by concern for the object: although a sentimental emotion may represent its object accurately, that would be in a certain sense an accident.

3. Since I have characterised emotion as partly constituted by cognition, this characterisation of sentimental emotion makes it closely related to the widely discussed phenomenon of self-deception. The main rift in the literature on self-deception is between 'intentionalist' accounts, according to which self-deception involves intentionally forming a desired belief without regard to the evidence, and 'revisionist', 'motivationist' or 'non-intentionalist' accounts, according to which self-deception does not involve such an intention, but rather merely the formation of a false belief motivated by a desire for it.[171] The non-intentionalists argue *inter alia* that consciously intending to believe something false involves knowing it to be false, meaning that the intentionalist ascribes to the self-deceiver the impossible mental state of believing *p* and not-*p*. Intentionalists usually reply by positing some kind of psychological compartmentalisation, and arguing *inter alia* that the non-intentionalists cannot explain how we end up with some desired but unevidenced beliefs rather than others.

Corresponding positions could be developed in the case of sentimentality. These would involve different glosses on the 'cause in a certain way' formula of my definition, generating two versions of my account. The intentionalist argues that the sentimentalist intentionally develops sentimental emotions,

[170] To be clear, I do not say that it is impossible that one actually has such concern, only that it is not the motivation behind the emotion, e.g. someone who has genuine affection for their spouse may also sometimes put on a demonstrative show of false affection.

[171] For arguments for intentionalism, see e.g. Amelié Rorty, 'User-Friendly Self-Deception', *Philosophy*, 69 (1994): 211–228 and José Luis Bermúdez, 'Self-Deception, Intentions, and Contradictory Beliefs', *Analysis* 60:4 (2000): 309–319. For arguments for non-intentionalism, see Alfred Mele, *Self-Deception Unmasked* (Princeton: Princeton University Press, 2001); Mark Johnston, 'Self-Deception and the Nature of Mind' in Brian McLaughlin and Amelié Rorty (eds), *Perspectives on Self-Deception* (Berkeley: University of California Press, 1998); Annette Barnes, *Seeing Through Self-Deception* (New York: Cambridge University Press, 1997).

either forgetting or somehow compartmentalising the original intention to allow the emotion to be enjoyed. The non-intentionalist argues that the sentimentalist never intends to have the sentimental emotion, but is nonetheless influenced in the process of forming the emotion by their desire for it.

The debate between intentionalism and non-intentionalism is a fascinating and difficult one. It is no easier to resolve in the case of sentimental emotion than it is in the case of belief. So I am going to remain neutral here between these different possibilities: the debate between intentionalism and non-intentionalism about sentimentality is a new area that my characterisation of it opens up, and which must for now be left to the future. On either characterisation, however, it will be true that one cannot be knowingly sentimental.[172] To know that one's emotion is sentimental is also to know that one did not form it on the basis of the evidence: given the constitutive aim of emotion discussed in Section 1.1, this knowledge *ipso facto* ends the emotion. So whether or not sentimentality emotion involves intention, it certainly involves a kind of self-ignorance. In the case of more superficial kinds of sentimentality, however, that self-ignorance can clearly be correspondingly superficial: I might sit down for a session of *Lassie* with a pretty good idea that the object will not merit my emotion, and merely suspend that awareness for the duration of the film.

4. In Section 5.1 I issued a promissory note that I would say something about Tanner's claim that certain feelings are sentimental regardless of what objects they may have. Tanner's main reason for this is the existence of emotion in instrumental music, which, he seems to imply, has no objects that it can misrepresent, and which must therefore be sentimental in some other way.

> Tchaikovsky is incessantly berated for the sentimentality of his symphonies, without a clear suggestion that they are exaggerated responses to some programmatic situation or to 'Fate', the ostensible subject of the Fourth Symphony, to which over-reaction would be difficult. The idea is rather that the way he is expressing himself is itself thoroughly regrettable - that no kind of object, or lack of one, would justify this kind of response or demonstration.[173]

Tanner goes on to say of Franck's Piano Quintet that, while want to characterise it as sentimental:

> That still doesn't mean that it has even an implicit object; the feeling it expresses, and in suitably responsive listeners evokes, are comparable no doubt to the feelings one has when there are objects around - objects of one's feelings, I mean - but objects being regarded sentimentally.[174]

My response to this is (a) to remain neutral on Tanner's argument and (b) to cautiously accept Tanner's conclusion. First, Tanner's argument. The nature of musical expression is a matter of enormous

---

[172] Some formulations of non-intentionalism have left loopholes whereby knowing self-deception is possible, but this surely implies that those formulations are mistaken. See Richard Holton, 'What is the Role of the Self in Self-Deception?', *Proceedings of the Aristotelian Society* 101:1 (2001): 53-69.

[173] Tanner, 'Sentimentality', p. 97.

[174] Ibid., p. 102.

difficulty, into which I do not propose to enter here.[175] Tanner seems to presuppose that (a) pure instrumental music *can* express emotion, that (b) it can do this without implying anything about the object of that emotion, but that (c) each type of emotion has a certain subjective experiential character that is independent of its content, such that it can still be individuated and described in considerable detail even when the emotion occurs in its objectless form in music. Perhaps this is correct - my account is not inconsistent with it - but each of these assumptions is fiercely contested in the literature, and I take no view on their truth.[176]

Second, Tanner's conclusion, namely that it is sentimental 'to feel in a certain way irrespective of the objects concerned'. The meaning of this is not immediately clear, even in context, but the most plausible interpretation is that Tanner is denying that sentimental emotions must misrepresent their objects, and affirming that, at least in some cases, sentimentality must lie in some other feature of emotion. If so, then of course I agree: I deny that misrepresentation is necessary for sentimentality, and agree that the sentimentality of the emotion lies elsewhere. What is perhaps less satisfying about my account here is that it entails that when we hear music as sentimental, we in some sense hear in it something about the motivation of an implied subject. I admit that this is a strange idea. But then, the very fact that we can hear emotion in a series of pitched tones is strange. As I have said, the philosophy of musical expression is an area of great difficulty. I submit that we should be careful of making virtually *any* claim about musical emotion criterial of a successful theory of sentimentality until we have a better understanding of what that emotion consists in, a project that lies, of course, beyond the scope of the present enquiry.

5. We began this chapter by mentioning Wilde's remark that the sentimentalist 'is one who desires to have the luxury of an emotion without paying for it'. We may conclude by returning to it. There may be more than one truth behind this metaphor. One is that, in many (though not all) cases, the sentimentalist represents objects as mattering to them when they actually do not care about those objects in the relevant way. In such cases, the sentimentalist is able to claim the credit for caring about objects without having the dispositions to make sacrifices that caring about objects involves. Compassion for the faraway malnourished children, unaccompanied by any attempt to help them, is a straightforward example. Second: as we have seen, the sentimentalist is not motivated in the appropriate

---

[175] For some key discussions, see Scruton, *The Aesthetics of Music*, esp. chs. 3, 6, 7 and 8; Stephen Davies, *Musical Meaning and Expression* (Ithaca NY: Cornell University Press, 1994); Peter Kivy, *Sound Sentiment: An Essay on the Musical Emotions* (Philadelphia PA: Temple University Press, 1989); Nick Zangwill, 'Against Emotion: Hanslick was Right About Music', *British Journal of Aesthetics* 44:1 (2004): 29-43; Malcolm Budd, *Values of Art: Pictures, Poetry and Music* (London: Penguin, 1995) and Jerrold Levinson, 'Music and Negative Emotion' in his *Music, Art and Metaphysics* (Ithaca NY: Cornell University Press, 1990 [1982]).

[176] Cf. Jefferson, 'What is Wrong with Sentimentality?', pp. 521-522 and Joel Feinberg, 'Sentiment and Sentimentality in Practical Ethics', p. 26.

way to ensure that their emotions are constitutively rational. I take it that ensuring that one's emotions are constitutively rational involves some kind of effort: one has to review a range of facts, sometimes complex, about how the world is and what matters about it. This need not be Tanner's 'fearsomely difficult process', but it does involve some work. If that is right, then one way in which the sentimentalist will avoid paying for their emotions could be that they will not put in this effort: since they just want the emotion for its own sake or for the sake of the role that it plays in their self-image, they will skip the work of checking whether it actually gets the world right. It might, accidentally, do so. But that will just be the sentimentalist's epistemic good luck.

5.4 Concluding Remarks

The existing literature on sentimentality gets a great deal right: virtually every proposal captures some important but non-definitional feature of sentimental emotion. But most accounts have overreached, being either too broad, like Tanner's, or too narrow, like the standard account. I have argued that it is possible to capture everything that is plausible in these accounts while avoiding their pitfalls with a parsimonious account according to which it is necessary and sufficient for an emotion's being sentimental that it be caused in a certain way by the subject's desire for the emotion. A crucial question for future enquiry concerns how that 'certain way' should be glossed, and in particular whether we should give an intentionalist or non-intentionalist account of it.

Conclusions

Virtually everyone in the emotion literature agrees that emotions are representations. The question is what sort of representations they are. The dominant model today is that they are something like perceptual states. Normal perceptual states represent trees, lamps, desks and rainbows: these states represent how things matter to us. Other contemporary accounts do not identify emotion with perceptual experience, but share with them the denial that emotions are beliefs or other cognitions. In this essay we have discussed, at greater or lesser length, the accounts of Prinz, Roberts, Tappolet and Döring in the first category, while we have touched on the accounts of Greenspan, Goldie, Brady and Helm in the second. All of these accounts, I have argued, encounter a common problem, which is that they cannot deal with the way in which our emotions are subject to rational standards. There is nothing irrational about seeing a stick bend upon entering water, but there is something irrational about being angry at those one knows to be innocent or frightened of that which one knows to be harmless.

Philosophers have tried to explain away this irrationality, arguing that what makes recalcitrant emotions seem more irrational than perceptual illusions is that our recalcitrant emotions consume more of our attention, or that they motivate us to more problematic actions, or that we care about them more, or that we have more control over our emotional dispositions. These explanations do capture elements of emotion's potential irrationality. Recalcitrant anger consumes our attention, it motivates us to problematic actions, it divides us against ourselves, and we can often affect our dispositions to have it. But, I have argued, there is a hard core of the irrationality of recalcitrant emotion that none of these explanations can reach. This is what I have called its constitutive irrationality. The closest the perceptual theory can come to accommodating the fact that emotions are candidates for constitutive irrationality is by identifying emotion with a kind of perceptual experience that has its own constitutive aim and can thus be constitutively irrational. But even this, I argued, cannot plausibly accommodate the constitutive aim of emotion. It might be true that it is constitutively rational to experience a tree as a tree, but it is definitely not constitutively rational to experience a stick as straight when it enters water. Inaccurate emotion is always constitutively irrational, but, even on the most generous account, there are constitutively rational inaccurate perceptual experiences. It follows that the contemporary orthodoxy in philosophy of emotion must be abandoned.

The fact that emotions have constitutive aims led us to reassess the belief-based theories of emotion which were defended by figures like Nussbaum, Marks, Taylor, Lyons and Neu in the 1970s and 80s, but which have generally been regarded as discredited. On these views, emotions are constituted by beliefs that the object has a certain evaluative property, or by combinations of such beliefs and other mental states like feelings and desires. The belief theory is consistent with the fact that emotions have constitutive aims, and so it is, in this respect, closer to being an adequate account of emotion than are the theories that have displaced it. Closer investigation revealed, however, that although beliefs, like emotions, have constitutive aims, those aims are different. One can acquire a

rational belief that an object has a certain evaluative property on the basis of reliable testimony, but, in many cases, the emotion that the belief theory takes to correspond to such a belief could not be made rational by such testimony alone. What was missing, it emerges, is a 'grasp' on what grounds the evaluative property: I cannot admire someone unless I have some grasp on what there is about them to admire. I explored an attempt to accommodate this fact in terms of a sophisticated belief theory, according to which this 'grasp' was constituted by more beliefs. I argued that this ultimately proves unsatisfactory.

Emotions cannot be perceptual states, because they have a constitutive aim of accuracy. But they cannot be beliefs, because their constitutive aim goes beyond accuracy in a way that the aim of belief does not. The true view must thus characterise emotions being candidates for constitutive rationality, but as being something other than belief, a logical space I labelled 'non-doxastic cognitivism'. This left us in an interesting position. Philosophy of emotion has traditionally been concerned with trying to characterise emotion in terms of some other, putatively better understood, mental state, of which perceptual experience and belief are the most obvious examples. But there seem to be no other mental states of which non-doxastic cognitivism is *uncontroversially* true. One state, however, is *controversially* so classifiable, namely understanding: some philosophers think that subjective understanding is constituted by networks of beliefs, but others think that while it does have a constitutive aim that includes accuracy, it is irreducible to belief. If that is right, then emotion might be a kind of understanding, specifically something like 'concernful understanding'. But if not, then the project of trying to reduce emotion to another mental state may have to be abandoned. If so, then ultimately the best theory of emotion is an 'emotion theory of emotion': we can of course say a good deal more about emotion than that, but we cannot, on this view, ultimately break it down into some more fundamental mental state.

One reason for the abandonment of doxastic cognitivist theories in recent years was their supposed inconsistency to accommodate the emotions of animals. It is claimed that animals cannot have beliefs, but that they do have emotions, and that their emotions must therefore constitute a counterexample to any belief theory. I argued, firstly, that there are ways for a cognitivist theory, doxastic or otherwise, to accommodate the presence of emotion in animals through claiming that similar evolutionary functions are performed using different mechanisms. Provided it is open to the possibility that human emotions are partly constituted by non-cognitive mental states, as both my account and most of the classical belief theories are, it can also explain the continuities between human and animal emotion in terms of the common presence of those states. Furthermore, a growing scientific and philosophical consensus has emerged that animals can have beliefs after all, vitiating the premiss of the objection. The second well known family of arguments against cognitivist theories focusses on how they describe 'recalcitrant' emotions that do not respond to reasons. I argued that there is no clear reason to suppose that emotions exhibit a strikingly higher level of recalcitrance than beliefs do, and

that a non-doxastic cognitivist theory could give a plausible characterisation of the level of irrationality that obtains when recalcitrant emotions do arise.

In Chapter One I explicitly left open the possibility that emotions have necessary non-cognitive components, and both in Chapter One and in Chapter Three I highlighted some of the ways that allowing other components into our account of emotion can strengthen it. In Chapter Four I finally offered a direct argument for the conclusion that emotions have necessary non-cognitive elements. This was what I called the 'oblivion argument'. Many perceptual and belief theories have sought to subsume the pleasure and pain involved in the emotion into its content, arguing, for instance, that the whole reason we do not like grieving is that we do not like loss. The problem with this is that it struggles to explain why people sometimes seek to pursue or avoid the mental states themselves, while knowing that to do so cannot affect their objects: the pursuit of oblivion is a clear example of this. I explored some possible responses, but concluded that they were unsuccessful, and that the pleasure or painfulness of emotion must be a further 'content-distinct' mental state that is in principle distinguishable from our cognitive states.

The conclusion of Chapter Four thus qualified the cognitivist theory of emotion that I had developed in the earlier chapters of the thesis. In Chapter Five, by contrast, I focussed once again on the norms governing emotion, examining the nature of sentimentality. Most accounts of sentimentality give two conditions for it, that it be in some way wrongly motivated and that it misrepresent the object in some way. I argued that the latter condition is unnecessary: sentimental emotions frequently misrepresent their objects, but it is not impossible to have sentimental emotions that represent their objects as mattering in a way they really do matter. What makes sentimental emotion sentimental is the way it arose: it is in some sense formed under the guidance of a desire for the emotion itself. Drawing on the literature on self-deception, I distinguished two ways in which we could understand this: it might be that the sentimental emotion is formed intentionally, the original intention being subsequently forgotten or compartmentalised, or that its formation was caused by the desire without there being any explicit intention.

* * * * *

At several points we encountered questions about emotion whose resolution depended upon the resolution of debates elsewhere in philosophy. Whether emotion should be characterised as a form of understanding, or whether it should be characterised as a basic mental state, cannot be resolved until we determine more about the nature of understanding. How cognitivist theories should describe the emotions of infants will remain unclear until we have a better developed account of the nature of infant minds at every stage of their development. Whether sentimentality should be understood as intentional or as merely motivated is likely to be solved together with the cognate debate about self-deception. I have also left largely untouched a number of other crucial issues about emotion. I have said almost

nothing about the kind of content that emotion must have: if we must care about the object of emotion, then what does this 'caring' consist in? Apart from its hedonic character, I have said little about any features that emotion may have apart from its constitutively rational one. Does emotion have components beyond cognition, pain and pleasure? Nor have I made any attempt to distinguish between the content different emotions. What is the difference between the thoughts involved in envy and jealousy, or liking and love? These are all deeply interesting questions. But for now, they must be left for the future.

Bibliography

Allen, Colin, and Marc Bekoff. *Species of Mind.* Cambridge, MA: MIT Press, 1997.

Armon-Jones, Claire. *Varieties of Affect.* New York: Harvester Wheatsheaf, 1991.

Baldwin, James. *Notes of a Native Son.* Boston: Beacon Press, 2012 [1966].

Barnes, Annette. *Seeing Through Self-Deception.* New York: Cambridge University Press, 1997.

Benbaji, Hagit. 'How is Recalcitrant Emotion Possible?' *Australasian Journal of Philosophy* 91:3 (2013): 577-599.

Bermúdez, José Luis. 'Self-Deception, Intentions, and Contradictory Beliefs'. *Analysis* 60, no. 4 (2000): 309–19.

———. *Thinking Without Words.* Oxford: Oxford University Press, 2003.

Boghossian, Peter. 'The Normativity of Content. *Philosophical Issues* 13:1 (2003), 31–45

Brady, Michael S. *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford: Oxford University Press, 2014.

———. 'Recalcitrant Emotions and Visual Illusions'. *American Philosophical Quarterly* 44, no. 3 (2007): 273–84.

———. 'The Irrationality of Recalcitrant Emotions'. *Philosophical Studies* 145, no. 3 (2009): 413–30.

Brentano, Franz. *Vom Ursprung sittlicher Erkenntnis.* Leipzig: Duncker & Humblot, 1889.

———. *Psychologie vom empirischen Standpunkt.* Leipzig: Duncker & Humblot, 1874.

Budd, Malcolm. *Values of Art: Pictures, Poetry and Music.* London: Penguin, 1995.

———. 'The Acquaintance Principle'. *British Journal of Aesthetics* 43, no. 4 (2003): 386–92.

Byrne, Alex. 'Introspection'. *Philosophical Topics* 33, no. 1 (2005): 79–104.

Calhoun, Cheshire. 'Cognitive Emotions?' In *What is an Emotion? Classic and Contemporary Philosophers on the Emotions*, edited by Robert C. Solomon. Oxford: Oxford University Press, 1983.

Carruthers, Peter. 'On Being Simple Minded'. *American Philosophical Quarterly* 41, no. 3 (2004): 205–20.

———. 'Meta-Cognition in Animals: A Skeptical Look'. *Mind and* Language 23 (2008): 58–89.

———. 'Invertebrate Concepts Confront the Generality Constraint (and Win)'. In *Philosophy of Animal Minds*, edited by Robert W. Lurz, 89–107. Cambridge: Cambridge University Press, 2009.

Churchland, Paul M. *Scientific Realism and the Plasticity of Mind*. Cambridge: Cambridge University Press, 1979.

Crane, Tim. 'Is Perception a Propositional Attitude?' *Philosophical Quarterly* 59, no. 236 (2009): 452–69.

Crisp, Roger. 'Review of Joel Kupperman's *Value... And What Follows*'. *Philosophy* 75, no. 3 (2000): 452–62.

———. 'Compassion and Beyond'. *Ethical Theory and Moral Practice* 11, no. 3 (2008): 233–46.

D'Arms, Justin, and Daniel Jacobson. 'The Moralistic Fallacy: On the "Appropriateness" of Emotions'. *Philosophy and Phenomenological Research* 61, no. 1 (2000): 65–90.

Damasio, Antonio. *Descartes' Error: Reason, Emotion and the Human Brain.* New York: Putnam, 1994.

Davidson, Donald. *Essays on Actions and Events.* Oxford: Clarendon, 1980.

———. 'Thought and Talk'. In *Inquiries into Truth and Interpretation.* Oxford: Clarendon, 1984.

———. 'Rational Animals'. In *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, edited by Ernest Lepore and Brian McLaughlin, 473–81. New York: Basil Blackwell, 1985.

———. 'The Emergence of Thought'. *Erkenntnis* 51 (1997): 7–17.

Davies, Stephen. *Musical Meaning and Expression.* Ithaca, NY: Cornell University Press, 1994.

De Sousa, Ronald. *The Rationality of Emotion.* Cambridge, MA: MIT Press, 1987.

Deigh, John. 'Cognitivism in the Theory of Emotions'. *Ethics* 104, no. 4 (1994): 824–54.

———. 'Primitive Emotions'. In *Thinking About Feeling: Contemporary Philosophers on Emotion.* Edited by Robert C. Solomon. Oxford: Oxford University Press, 2004.

Dennett, Daniel. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

Deonna, Julien A. 'Emotion, Perception and Perspective'. *Dialectica* 60, no. 1 (2006): 24–49.

Deonna, Julien A., and Fabrice Teroni. *The Emotions: A Philosophical Introduction.* London: Routledge, 2012.

Döring, Sabine A. 'Explaining Action by Emotion'. *The Philosophical Quarterly* 53, no. 211 (2003): 214–30.

———. 'Why Recalcitrant Emotions are Not Irrational'. In *Emotion and Value*, edited by Sabine Roeser and Cain Todd. Oxford: Oxford University Press, 2014.

———. 'What's Wrong with Recalcitrant Emotions? From Irrationality to Challenge to Agential Identity'. *Dialectica* 69:3 (2015): 381-402.

Dretske, Fred. *Knowledge and the Flow of Information.* Cambridge, MA: MIT Press, 1982.

———. 'Misrepresentation'. In *Belief: Form, Content and Function*, edited by Radu Bogdan, 17–36. Oxford: Oxford University Press, 1986.

Ekman, Paul. 'Are There Basic Emotions?' *Psychological Review* 99, no. 3 (1992): 550–3.

———. 'Basic Emotions'. In *Handbook of Cognition and Emotion*, edited by Tim Dalgleish and Mike Power, 45–60. New York: Wiley, 1999.

Elgin, Catherine. 'True Enough'. *Philosophical Issues* 14 (2004): 113–31.

Ewing, A. C. *The Definition of Good.* London: Routledge Kegan and Paul, 1948.

Finn, J.K., T. Tregenza and N. Tregenza. 'Defensive tool use in a coconut-carrying octopus'. *Current Biology* 19:23 (2009): 1069-1070.

Fiorito, Graziano and Pietro Scotto. 'Observational Learning in *Octopus vulgaris*'. *Science* 256:5056 (1992): 545–547.

Fodor, Jerry A. *Psychosemantics*. Cambridge, MA: MIT Press, 1987.

———. *The Elm and the Expert*. Cambridge, MA: MIT Press, 1994.

Frankfurt, Harry. 'On Bullshit'. *Raritan Quarterly Review* 6, no. 2 (1986): 81–100.

Frijda, Nico. 'Varieties of Affect: Emotions and Episodes, Moods and Sentiments'. In *The Nature of Emotion: Fundamental Questions*, edited by Paul Ekman and Richard J. Davison, 197–202. New York: Oxford University Press, 1994.

Gaus, Gerald. *Value and Justification: The Foundations of Liberal Theory.* Cambridge: Cambridge University Press, 1990.

Gendler, Tamas Szabó, and Karson Kovakovich. 'Genuine Rational Fictional Emotions'. In *Contemporary Debates in Aesthetics and the Philosophy of Art*, edited by Matthew Kieran. Oxford: Blackwell, 2006.

Gibbard, Allen. 'Truth and Correct Belief'. *Philosophical Issues* 15, no. 1 (2005): 338–50.

Glock, Hans-Johann. 'Animal Rationality and Belief'. In *The Routledge Handbook of Philosophy of Animal Minds*, edited by Kristin Andrews and Jacob Beck, 89–99. London: Routledge, 2017.

Goldie, Peter. *The Emotions: A Philosophical Exploration.* Oxford: Oxford University Press, 2000.

———. 'Emotion, Feeling, and Knowledge of the World'. In *Contemporary Philosophers on Emotion*, edited by Robert Solomon. Oxford: Oxford University Press, 2004.

———. 'Getting Feelings into Emotional Experience in the Right Way'. *Emotion Review* 1, no. 3 (2009): 232–9.

———. 'Review: Gut Reactions: A Perceptual Theory of Emotion'. *Mind* 115, no. 458 (2006): 453–7.

Gordon, Robert. *The Structure of Emotions: Investigations in Cognitive Psychology.* Cambridge: Cambridge University Press, 1987.

Greenspan, Patricia. *Emotions and Reasons.* London: Routledge, 1988.

Griffin, Donald R. *Animal Thinking.* Cambridge, MA: Harvard University Press, 1984.

———. *Animal Minds: Beyond Cognition to Consciousness*. Chicago: University of Chicago Press, 2001.

Griffiths, Paul E. 'The Degeneration of the Cognitive Theory of Emotions'. *Philosophical Psychology* 2, no. 3 (1989): 297–313.

———. 'Review: Gut Reactions: A Perceptual Theory of Emotion'. *British Journal of the Philosophy of Science* 59, no. 3 (2008): 559–67.

Griffiths, Paul and Andrea Scarantino, 'Emotions in the Wild: The Situated Perspective on Emotion'. In Philip Robbins and Murat Aydede (eds.), *The Cambridge Handbook of Situated Cognition.* Cambridge: Cambridge University Press, 2009.

Grimm, Peter. 'Is Understanding a Species of Knowledge?' *British Journal for the Philosophy of Science* 57, no. 3 (2006): 515–35.

Grzankowski, Alex. 'The Real Trouble with Recalcitrant Emotions'. *Erkenntnis.* 82:3 (2017): 741-651.

Hanson, N.R. *Patterns of Discovery* Cambridge: Cambridge University Press, 1958.

Hawthorne, John, and Tamar Szabo Gendler, eds. *Perceptual Experience.* Oxford: Oxford University Press, 2006.

Helm, Bennett W. *Emotional Reason: Deliberation, Motivation and the Nature of Value.* Cambridge: Cambridge University Press, 2001.

———. 'Emotions as Evaluative Feelings'. *Emotion Review* 1, no. 3 (2009): 248–55.

———. 'Love, Identification and the Emotions'. *American Philosophical Quarterly* 46, no. 1 (2009): 39–59.

———. 'Emotions and Recalcitrance: Reevaluating the Perceptual Model'. *Dialectica* 69, no. 3 (2015): 417–33.

Hills, Alison. 'Moral Testimony and Moral Epistemology'. *Ethics* 120, no. 1 (2009): 94–127.

———. 'Understanding Why'. *Noûs* 50, no. 4 (2016): 661–88.

Holton, Richard. 'What is the Role of the Self in Self-Deception?' *Proceedings of the Aristotelian Society* 101, no. 1 (2001): 53–69.

Hopkins, Robert. 'How to be a Pessimist About Aesthetic Testimony'. *The Journal of Philosophy* 108:3 (2011): 138-157.

Hopkins, Gerard Manley. *Major Works*, edited by Catherine Phillips. Oxford: Oxford University Press, 1987.

Howard, Scott Alexander. 'Lyrical Emotions and Sentimentality'. *Philosophical Quarterly* 62, no. 248 (2012): 546–68.

James, William. 'What is an Emotion?' *Mind* 9, no. 34 (1884): 199–205.

———. *Psychology*. London: Macmillan and Co, 1892.

Jefferson, Mark. 'What is Wrong With Sentimentality'. *Mind* 92, no. 368 (1983): 519–29.

Johnston, Mark. 'Self-Deception and the Nature of Mind'. In *Perspectives on Self-Deception*, edited by Brian McLaughlin and Amelié Rorty, 63–91. Berkeley: University of California Press, 1998.

———. 'The Authority of Affect'. *Philosophy and Phenomenological Research* 63, no. 1 (2001): 181–214.

Joyce, Richard. 'Rational Fear of Monsters'. *British Journal of Aesthetics* 40, no. 2 (2000): 209–24.

Kahn, Charles H. 'Plato's Theory of Desire'. *The Review of Metaphysics* 41, no. 1 (1987): 77–103.

Kenny, Anthony. *Action, Emotion and Will*, 2nd edn. London: Routledge, 2003 [1963].

Kieswetter, Benjamin. *The Normativity of Rationality.* Oxford: Oxford University Press, 2017.

Kivy, Peter. *Sound Sentiment: An Essay on the Musical Emotions*. Philadelphia, PA: Temple University Press, 1989.

Kvanvig, Jonathan. *The Value of Knowledge and the Pursuit of Understanding.* Cambridge: Cambridge University Press, 2003.

Laetz, Brian. 'A Modest Defense of Aesthetic Testimony', *Journal of Aesthetics and Art Criticism* 66:4 (2008): 355-363.

Lamarque, Peter. 'How Can We Fear and Pity Fictions?' *British Journal of Aesthetics* 21 (2000): 291–304.

Lange, Carl. 'The Mechanism of Emotions'. In *The Classical Psychologists*, edited by Rand Benjamin, 672–84. Boston: Houghton Mifflin, 1912 [1885].

Langton, Rae. 'Objective and Unconditioned Value'. *Philosophical Review* 116, no. 2 (2007): 157–85.

Lazarus, Richard S. 'The Stable and Unstable in Emotion'. In *The Fundamental Nature of Emotion*, edited by Paul Ekman and Richard J. Davidson, 79–85. Oxford: Oxford University Press, 1994.

Lazarus, Richard S., and Bernice N. Lazarus. *Passion and Reason: Making Sense of Our Emotions.* Oxford: Oxford University Press, 1994.

Leavis, F.R. '"Thought" and Emotional Quality', *Scrutiny* 8:1 (1945): 53-71.

———. 'Reality and Sincerity'. *Scrutiny* 19:2 (1952): 90-8.

Levinson, Jerrold. 'Music and Negative Emotion'. In *Music, Art and Metaphysics.* Ithaca NY: Cornell University Press, 1990 [1982].

Linehan, Elisabeth A. 'Ignorance, Self-Deception and Moral Responsibility'. *Journal of Value Inquiry* 16 (1982): 101–15.

Lipton, Peter. 'Understanding Without Explanation'. In *Scientific Understanding: Philosophical Perspectives*, edited by Henk de Regt, Sabina Leonelli, and Kai Eigner, 43–63. Pittsburgh: University of Pittsburgh Press, 2009.

Livingstone, Paisley. 'On an Apparent Truism in Aesthetics'. *British Journal of Aesthetics* 43, no. 3 (2003): 260–78.

Lormand, Eric. 'Towards a Theory of Moods'. *Philosophical Studies* 47 (1985): 385–407.

Lurz, Robert W. 'In Defense of Wordless Thoughts about Thoughts'. *Mind and Language* 22 (2007): 270–96.

———. 'The Philosophy of Animal Minds: An Introduction'. In *Philosophy of Animal Minds*, edited by Robert W. Lurz, 1–14. Cambridge: Cambridge University Press, 2009.

Marks, Joel. 'A Theory of Emotion'. *Philosophical Studies* 42, no. 1 (1982): 227–42.

Martin, Michael. 'The Transparency of Experience'. *Mind and Language* 17 (2002): 376–425.

Matravers, Derek. 'The Challenge of Irrationalism, and How Not to Meet It'. In *Contemporary Debates in Aesthetics and the Philosophy of Art*, edited by Matthew Kieran, 254–64. Oxford: Blackwell, 2006.

McDowell, John. *Mind and World.* Cambridge, MA: Harvard University Press, 1994.

Mele, Alfred. *Self-Deception Unmasked*. Princeton: Princeton University Press, 2001.

Meskin, Aaron. 'Aesthetic Testimony: Can We Learn from Others about Beauty in Art?' *Philosophy and Phenomenological Research* 69, no. 1 (2004): 65–91.

———. 'Solving the Puzzle of Aesthetic Testimony'. In *Knowing Art: Essays in Aesthetics and Epistemology*, ed. Dominic Mclver Lopes and Matthew Kieran. Dordrecht: Springer, 2006.

Midgley, Mary. 'Brutality and Sentimentality'. *Philosophy* 54, no. 209 (1979): 353–9.

Millikan, Ruth. *Language, Thought and Other Biological Categories.* Cambridge, MA: MIT Press, 1984.

Moran, Richard. 'Frankfurt on Identification: Ambiguities of Identity in Mental Life'. In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Harry Overton, 189–217. Cambridge, MA: MIT Press, 2002.

Morillo, Carolyn. 'The Reward Event and Motivation'. *Journal of Philosophy* 87 (1990): 169–86.

Nelkin, Dana Kay. 'Responsibility and Self-Deception: A Framework'. *Humana.Mente Journal of Philosophical Studies* 20 (2012): 117–39.

Neu, Jerome. *A Tear is an Intellectual Thing.* Oxford: Oxford University Press, 2002.

Nussbaum, Martha. *The Therapy of Desire: Theory and Practice in Hellenistic Ethics.* Oxford: Oxford University Press, 1990.

———. *Upheavals of Thought: The Intelligence of Emotions.* Cambridge: Cambridge University Press, 2001.

Oakley, Justin. *Morality and the Emotions*. London: Routledge, 1992.

Packard, A. 'Cephalopods and fish: The limits of convergence'. *Biological Reviews* 47:2 (1972): 241–307.

Price, Carolyn. 'Affect Without Object: Moods and Objectless Emotions'. *European Journal of Analytical Philosophy* 2, no. 1 (2006): 49–68.

Prichard, Duncan. 'Knowing the Answer, Understanding and Epistemic Value'. *Grazer Philosophische Studien* 77, no. 1 (2008): 325–39.

Prinz, Jesse. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press, 2004.

———. 'Embodied Emotions'. In *Thinking About Feeling: Contemporary Philosophers on Emotion*, edited by Robert Solomon. Oxford: Oxford University Press, 2004.

———. 'Are Emotions Feelings?' *Journal of Consciousness Studies* 12, no. 8–10 (2005): 9–25.

———. 'Is Emotion a Form of Perception?', *Canadian Journal of Philosophy* 36 (2006): 137–60.

Pugmire, David. *Rediscovering Emotion.* Edinburgh: Edinburgh University Press, 1999.

Radford, Colin. 'How Can We be Moved by the Fate of Anna Karenina?' *Proceedings of the Aristotelian Society* 49 (1975): 67–80.

Rawls, John. *A Theory of Justice.* Cambridge, MA: Harvard University Press, 1971.

Riaz, Amber. 'Moral Understanding and Knowledge'. *Philosophical Studies* 172, no. 1 (2015): 113–28.

Richards, I.A. *Principles of Literary Criticism.* London: Routledge & Kegan Paul, 1926 [1924].

Roberts, Robert C. 'What an Emotion is: A Sketch'. *Philosophical Review* 97, no. 2 (1988): 183–209.

———. *Emotions: An Essay in Aid of Moral Psychology.* Cambridge: Cambridge University Press, 2003.

———. *Emotions in the Moral Life.* Cambridge: Cambridge University Press, 2013.

———. 'Propositions and Animal Emotions'. *Philosophy* 71, no. 275 (1996): 147–56.

———. 'The Sophistication of Non-Human Emotion'. In *The Philosophy of Animal Minds*, edited by Robert W. Lurz, 218–36. Cambridge: Cambridge University Press, 2009.

Robinson, Jenefer. 'Startle'. *The Journal of Philosophy* 92, no. 2 (1995): 53–74.

Rorty, Amelié. 'User-Friendly Self-Deception'. *Philosophy* 69 (1994): 211–28.

Ruskin, John. *Modern Painters.* London: Smith, Elder and Co., 1843-1860. 5 vols.

Sartre, Jean-Paul. *L'être et le néant: Essai d'ontologie phénoménologique.* Paris: Gallimard, 2001 [1943].

Savile, Anthony. *The Test of Time: An Essay in Philosophical Aesthetics*. Oxford: Oxford University Press, 1982.

Scanlon, T. M. *What We Owe to Each Other.* Cambridge, MA: Harvard University Press, 1998.

Schueler, George F. *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge, MA: MIT Press, 1995.

Scruton, Roger. *Art and Imagination: A Study in the Philosophy of Mind.* London: Methuen, 1976.

———. 'Fantasy, Imagination and the Screen'. *Grazer Philosophische Studien* 19:1 (1983): 35-46.

———. *The Aesthetics of Music.* Oxford: Oxford University Press, 1999.

———. *Modern Culture.* New York: Continuum, 2005.

———. *The Aesthetics of Architecture*, 2nd edn. Princeton: Princeton University Press, 2013 [1979].

Searle, John. *Intentionality: An Essay in the Philosophy of Mind.* Cambridge: Cambridge University Press, 1983.

———. *The Rediscovery of Mind*. Cambridge, MA: MIT Press, 1992.

———. 'Animal Minds'. *Midwest Studies in Philosophy* 19, no. 1 (1994): 206–19.

Sellars, Wilfrid. 'Intentionality and the Mental'. In *Minnesota Studies in the Philosophy of Science*, Volume 2, edited by Herbert Feigl, Michael Scriven, and Grover Maxwell, 507–33. Minneapolis, MN: University of Minnesota Press, 1957.

Shah, Nishi. 'How Truth Governs Belief', *Philosophical Review* 112:4 (2003).

Shah, Nishi, and David Velleman. 'Doxastic Deliberation'. *Philosophical Review* 114, no. 4 (2005): 497–534.

Shargel, Daniel and Jesse Prinz. 'An Enactivist Theory of Emotional Content'. In Hichem Naar and Fabrice Teroni, *The Ontology of Emotions*. Cambridge: Cambridge University Press, 2018.

Sizer, Laura. 'Towards a Computational Theory of Mood'. *British Journal for the Philosophy of Science* 51 (2000): 743–69.

Sliwa, Paulina. 'Understanding and Knowing'. *Proceedings of the Aristotelian Society* 65, no. 1 (2015): 57–74.

———. 'Moral Understanding as Knowing Right from Wrong'. *Ethics* 127, no. 3 (2017): 521–52.

Snowdon, Paul. 'Perception, Vision and Causation'. *Proceedings of the Aristotelian Society* 81 (1980–81): 175–92.

Solomon, Robert C. *The Passions.* New York: Doubleday, 1976.

———. 'On Emotions as Judgments'. *American Philosophical Quarterly* 25, no. 2 (1988): 183–91.

———. 'Emotions, Thoughts, and Feelings: Emotions as Engagements with the World'. In *Thinking About Feeling: Contemporary Philosophers on Emotions*, edited by Robert C. Solomon, 76–90. Oxford: Oxford University Press, 2001.

———. *In Defence of Sentimentality.* Oxford: Oxford University Press, 2004.

Sorabji, Richard. *Animal Minds and Human Morals: The Origins of the Western Debate.* Ithaca, NY: Cornell University Press, 1995.

Steglich-Petersen, Asbjørn. 'No Norm Needed: On the Aim of Belief'. *Philosophical Quarterly* 56, no. 225 (2005): 499–516.

Stephan, Achim. 'Moods in Layers'. *Philosophia* 45, no. 4 (2017): 1481–95.

Stich, Stephen. 'Animals Have Beliefs?' *Australasian Journal of Philosophy* 57 (1979): 15–28.

Strawson, Peter. 'Perception and its Objects'. In *Perception and Identity: Essays Presented to A. J. Ayer*, edited by G. F. McDonald, 41–60. London: Macmillan, 1979.

Strevens, Michael. 'No Understanding Without Explanation'. *Studies in History and Philosophy of Science, Part A* 44, no. 3 (2013): 510–15.

Tanner, Michael. 'Sentimentality'. In *Art and Morality*, edited by José Luis Bermúdez and Sebastian Gardner, 95–110. London: Routledge, 2003 [1976–77].

———. 'Ethics and Aesthetics Are - ?' In José Luis Bermúdez and Sebastian Gardner (eds), *Art and Morality.* New York: Routledge, 2003.

———. 'Review: In Defense of Sentimentality'. *British Journal of Aesthetics* 46, no. 3 (2006): 312–13.

Tappolet, Christine. 'Emotions, Perceptions and Emotional Illusions'. In *Perceptual Illusions: Philosophical and Psychological Essays*, edited by Clotilde Calabi, 205–22. Basingstoke: Palgrave Macmillan, 2012.

———. *Emotions, Values and Agency.* Oxford: Oxford University Press, 2016.

Taylor, Gabriele. 'Justifying the Emotions'. *Mind* 84, no. 1 (1975): 390–402.

Tye, Michael. 'The Problem of Simple Minds: Is There Anything it is Like to be a Honey Bee?' *Philosophical Studies* 88 (1997): 289–317.

Velleman, David. 'On the Aim of Belief. In David Velleman (ed.), *The Possibility of Practical Reason*. Oxford: Oxford University Press, 2000.

Walton, Kendall L. 'Fearing Fictions'. *The Journal of Philosophy* 75 (1978): 5-27.

Wilde, Oscar. 'De Profundis'. In *De Profundis and Other Writings*, edited by Coim Tóibín. London: Penguin, 2013 [1897].

Williamson, Timothy. *Knowledge and its Limits*. Oxford: Oxford University Press, 2000.

Wollheim, Ricard. *Art and Its Objects*, 2nd edn. New York: Cambridge, 1980 [1968].

Wood, Allen W. *Unsettling Obligations: Essays on Reason, Reality, and the Ethics of Belief*. Stanford, CA: CSLI, 2002.

———. 'The Duty to Believe According to the Evidence'. *International Journal for Philosophy of Religion* 63, no. 1 (2008): 7–24.

Zagzebski, Linda. 'Recovering Understanding'. In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, edited by Matthias Steup. New York: Oxford University Press, 2001.

Zangwill, Nick. 'Against Emotion: Hanslick was Right About Music'. *British Journal of Aesthetics* 44, no. 1 (2004): 29–43.

Zeman, Adam, Micaela Dewar and Sergio della Sala, 'Lives Without Imagery: Congenital Aphantasia', *Cortex* 73 (2015): 378-380.