

Desire and Practical Rationality

Ashley Shaw

Submitted in Partial Fulfilment of the Requirements
for the Degree of
Ph.D. in Philosophy



June, 2019

Declaration

I, Ashley Shaw, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis

Ashley Shaw

Abstract

Many of our desires spontaneously arise within us without prior decision or deliberation. Their presence exerts an influence on us, shaping our perspective and informing our decisions about what to do. This thesis starts by bringing a problem into view, one concerning how to explain the rational significance of such desires and the role they play in guiding rational action. It is often taken for granted that such desires can contribute to what it is rational for one to do. Yet predominant philosophical accounts of desire fail to provide a satisfying explanation of how such desires matter to practical rationality. Existing accounts of desire have centralised aspects of desire that might be enlisted to explain the rational significance of desire. These include their connection to motivation, affect and evaluation. These existing accounts, I argue, either fail to capture the contribution that such desires can make to rational agency, or threaten to leave us with a distorted and overly intellectualised account of desire.

I develop an account of such desires as serving to *attune* us to our reasons for action. I motivate the view that our conative system is part of our overall competence to recognise and respond to normative reasons for action in virtue of their biological function. The result will be an account on which desires are part of our capacity to respond rationally to our normative reasons for action.

Impact Statement

This thesis investigates the role that desires play in our capacity for practical rationality and to respond to normative reasons for action. This thesis connects a range of philosophical sub-disciplines: Philosophy of Mind, Philosophy of Action, Value Theory and Practical Reason (including Reasons for Action and Rationality). The central and organising problem has been previously raised in a more narrow fashion in the philosophical literature which this thesis has sought to develop and generalise. The negative part of the thesis (Chapters 3 and 4) engages with recently published work on these areas. The positive part of the thesis (Chapters 5 and 6) draws on work concerning desire, appetite, and motivation in Philosophical Psychology, Philosophy of Biology, Epistemology (esp. work on perceptual entitlement), to develop an account of how desires can contribute to what it is rational to do.

Contents

Abstract	4
Acknowledgements	8
Introduction	9
1 What it's Rational to Do	13
1.1 Reasons	13
1.2 Rationality	17
1.3 Practical Rationality and our Apparent Reasons	22
2 Desire and The Problem of Conative Significance	33
2.1 The Nature of Desire	34
2.2 The Problem of Conative Significance	52
3 Desire, Pleasure and Affect	63
3.1 Desire and Pleasure	64
3.2 Desire and Affect	80
3.3 The Regulatory Role of Affect	92
4 Desire and the Good	99
4.1 The Guise of the Good	100
4.2 Belief as a Model for Desire	102
4.3 Perception as a Model for Desire	111
5 The Function of Desire	123
5.1 Etiology and Function	124
5.2 The Guise of the Biological Good	129
5.3 Normal Functioning and Normal Constitution	137

5.4	'Representing as Good'	148
6	The Rational Role of Desire	153
6.1	The Road Behind, The Road Ahead	154
6.2	Competences and our Apparent Reasons	155
6.3	The Rationality of Acting on Desire	163
	Conclusion	178
	Bibliography	182

Acknowledgements

First and foremost, I would like to thank my primary supervisor, Lucy O'Brien. I am lucky to have been a student of hers. This thesis would not exist if not for Lucy's careful, critical engagement with the material contained within at every stage of development. One of her many talents as a supervisor is the ability to excavate philosophically promising ideas and instil in one the confidence to develop them rigorously. Often when progress ground to halt, a supervision with Lucy never failed to get me writing again.

I would also like to thank Mike Martin who provided important direction during the early stages of this project and for many rich and constructive supervisions which helped improve my understanding of the issues tackled here. I am grateful to Ulrike Heuer who provided important suggestions on how to improve some nascent ideas.

I have benefitted from discussing ideas in this thesis with Per Al-gander, Showkat Ali, Julian Bacharach, Vanessa Carr, Kirstine la Cour, Catherine Dale, Pete Faulconbridge, Alex Geddes, Alec Hinshelwood, Charles Jansen, Andrew Knox, David Olbrich, Edgar Phillips, Richard Rowland and Laura Silva. I would like to single out Poppy Mankowitz for numerous comments and suggestions that improved the clarity of this thesis. I would like to thank Arica Shaw for all of her love and support.

Finally, I would like to thank the AHRC London Arts and Humanities Partnership for a Research Studentship that funded my PhD.

Introduction

Many of our desires spontaneously arise within us without prior decision or deliberation. A desire to eat, take a stroll in the park or spend the afternoon playing chess can arise out of the blue. When they do, their presence exerts an influence on us; these desires shape our decisions about what to do. I call these our *basic* desires.

It is often taken for granted that our basic desires can contribute to what it is rational for one to do. Other things being equal, these desires seem to have the capacity to rationally initiate action, or the formation of intentions, aimed at satisfying the desire in question when they arise. This thesis starts by bringing a problem into view, one concerning how to explain the rational significance of such desires and their role in our capacity to act rationally. The Problem of Conative Significance, as I call it, is driven to a large extent by our inability to reconcile this intuition with predominant conceptions of the nature of desire. For this reason, our investigation into the rational significance of desire will be structured by an exploration of the different features ordinarily taken to be essential to, or part of, basic desire.

In what follows, I provide a broad overview of the thesis as a whole. I will explain the general shape of the thesis and provide a review of the individual lines of investigation that each chapter pursues, explaining at each stage the main positive and negative claims defended.

The Shape of Things To Come

The set up and development of the overarching problem will span Chapters 1 and 2.

Our enquiry will begin with an examination of the concept of practical rationality. In Chapter 1 (“What it’s Rational to Do’), I outline several

accounts of rationality and motivate the importance of an account of rationality that emphasises a relationship with normative reasons for action. In slogan form: practical rationality involves choices and acts that manifest sensitivity to our normative reasons for action. The picture of practical rationality sketched here will serve as the working account of practical rationality that informs the next five chapters; further developments are undertaken in final chapter of the thesis. As part of this, I explain and clarify a crucial notion, that of a reason for action. I motivate a ‘factualist’ account of reasons for action and explain the explanatory, normative and motivating roles that reasons can play. I end with a discussion about the way in which mental states can rationally support (‘rationalise’) certain psychological transitions such as the formation of beliefs or intentions by facilitating our ability to respond to reasons for action.

Chapter 2 (“The Problem of Conative Significance”) develops the central problem of the thesis. I begin by outlining what is perhaps the predominant account of desire, the motivational view. I clarify some important distinctions concerning desire that will be relied on in the rest of this work; in particular, the distinction between basic and non-basic desires. Having explained the target notion of desire, I motivate a natural and widely held intuition about the rational significance of desire. The problem I develop, The Problem of Conative Significance, is one of how to explain how our understanding of the nature of desire can vindicate such intuitions. Versions of this problem have been pressed in the literature. Most notably, Warren Quinn has argued that the predominant motivational view of desire fails to explain the power of desire to rationalise choice and action. Through an examination of Quinn’s negative argument, I motivate two constraints on answers to the Problem of Conative Significance.

What is left to be addressed in the rest of this thesis is the burden of explaining how we can provide a plausible and psychologically realistic account of the role of desire in rational agency that satisfies these two constraints. The following two chapters are concerned with extant attempts in the literature to respond to the Problem of Conative Significance.

In Chapter 3 (“Desire, Pleasure and Affect”), I consider a range of

views that centralise pleasure and affect in our understanding of desire. According to hedonic accounts of desire, at least some desires are rationally significant on account of their connection to states of bodily discomfort and pleasurable experiences. I raise a range of problems for two hedonic accounts; chiefly, I argue that even granting that desires are connected with these states of feeling, the rational significance of desire cannot be captured simply by reference to the reasons there are to seek pleasure or relief from discomfort. Next, I consider a different kind of account that emphasises the relationship between basic desires and conscious affective experiences. I argue that explanations of the rational significance of desire that appeal to the intrinsic features of such experiences do not provide us with an adequate final answer to the rational significance of desire. Nevertheless, as I will elaborate in a following chapter, I suggest that affect plays some role in regulating our basic desires.

Chapter 4, (“Desire and the Good”) considers a family of approaches to solving the Problem of Conative Significance that connect desires with representations of courses of action as normatively favoured, for instance, by being *good* or backed by *reasons*. I begin by considering two views that pursue this approach, aiming to explain the Problem of Conative Significance by appeal to a cognitivist or ‘perceptualist’ accounts of basic desire as being under the ‘guise of the good’. The first identifies desires with *beliefs* possessing normative content. I argue first that basic desires have features that cannot be explained if they were to be identified with beliefs and should be rejected for this reason. Second, I argue that the ‘desire as normative belief’ account freights basic desires with a kind of content that makes for an unacceptably intellectualised account of basic desire. The second approach pursues an analogy with perception. Though these strategies are promising, I argue that, as formulated, extant accounts either fall prey to the same problems or fails to provide an account of basic desire that informatively moves us beyond an analogy.

The negative work done in Chapters 3 and 4 clears the ground for the final two chapters which develops the positive proposal of this thesis which builds on the insights of the perceptual approach by focussing on the teleo-functional role of desires.

Chapter 5 (“The Function of Desire”) develops the positive proposal

of this thesis. I refocus attention on the mechanisms that underlie the production and regulation of our basic desires. Pursuing a position developed by Dennis Stampe with machinery developed by Ruth Millikan, I make the case that these mechanisms have the etiological function to produce desires that are conducive to the good of the organism.

Finally, in Chapter 6 (“The Rational Role of Desire”), I explain how the positive account provided addresses the Problem of Conative Significance. I begin by developing the connection sketched in Chapter 1 of the relationship between rationality and reasons, developing a view of rationality as responsiveness to our apparent reasons. Drawing on the work in Chapter 5, I argue that it is plausible to regard occurrent desires as manifestations of reasons-sensitive competences. I address several outstanding issues concerning whether the view can explain how desires contribute to the rational intelligibility of acting on desire and our reasons for acting when we act on desire.

Chapter 1

What it's Rational to Do

As rational creatures, we can decide what we do on the basis of considerations that we recognise to count in favour of various courses of action. The considerations that *do* count in favour of these courses of actions are *reasons*. Reasons play a role in guiding our decisions about what to do. When we are motivated to act by a reason, our reasons can explain our actions and *rationalise* our decision to pursue those courses of action. There is a connection then between acting rationally and responding to our reasons for action.

This chapter has two aims. First, I seek to outline and clarify these two central concepts as both will be regularly employed throughout this thesis. The second aim will be to articulate the connections between these two notions. This is important since I will also be relying on connections between these two concepts in order to articulate both the central problem of this thesis discussed in the following chapter and its solution.

1.1 Reasons

Suppose you are trying to decide whether to pursue one course of action ϕ or another ψ . In order to decide, one might turn one's mind to things that would support or favour ϕ -ing and those that support ψ -ing and try to weigh these reasons up. Things that count in favour of actions are *normative* reasons for action. We appeal to normative reasons to *justify* actions. Take for instance claims to the effect that some reasons obtain, e.g. "there is a reason for Uma to leave the building", or a claim

that identifies a consideration as a reason for someone to act in a certain way, e.g. “that there is a fire is a reason for Uma to leave the building”. For this reason, normative reasons are sometimes also called *justifying* reasons.

On the standard way of explaining the concept, a normative reason, to some degree, *counts in favour* of a certain response (e.g. leaving the building, intending to leave the building, believing that such and such, etc.) (Scanlon, 1998, p. 17).¹ It will be helpful also to articulate the notion in relational terms: something that stands in a certain relation, the *favouring-* or *reasons-relation*, to an agent *A* and a response ϕ .²

What kind of things stand in the favouring-relation? The standard view is that the elements that stand in normative favouring-relations are facts (Parfit, 2011; Scanlon, 1998; Alvarez, 2010; Raz, 2011).³ This position is highly credible. Facts are uniquely suited to play the role of normative reasons. Normative reasons are supposed to be considerations that *justify* an agent’s action and in order for something to justify, it has to be a fact.

What kinds of things are *favoured* by normative reasons? As we have seen, there can be reasons for *actions* but also, *inter alia*, reasons for *belief*, reasons to *feel*, reasons to *desire* and reasons to *intend*. Since my interest will primarily be based on *practical* reasons, most of my attention will be focussed on the reasons for action, but also reasons for desire and reasons to intend.

¹This is not offered as any kind of reductive analysis or definition of the notion of a reason. As Scanlon notes “any attempt to explain what it is to be a reason for something seems to lead [one] back to the same idea: a consideration that counts in favor of it” (Scanlon, 1998, p. 17). With many philosophers, I will treat the notion of a reason and the ‘favouring’ relation as primitive.

²It is plausible, following Scanlon (2014, p. 31), that the reasons-relation is relative also to *circumstance*, such that reasons are considerations that count in favour of an agent’s responding in a certain way in a particular *circumstance*.

³Within this position, there is a sub-dispute which I will stay largely neutral with respect to concerning how we should think of facts. Are facts exemplifications of properties by objects? States of affairs that obtain? True propositions? For discussion, see Dancy (2000, pp. 112–20).

1.1.1 Strength and Weights

Normative reasons can individually vary in strength, so some reasons can be stronger or weaker to various degrees compared to other reasons.⁴ If a reason favours a response ϕ to some non-zero degree, it is a *pro tanto* reason to ϕ . These reasons can work together, each contributing some weight to pursuing a course of action.

Sets of reasons can outweigh other sets of reasons. For example, the fact that my favourite band is performing at a local venue is a reason to attend, but it is *outweighed* by the fact I promised my manager that I would work that night. Normative reasons can also interact with other reasons by affecting their weights, for instance some reasons do not just outweigh other reasons but can reduce the strength of, or *undercut*, other reasons; in some cases, some reasons are said to reduce the strength of other reasons to zero, that is *disable* these reasons.

Some reasons can be *decisive* or *conclusive* reasons to act. When a reason is a decisive reason to ϕ , it is stronger than any reasons there are to respond in some other way, including refraining from ϕ -ing (Parfit, 2011, p. 32). For example, the fact that I promised my manager I would work is a decisive reason for me not to go to the gig. There are cases when the balance of reasons do not determine a unique course of action; that is, there is no one course of action for which there are decisive reasons, but rather, there is *sufficient* reason to pursue two or more courses of action. In such cases, the options backed by sufficient reason constitute *eligible* options for choice (Raz, 1999, Ch.2-3).

1.1.2 Reasons, Correctness and Objective Oughts

Normative reasons play a role in determining what we ought to do. We have to be careful to specify this more carefully since ‘ought’ is a highly context-sensitive term. Here I want to focus on uses of ‘ought’ which are fixed solely as a function of *the total balance of normative reasons*. In a context where the total balance of reasons is such that there is a decisive reason for A to ϕ , then A ought to ϕ . These are cases in which ‘ought’ is ‘decisive reason-implying’ (Parfit, 2011, p. 33) or *objective*. Objective

⁴For a collection that extensively discusses the weights of reasons, see Lord and Maguire (2016).

oughts contrast with subjective oughts which depend on an agent's beliefs, desires and so on. We can illustrate objective oughts by considering the kind of *advice* we would give from a God's eye perspective. For example, if Uma is allergic to nuts and the cake she plans to eat, unbeknownst to her, contains nuts, she objectively ought not eat the cake. Relative to an agent A and a circumstance c , I will refer to the response favoured by the totality of reasons as the response that would be *correct* for A to do in c .

1.1.3 The Normativity of Practical Reasons

I want to end this brief discussion of normative reasons by backgrounding a foundational 'meta-normative' issue about the normativity of reasons, specifically, the fiercely debated question that asks *in virtue of what* certain facts constitute reasons for us to act.

A reason R for an agent A to act in way ϕ is a fact that counts in favour of ϕ -ing. For example, the fact that going to the shop will allow me to eat ice cream is a reason there is for me to go to the shop. According to what we can call the *Pure Desire-Based* account, the most fundamental explanation of why some R constitutes a reason to ϕ consists in the fact that R promotes the satisfaction of some (actual/counterfactual) desires that one has (cf. Williams, 1979; Schroeder, 2007). What explains why the fact that going to the shop will allow me to eat ice cream favours my going to the shop is ultimately that it will promote some (actual/counterfactual) desire I have for ice cream. I call such desire-based account *pure* because they seek to explain every normative reason for action by reference to (actual/counterfactual) desires of the agent.

In contrast, what we can call the *Pure Value-Based* account says that the most fundamental explanation of why some R constitutes a reason for A to ϕ consists in the fact that R conduces or promotes one's ϕ -ing where ϕ -ing is valuable in some respect (e.g. Quinn, 1994a; Parfit, 2011; Raz, 1999). For example, what explains why the fact that going to the shop will allow me to eat ice cream favours my going to the shop is ultimately that eating ice cream is valuable, e.g. because it will bring pleasure, or that its sugar content meets a biological need. Pure Value-Based accounts are typically said to provide an 'externalist' grounding to

reasons because whether a course of action is valuable on this view is a mind-independent fact about the nature of the relevant course of action.

There are a range of other positions, for instance *Constructivist* accounts that aim to explain reasons in terms of the outcomes of rational reasoning or procedures.⁵ In between these poles are hybrid positions which appeal to multiple sources of normativity and allow that different reasons are grounded in different sources (see e.g. Chang, 2013). This thesis does not seek to defend a particular position on this front. Nevertheless, it is important to briefly background this issue because many of positions and issues concerning, for instance, the nature of desire, which I will discuss in the following chapter, are developed with an eye to defending or attacking various answers to this ‘meta-normative’ question.

1.2 Rationality

With this important notion of a normative reason for action in place, we turn now to *rationality*.

We use the adjective ‘rational’ to apply to a myriad group of entities. To focus our attention, I want to isolate first purely *capacity-ascribing* uses of ‘rational’. We can say for instance that a certain *individual* or more typically a *kind* is rational, by which we mean that they possess a certain set of capacities, abilities or powers (the power of Reason). For example, we might say that humans are rational or possess the power of Reason. Correlative with this are uses of ‘arational’ to withhold or deny the ascription of such powers, as when we say for instance that paramecia or caterpillars are arational creatures.

There are also *standard*-related uses of ‘rational’. Here the targets are more wide ranging, they can include individuals but also actions (mental or physical), (sets of) mental states or processes, as well as transitions or relations between mental states and actions. Whether an action, mental state, or some relation between these elements is rational or irrational turns on how it relates to standards or norms associated with rationality.⁶

⁵Kant of course is the key influence for contemporary defences of constructivist accounts, which include for example, Korsgaard (1996) and Street (2008).

⁶There is a connection back from these assessments of the rationality to assessments of the rationality of individuals. For instance, when an individual’s action is irrational or when she forms a mental state irrationally, we can then say that the agent is being

Here I will mainly be concerned with standard-related uses of ‘rational’. Given our focus on practical rationality, I will primarily focus on cases in which targets of rational assessment focus on rational *agents*, their *actions* and the psychological precursors to action, e.g. decisions, choices or intentions to pursue a course of action, and the relations or transitions between these elements.⁷

In this section, I want to focus on a standard of rationality that directly pertains to apparent reasons for action. The starting point for me will be the thought that there is a close connection between acting rationally and responding to one’s (apparent) reasons. I am interested here in how possession of reasons for action can rationalise intention and action, and how various mental states contribute to what it is rational to do by affecting our apparent reasons for action. I will not be interested in providing a general account of rationality, as such I will remain silent on several important issues concerning the nature of rationality. For instance, it is plausible that full rationality requires that one not possess certain ‘incoherent’ combinations of attitudes. An issue I will not be engaging with will be about how to understand these ‘structural’ standards of rationality and whether or not they can be explained by reference to reasons we have to be coherent as such.⁸ Here I will be assuming that such coherence requirements do not exhaust all the requirements of rationality and that there are some standards for being rational that are a function of our apparent reasons for action.⁹

1.2.1 Rationality and Mental State Transitions

In order to clarify the standard of rationality I will be interested in, focus here on psychological transitions from a set of ‘base’ states ‘ M_1, M_2, \dots ’ to a ‘concluding’ state M_n .¹⁰ When we evaluate the rationality

irrational, or less severely, that she is being less than ideally rational in pursuing that course of action.

⁷Some theorists, e.g. John Broome, have claimed that assessments of rationality concern only things over which we have direct control and so rationality strictly only concerns mental states, not actions (Broome, 2013, p. 89).

⁸For discussion, see Broome (2007), Kauppinen (forth), Kiesewetter (2017), Kolodny (2005), Lord (2014), Way (2018) and Wedgwood (2017).

⁹For a defence of this point against, for instance, Broome (2013, Ch.1), see Lord (2014).

¹⁰This terminology comes from Peacocke (1999) and O’Brien (2005).

of a transition between the base states and the concluding state, we are interested, *inter alia*, in:

- the *contents* of M_{1-n} ,
- the *type* of mental states M_{1-n} are

For instance, if a subject believes that p and believes that *if p , then q* , then an explanation of why it is rational for her to believe that q must refer to the fact that she believes rather than imagines that p , and that it is a belief with the content that p rather than that not- p . In the practical realm, if one desires to ϕ and believes that ψ -ing conduces to ϕ -ing, then similarly, it will matter to the rationality of a corresponding choice to ψ that it is a *desire* with the content it has and a means-ends *belief* with the content it has that forms the base of the psychological transition. Further, we are interested also in:

- the relationship between the base states and the concluding state(s)

For instance, if a subject forms the belief that q by accident because she has a strange one-off neurological condition which causes her to believe that q whenever she believes that p and *if p , then q* , it would not be correct to describe this transition as *rational*.

1.2.2 Rationality and ‘External’ Correctness

When we evaluate the rationality of a psychological transition, we are centrally interested in the nature and relation between psychological states and their contents as opposed to the immediate relationship between these mental states and the external world. For instance, in the above cases, the falsity of the resulting belief that q would not *ipso facto* impugn the rationality of the transition from the base states provided the base states were themselves rationally formed and suitably connected with the resulting belief that q . Similarly, the fact that the agent’s intention to ψ would not be, all things considered, something there is sufficient reason to do would not *ipso facto* impugn the rationality of the intention. However, this is not to say that the standards we bring to bear when we evaluate psychological transitions as rational have *nothing* to do with the

external world, for instance, conduciveness to truth or acting with reason. I will shortly suggest that transitions can be rational if they employ competences to respond to objectively determined truth- or value-related reasons.

This point can be made by distinguishing the standards of rationality from the standards of correctness. In the previous section, we defined the notion of *correctness* where a response ϕ is correct just in case it is the response favoured by the total balance of normative reason. A response ϕ 's being correct is neither a necessary nor sufficient condition for ϕ 's being a rational response. To see that it is not necessary, note that one can rationally form the belief that it has been raining outside even if one falsely believes that the streets outside are wet on the basis of perception. To see that it is not sufficient, recall that rational assessment of a response is sensitive also to the *relationship* between the response and the base states. For example, if I form the belief that it has been raining based on wishful thinking, my belief would be correct though not rational even if I also believe that the streets are wet.

1.2.3 Rationality and Styles of Warrant

Evaluation of rationality has an internalist or perspective-dependent character. In order to further clarify what this more 'internal' aspect consists in, I want to turn to considering the relationship between rationality and a notion of warrant familiar in contemporary debates in epistemology, in particular, focussing on the work of Tyler Burge and Christopher Peacocke. I will begin by considering epistemic warrant and its relationship with rationality since this has been the focus of both philosophers; once the idea is on the table, I will return to practical rationality.

Epistemic warrant is a normative property which truth-apt states like beliefs and judgements possess when formed and maintained in a particular way that makes them reliably conducive to knowledge. The particular way in question turns on the nature and/or content of the states underlying the formation of belief. Like epistemic rationality, it is not required for a belief to be warranted that it be *true*. What is required is that it is, in virtue of its etiology, "well positioned to achieve or indicate truth in normal circumstances, given certain limitations in

the individuals abilities and perspective” (Burge, 2003, p. 506).

Burge outlines two sub-species of epistemic warrant: (i) entitlement and (ii) justification (Burge, 2003). Whether a belief is one a subject is entitled to form turns on whether it is formed on the basis of psychological antecedents that have a nature and content that make the formation of the belief reliably truth-conducive under certain conditions.¹¹ In contrast, justifications are a class of warrants that warrant the formation of a belief in virtue of a reason being represented as the basis for forming the belief. One key differentia of justifications is the requirement that the subject have *access* to the item that constitute the reason for belief. The relevant notion of conceptual accessibility operative in Burge (2003) is simply that the subject has the conceptual competence to entertain the putatively reason-providing consideration. In contrast, according to Burge, some warranted beliefs need not involve the grasp of any reason, as when they are ones to which the subject is entitled.¹²

¹¹It is important that a source of entitlement provides warrant not *simply* by appeal to its reliability (although entitlement generating sources are normally reliable under normal conditions). According to Burge, the reliability of the transition from entitling state to belief must be *grounded* in the nature and intentional content of both psychological states. For instance, on Burge’s view about how perception can confer entitlements to perceptual beliefs, he appeals to the idea that perceptions share a representational function, and have their contents determined anti-individualistically. The normal conditions in question “[are] the one[s] by reference to which the perceptual content of the perceptual state is explained and established” (Burge, 2003, p. 532). For more details, see Burge (2003, pp. 530–7) and Burge (forth. pp.4-5) in which Burge describes this as the representation function of perception-belief transitions. These aspect of entitlements are similarly endorsed in Peacocke (2002) and Peacocke (2003), both of which contains an analogue of both the reliability and nature-content grounding condition, what Peacocke titles the ‘Special Truth-Conduciveness thesis’ and ‘Rationalist Dependence thesis’.

¹²As we see here, Burge defines entitlements *via negativa*: they are warrants that do not require that the subject represent the reason as a proposition. Some epistemologists have mistakenly interpreted Burge as developing a more demanding notion of justification, e.g. Silins (2012) and Casullo (2007). Silins interprets the notion of justification as satisfying high internalist standards such that the subject must conceptualise the putatively justifying reason *as such*, that is, she must represent the reason R (have “access₁” to R) and represent R as a reason (have “access₂” to R) (Silins, 2012, p. 247). In contrast, Silins holds that “[e]ntitlements should be defined as warrants that one can have *without satisfying any... access conditions*” (Silins, 2012, 247–8, emphasis added). Burge has explicitly denied this reading of his work in clarificatory remarks, see Burge (2013) and Burge (forth.). Burge’s views have changed since “Perceptual Entitlement” in light of considerations to do with modular reasoning at the sub-personal level. These involve representations “in an individual’s psychology that are not consciously thinkable by the individual” (Burge, forth. p.121). See Burge (2013, pp. 306–41) and Burge (forth.) where the entitlement-justification contrast is drawn in a subtly different way as a result.

In the epistemic realm, Burge and Peacocke tie rationally held attitudes with warrant in general, such that attitudes one is entitled to have are said to be ones that it would be *rational* for the subject to form (Peacocke, 2002, p. 385). On Burge’s view, perceptual belief formation can be rational even if not formed, in his sense, *for reasons*.

Things however are slightly different with respect to matters of rational *action* and practical rationality more generally. There is a close connection between intentional action *per se* and *reasons*. When we act intentionally, we typically know why we act as we do. Anscombe emphasised this in distinguishing intentional action as actions for which “a certain sense of the question “Why?” is given application; the sense is of course that in which the answer, if positive, gives a reason for acting” (Anscombe, 1957, p. 9). The answer to Anscombe’s *why*-question gives the *subject’s reason* for action, the thing in light of which the subject acted. Given our focus on practical rationality, we will need to further elaborate the connection between rationality and acting for reasons. Nevertheless, as we will see, practical rationality, like warrant more generally, involves a certain kind of competent exercise of one’s capacity to achieve success, in this case, to act for reasons, given limitations in the agent’s ability or perspective.

1.3 Practical Rationality and our Apparent Reasons

The previous section tells us that when we rationally evaluate various actions or intentions, we are interested in how the action or intention is undergirded by reasons that constitute the subject’s reasons to act or intend. I take a closer look in this section at the relationship between rationality and reasons for action. I begin first by considering what is required for one to act *for* a reason and explain how reasons feature in explanations of action that simultaneously rationalise that act. This will give us the opportunity to discuss the notion of a motivating reason and clear up some contentious issues surrounding how to understand motivating reasons for action. Finally, I outline the notion of an apparent reason for action and explain how rationality relates to one’s apparent

reasons for action.

1.3.1 Acting for Normative Reasons

I want to begin with the idea of acting for a reason. Reasons are facts that favour certain courses of action. Under certain conditions, we are in a position to respond to normative reasons. It is important to distinguish when action constitutes a proper *response* to a normative reason R from when that action merely is in accordance with R and the act it favours. What is the difference? Acting merely in accordance with a reason to ϕ requires only that we ϕ and that there be a reason to ϕ . In contrast, when we act in *response* to a reason there is to ϕ by ϕ -ing, we are sensitive to:

- (i) the *fact* that favours our ϕ -ing, and
- (ii) the *favouring-relation* that the fact stands in to our responding in the way we do.

To illustrate this, suppose that there is an aggressive bear in front of one. This fact is a reason to make oneself look bigger. Correlative with each of these two conditions, there are two ways in which we can come up short when we fail respond to our normative reasons.

First, one might simply fail to register the fact that there is an aggressive bear in front of one. In this case, one fails to be aware about the *facts* the obtaining of which favours a certain course of action. If one is not aware of the fact that favours a certain response, then one cannot respond to that fact. Plausibly, proper responsiveness to reasons requires that one stand in some *epistemic* relation to the fact that constitutes the normative reason there is for one to act. There is some dispute about what this epistemic relation consists in; for now, let us suppose that the weakest relation that one must stand in to the favouring fact consists in a belief that the fact obtains.¹³

Second, one might be aware of the fact that obtains, but fail to recognise or understand that this fact counts in favour of acting a certain way.

¹³See Lord (2018) for a nice survey of the various options, including the distinction between ‘high bar’ and ‘low bar’ epistemic relations, some require that reasons simply comprise the content of some presentational state (e.g. Schroeder, 2008a) whilst others require that one know the reason in question (cf. Hawthorne and Magidor, 2018).

A child ignorant of the danger might, for instance, accidentally act in accordance with what there is reason to do (make oneself look as big as possible) by trying to hug the bear (e.g. because he thinks it looks cuddly), but he fails to *properly respond* to the reason since his actions do not manifest an appropriate sensitivity to what that fact favours. In addition to the epistemic condition, the following is a plausible condition on the action undertaken in the belief that the reason obtains: the act manifests a sensitivity to the favouring relation that the fact stands in to the response. For now, I capture this sensitivity condition with the claim that the subject ‘takes’ R to be a reason, where ‘taking’ is simply a placeholder notion.¹⁴

When these two conditions are satisfied, we are said to possess or *have* the reason. It is important to distinguish this substantive notion of possession from a weaker sense of ‘have’ on which the mere existence of a normative reason R for A to ϕ is one that a subject ‘has’. Schroeder (2008a) calls this the ‘pleonastic’ sense of *having*. For example, there are extremely good reasons for anyone to quit smoking because smoking drastically increases the risk of cancer. Bertrand Russell, who knew nothing about the risks about smoking, therefore, in this pleonastic sense, ‘has’ a good reason to stop smoking. This however is not the sense of having that is relevant here, and as we will see, only reasons that we non-pleonastically possess will be centrally relevant to assessments of rationality. Russell, given his perspective, cannot be accused of irrationality given that he knew nothing about the risks. This is reflected in the propriety of blame; it would be out of place to say that Russell only had himself to blame should he become ill. Having a reason puts a subject in a position to act *for* that reason; the reason can be *her* reason for responding as she does. Philosophers call such reasons an agent’s ‘motivating’, ‘personal’ or ‘operative’ reasons for acting.

¹⁴I will hold off, until §1.3, discussing the various conditions that have been proposed to capture what I have glossed here as ‘appropriate sensitivity’ or ‘taking’; some philosophers require demanding conditions, such as that the subject *believe* the fact in question be a *reason* (Raz, 2011). I favour a weaker condition articulated in terms of reasons-sensitive competencies.

1.3.2 Reasons and Rationalising Explanation

Earlier we said that when we rationally assess a transition from a set of base states to a concluding state, we are interested in (i) the nature of both the base states and the concluding state and (ii) the relation between these states. I want to connect this thought now to the notion of acting for reasons.

Suppose we see that Uma is walking to the kitchen. Consider the following two ways we can explain her action (mutatis mutandis for an intention):

Psychological Explanation: Uma is walking to the kitchen because she wants to eat a snack and believes that there is a snack for her to eat in the kitchen.

Reasons-citing Explanation: Uma is walking to the kitchen because there is a snack for her to eat in the kitchen.

Both of these are what philosophers call *rationalising* explanations of action because they provide an explanation of action that affords an understanding of the action as rational. These are related. When the explanation cites Uma's reason as being that there is a snack in the kitchen, it rationalises her action against the presupposition that Uma *believes* that there is a snack to eat in the kitchen, that she has eating a snack as a goal of hers, e.g. if she desires to eat a snack, and that her action is suitably related to these two psychological elements. We see her action as rational on account of the nature of the relevant psychological states, their contents and the relationship with her action.

When the explanation cites Uma's psychological states, we are able to, extrapolating from the content of her beliefs, discern Uma's *reason*, viz. the fact that there is a snack for her to eat in the kitchen, presupposing that, in acting, she manifests a sensitivity to the relationship between this fact and the goal we take her to have, viz. to eat a snack.¹⁵

¹⁵As Davidson puts it, such explanations rationalise by "lead[ing] us to see something the agent saw, or thought he saw, in his action— some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable" (Davidson, 1963, p. 685). So, when we learn Uma's reason for walking to the kitchen, we are led to see, against some background assumptions, some feature of walking to the kitchen that *Uma* saw in the pursued course of action.

The role of belief in particular here serves to introduce reasons into our psychology. As Stampe writes, “[w]hen we reason from our beliefs it is from *what* we believe— the object of our beliefs— that we reason: the facts as we believe them to be” (Stampe, 1987, p. 337).

It is important not to confuse an *agent’s reason* (her motivating reason) from the set of base *mental states* that characterise the relevant psychological transition undergirding her action. Though we can construct a rationalising explanation of Uma’s action in which the explanans comprise either *her reason* or the *mental states* that comprise the base state, these are two distinct kinds of items. Uma’s reason is that there is a snack in the kitchen, a fact that counts in favour of her walking to the kitchen. In this case, her reason is a good one, it is a *normative* reason that favours her action. It is extremely natural to think that, at least in successful cases when the subject correctly recognises a fact and the act it favours, an agent’s reason for action is simply the fact that does the favouring (Dancy, 2000, pp. 103–4).¹⁶ But her *believing* and her *desiring* is not the thing that we typically regard as counting in favour of actions.

Though it is true that in order for something to be an agent’s reason for acting she must stand in some epistemic relation to that reason, it does not follow that her reason, the reason that motivates her, is to be *identified* with the mental states that enabled the subject to act for that reason (Dancy, 2000, p. 127). Nor should we be misled by the fact that we can truly assert that Uma’s reason is her belief that there is a snack in the kitchen. In general, phrases of the form “A’s belief that *p*” can be read objectively such that the phrase denotes *what she believes* (i.e. that *p*), or it can be read attitudinally where the phrase denotes a psychological state, her *believing* that *p* (cf. Alvarez, 2008, Alvarez, 2010). So though it is true that for a reason to motivate, one must believe that it obtains, it does not follow that it is her believing that *p* which constitutes her motivating reason.¹⁷

Naturally in this case what the agent saw that we are ‘led to see’ is that walking to the kitchen would allow her to eat a snack, something she wants to do.

¹⁶As Ulrike Heuer puts it, in cases where a subject acts for a good reason, “that very reason is also the explanation (or at least part of the explanation) of why she did what she did. Normative or justificatory and explanatory reasons are the same reasons in such a case and not different kinds of reasons altogether” (Heuer, 2004, p. 45).

¹⁷These considerations strike me as decisive against the view in the literature known

Unlike some philosophers, I will resist bifurcating the notion of a motivating reason. For instance, some philosophers distinguish *two* senses of ‘reason’ (Ginsborg, 2006, p. 289). Ginsborg, employing the subscript strategy, distinguishes between a ‘reason₁’ for belief which is “[a] fact which presents itself to the subject as favouring belief” (Ginsborg, 2006, p. 290) and ‘reason₂’ for belief which is “a psychological state, typically another belief, in the light of which her original belief can be recognised, from a third-person perspective, as rational” (Ginsborg, 2006, p. 290). I will take motivating reasons, and reasons for belief, to be solely of the first kind. An agent *A*’s reason for ϕ -ing is some fact, one that she represents as obtaining, which she is appropriately sensitive to as favouring ϕ -ing, which explains in the right way her ϕ -ing. The framework here then is a form of *Factualism*, as opposed to Psychologism, about motivating reasons.

In general then, psychological explanations are not *reasons*-citing explanation, though, as noted above, they can lead one easily to discern the reasons in question. This has to be carefully understood so as not to rule out cases in which a fact about one’s psychology *is* a reason for one to pursue a course of action, for which a subject might pursue that course of action. For example, one might go to the psychiatrist because one believes that there are pink dragons living inside one’s ear (example adapted from Hyman, 1999, p. 444). This kind of fact is clearly not what is going on in Uma’s case. Another thing that it is important not to rule out in this framework is that there are also cases in which the manner in which a subject’s conscious perspective on a content can constitute a reason for a certain response. For instance, in consciously judging that *p*, the phenomenal character of *judging* that *p* can constitute a reason for

as *Psychologism* about the category of ‘motivating’ reasons. A central proponent of psychologism is Michael Smith who defends the view that an agent’s motivating reason for action consists in mental states, specifically, a belief-desire pair. Smith’s form of psychologism, taking his lead from Donald Davidson, flows from a ‘causalist’ view about action explanation according to which actions are distinctive in being caused ‘in the right way’ by mental states, centrally states of belief and desire (the ‘standard story’ of action) (Smith, 2012a, p. 387). For Smith, “the distinctive feature of a motivating reason to ϕ is that, in virtue of having such a reason, an agent is in a state that is *explanatory* of her ϕ -ing” (Smith, 1994, p. 96). For motivating reasons to play an “explanatory role in producing action” motivating reasons have to be, as he puts it, “psychologically real” (Smith, 1994, p. 96). For this reason Smith thinks that motivating reasons are “the desire and belief pairs that figure in the standard story” (Smith, 2012a, p. 388).

one to self-ascribe the judgement that p (cf. Peacocke, 1999; Peacocke, 1999, ch.4; O'Brien, 2005, McHugh, 2012; *pace* Byrne, 2011).

1.3.3 Error Cases, Motivating Reasons and Apparent Reasons

So far, the discussion has focussed on cases in which a subject acts for a reason, a fact she correctly represents as obtaining to which she is appropriately sensitive to as favouring the course of action she undertakes. We can represent the transition here as one in which a subject moves from a set of base states containing, inter alia, a belief the content of which represents the reason-constituting fact to an intention or an action. In this section, I consider cases in which subjects act on a *false* belief, where she incorrectly represents a state of affairs as obtaining.

Consider a variant of the case in which Uma falsely, but let us say justifiably, believes that there is a snack in the kitchen for her to eat. In this case, there no fact that corresponds to the representational content of her belief that there is a snack to eat in the kitchen. Given the way that I have set things out, there is no fact to which she stands in the appropriate epistemic relation that is a normative reason for her to go to the kitchen. Nevertheless, many philosophers have claimed that Uma, nevertheless, acts for a reason when she goes to the kitchen. This might seem to count decisively against Factualism about motivating reasons, for if a motivating reason R is a fact that count in favour of an action ϕ , then R obtains, but in error cases like this one, A 's reason for ϕ -ing fails to obtain.¹⁸

This has given rise to a dispute about whether or not Factualism is the right account of motivating reasons. Some philosophers have taken cases like this to motivate an alternative position on which motivating reasons are simply *propositions* that the subject believes, which are such that, if true, would favour her action, a relation that she is appropriately sensitive to. Call this Propositionalism about motivating reasons. I do not aim to vindicate the Factualist here; none of what follows in this thesis hinges on the proper resolution to this issue. Nevertheless, I

¹⁸Variants of this problem are widely discussed in literature; for a representative sample of various positions and how they grapple with this problem, see Williams (1979), Alvarez (2016), Comesaña and McGrath (2014) and Hornsby (2008).

will persist in framing notions like motivating reasons as the Factualist understands them, and to that end, I should articulate some way in which Factualists can make sense of error cases. The Factualist has to say something about what motivates the subject in cases of error, specifically what it is in light of which a subject acts when she acts on a false belief. It is this issue to which we now turn.¹⁹

I want to introduce some terminology now that allows us to make progress. Rather than call that in light of which the subject acts in error cases, a *reason* that motivated the subject, let us call it an *apparent* reasons that motivated her (cf. Alvarez, 2016; Parfit, 2011)

An apparent reason is something that the subject ‘takes’ to count in favour of her action. Apparent reasons can be normative reasons or *merely* apparent reasons. As Parfit puts it, an apparent reason might be a “real reason” (that is, a normative reason) or something that “merely appears to be a reason” (Parfit, 2011, p. 35).²⁰ If a subject take a consideration that R to count in favour of ϕ -ing, and R is a fact that counts in favour of her ϕ -ing, then R is a normative reason there is for her to ϕ . If a subject takes R to favour ϕ -ing, but R is not a fact that favours ϕ -ing either because (i) it is not a fact that R or (ii) R does not favour ϕ -ing, then R is a *merely* apparent reason. If a subject acts for a merely apparent reason that is not a normative reason, I will call this the agent’s apparent reason for action, one that was merely apparent.²¹ Here we can remain silent on whether a merely apparent reason that motivates an agent can be described as an *agent’s reason*.

When is something an apparent reason for action? Here I have glossed apparent reasons as considerations taken to count in favour of actions. However, as earlier noted, ‘taking’ here is a placeholder notion used to capture the idea that in order for one to have a reason to ϕ , one must

¹⁹A proper resolution of this issue will come by re-examining how the theoretical category of motivating reason is demarcated, in particular, whether or not the notion of motivating reason can fruitfully be articulated in a way that is divorced from the notion of a normative reason. This work is outside the scope of this thesis.

²⁰Parfit later goes on to say that “[w]hen we have. . . a reason or apparent reason, and we act for this reason, this becomes our motivating reason”. Alvarez would presumably dispute this, see (Alvarez, 2016, pp. 3304–7).

²¹This phrase ‘A’s apparent reason for ϕ -ing’ is slightly awkward since it might be read non-neutrally as talking about a reason that appears to be A’s reason, rather than, as I intend, an apparent reason that motivated the agent. It is only the later that I will use this phrase to denote.

not just (i) stand in the appropriate epistemic relation to that reason but also (ii) be appropriately sensitive to the favouring relation that the reason stands to ϕ -ing. I will reserve discussion of how to capture this notion of appropriate sensitivity until later; for now, I want to simply make a negative point against a prominent account of apparent reasons which I claim fails to respect this idea.

In his gloss on apparent reasons, Pfitzner writes that “we have some apparent reason when we have some belief whose truth would give us that reason” (Pfitzner, 2001, p. 25). According to Pfitzner, we respond to these reasons if our awareness of these reasons “lead[s] us to do, or try to do, what we have reason to do” (Pfitzner, 2001, p. 31). This however is too weak a characterisation of what it is for one to respond to an apparent reason one has. If Pfitzner is right, then R is an apparent reason for A to ϕ just in case we have some belief whose truth is a reason to ϕ and our ϕ -ing constitutes a response to R if the corresponding belief leads us to ϕ , at least absent some explication of ‘lead to’. But this cannot be sufficient for one to have an apparent reason and respond to that apparent reason given the ‘appropriate sensitivity’ condition I outlined in §1.3.1. Recall the case in which the gormless child believes that there is an aggressive bear in front of him but fails to understand the danger it poses. His belief is one whose truth is a reason to stretch out his arms, and his doing just that is explained causally by that belief. However, his actions are not a manifestation of a sensitivity to the reason’s character *as a reason* to stretch out his arms. This point applies also to reasons for belief. Suppose Tom, someone of low mathematical ability, believes that p , where p is some true mathematical proposition. And suppose that p entails q , where q is some arcane mathematical proposition that Tom is in no position to deduce. Though it is true that Tom’s belief that p is one whose truth is a reason to believe that q , it does not constitute an *apparent* reason for Tom to believe that q .²²

1.3.4 Rationality and our Apparent Reasons

As I emphasised earlier, rationality is not simply *external correctness*. An intention or an action can be correct, that is, in accordance with the

²²For further discussion, see Sylvan (2014, §3).

act favoured by the total balance of reasons without being *rational* for the subject to have intended or done. Conversely, it can be rational for a subject to intend or do something that would be contrary to what is correct for her to do. This point can be put in terms of objective *oughts*. What it is rational for one to do need not be what one objectively ought to do, nor is what one objectively ought to do what it would be rational for the subject to do. The case of error just considered demonstrates all these points nicely: given that going to the kitchen would be a waste of time for Uma, it would not be correct for her to do so, so hence she objectively ought not do it. But given her false belief that there is a snack to be eaten there, she has a merely apparent reason to do so which can make it rational for her to do so.

This thought needs to be qualified further. Whether it would be rational for someone to (intend to) ϕ is sensitive then not to the overall balance of reasons external to the perspective of the subject but the agent's set of *apparent* reasons. This set of apparent reasons will include: (i) facts she believes to be the case, that she is appropriately sensitive to as favouring her actions and (ii) states of affairs that are not the case that she takes to favour her action (her merely apparent reasons). The exact connection between what it would be rational to do and one's apparent reasons will need to take into account the agent's total perspective on what apparent reasons there are for her to act, and how their weights and strengths agglomerate to determine the total balance of *apparent* reasons. The following is a plausible connection between apparent decisive reasons and an important standard of practical rationality:

- If, from A 's perspective, she has apparent decisive reasons to ϕ , then it is rationally required that A intend to ϕ .

A subject who fails to respond to her apparent decisive reasons is less than fully rational (Parfit, 2011, p. 34). As mentioned earlier, it is not always the case that the total balance of reasons are decisive, that is, determine a unique outcome as favoured. In such cases, it is possible that more than one option is favoured by sufficient reasons; in this case:

- If, from A 's perspective, she has sufficient apparent reason to ϕ , then it would be rational for A to intend to ϕ .

In such cases, the range of options backed by sufficient apparent reasons are said to be rationally eligible options for the subject to pursue. Given that there is no decisive reason to select between the range of eligible options, it is not a failure of rationality if one picks one over any other eligible option.

Conclusion

The aim of this chapter has been to lay down some conceptual foundations, (partially) explicating concepts like reasons and rationality which will be extensively utilised in the rest of this thesis. I provided some background about normative reasons for action (§1.1), started work demarcating a particular standard of rationality related to one's reasons for action, and distinguished between this standard of rationality and the notion of correctness (§1.2). Next, I further clarified the relationship between practical rationality and reasons for action (§1.3). I clarified some constraints on acting for a normative reason, explained the role of reasons in rationalising explanations of action and introduced the notion of an apparent reasons for action. Finally, I outlined two plausible norms of rationality that relate to one's apparent reasons.

Chapter 2

Desire and The Problem of Conative Significance

Many of our desires arise spontaneously without decision or deliberation. It is a pervasive part of life that desires, as we say, strike like a bolt out of the blue. These desires exert an influence on the will, inclining us to act in certain ways: they shape and guide choice. Uma, for example, who has been engrossed in her book all morning suddenly feels the onset of a desire to eat and begins to think about what to have for lunch. “What should I have?”, Uma ponders as she makes her way to the kitchen. “I could have some soup, or maybe I’ll fix myself a sandwich”, she thinks. Desire strikes again; just as easily as the thought crossed her mind that that some soup or a sandwich would suit her need, one of those options is cast aside when thoughts of a nice cheese sandwich occupy her mind. “Yes, that’s it. A sandwich is what I want”, she thinks.

Reflection on this mundane case motivates the thought that desires can seem to make a difference to one’s practical situation in ways that make certain responses rationally intelligible. From Uma’s perspective, her decisions and actions seem perfectly appropriate responses. For observers, desire seems to play an important role in rationalising her behaviour. Desire, it seems, contributes to an explanation of what it would be rational for one to do. Desire appear to have, as I will put it, a certain *rational significance*.

The aim of this chapter will be articulate the truth in this simple thought, and bring into view a problem that will be the overarching

concern of this thesis. The problem is that of explaining the rational significance of desires in the rationalisation of action, that is, to explain how and why desires can bear on what it would be rational to do. We can see this problem as generated by a confluence of two issues: (i) how we understand the notion of desire and (ii) how we understand the nature of rational action.

To bring the problem into view requires that we first outline the state of understanding concerning the nature of desire. In §2.1 I provide some background about our ordinary notion of desire and wanting, and relate this to a distinction between desires understood very broadly under the heading ‘pro-attitudes’ and more narrow classes of desire, ‘desires proper’. I further demarcate a category of desire that will be the main focus of this thesis, what I will call *basic* desires. I turn next to a predominant understanding of desire as constitutively connected with action, the motivational account of desire, and explain how this account is typically explicated within a functionalist framework.

§2.2 then returns to the core intuition just outlined. I develop a problem— the Problem of Conative Significance— that seems to raise a question about how to explain these natural intuitions, at least in conjunction with the standard motivational view of desire. I relate the problem developed to similar issues raised in the literature by Warren Quinn. I motivate several important constraints on the solution to the problem.

2.1 The Nature of Desire

‘Desire’ and ‘want’ are employed in natural language in a number of ways. There are a number of important differences that we can draw between these uses of ‘desire’ and the aim of this section will be to bring out these differences and systematise the difference concepts of desire expressed in order to narrow down the class of desires that I am interested in. Before we turn to these differences however, it will be important to begin with one feature of desire that comes closest to universal consensus, their intentionality.

2.1.1 Attitude and Content

Desires are psychological states that are *about* or *directed* at individuals, properties or states of affairs (Brentano, 1874). For every desire, there corresponds a *desideratum*, what it is that is wanted. I will call this the *object* or *content* of the desire. The surface form of desire or want ascriptions seems to countenance a diversity of objects such as worldly objects, act-types and propositions:

1. Tom wants/desires a hamburger.
2. Tom wants/desires to play tennis.
3. Tom desires that he is successful in his job application.

In (1), the verbs ‘want’ or ‘desire’ embed a determiner phrase, whereas in (2) and (3), they grammatically embed a non-finite infinitival clause and a complementiser in the form of a that-clause respectively. Despite these distinct surface forms, the standard view is that these ascriptions ascribe states that have objects that are uniformly *propositional* in shape. Desire ascriptions that fail to specify a propositional object fail to determine a unique desire. But in most contexts, it will be clear which desire it is that is being ascribed. For instance, (2) is usually taken to ascribe the desire that Tom wants *to eat* a hamburger, rather than *to caress*, or *to throw* a hamburger.²³ Once we have narrowed down an act-type ϕ , it is typically thought that a desire to ϕ can be captured in propositional form as a desire *that one* ϕ .²⁴ In this case, what Tom wants is that he eat a hamburger.

Obviously, many psychological states have content so we need to look to other features to help understand what is distinctive about desire. It is helpful to frame the inquiry with two central questions, focussing on the first:

²³Recently, there has been debate about whether some attitudes are irreducibly non-propositional in nature. See Ben Yami (1997), Grzankowski (2015), Montague (2007) and Thagard (2006). For an explicit defence of the propositional view more generally, see Sinhababu (2015). This thesis will largely remain neutral with respect to this issue.

²⁴It is important to note that the relevant content is very plausibly *de se* in nature. This is naturally complimented on the linguistic front where many syntactic theories analyse the non-finite clause as containing a phonologically null pronoun in subject position, subject-control PRO. For discussion, see Chierchia (1989) and Ninan (2010).

Attitude Question: what is it for a subject to be in a state of *desire* with content *C* (rather than some other attitude with the same content *C*)?

Content Question: What is it to have a desire with some content *C* (rather than a desire with a distinct content *C*)?

An answer to both these questions is part of any satisfactory final account of desire and mental states more generally. This thesis will focus more on the Attitude Question, on how to understand what is essential to desire that sets it apart from other mental states. Far less will be said about the Content Question, the formidable issue of explaining how it is that mental states acquire the content they do. Nevertheless some of the answers we will consider to the Attitude Question are naturally paired with certain views about how to settle the Content Question.

2.1.2 Pro-Attitudes and Desires Proper

Before we go any further, it is important to draw a well-known distinction between a broad and a narrow sense of ‘desire’ or ‘want’.²⁵ Discussing a point famously made in Nagel (1970, p. 29), Schuler writes that:

[T]here is a fault line that cuts across uses of the term ‘desire’. It is the one created when one asks oneself whether it is possible for someone to intentionally do something that he or she has no desire to do. In one sense of ‘desire’, this is simply not possible. (Schueler, 1995, p. 29)

An example will help bring this point out. Suppose I have to give myself a painful insulin injection. I know if I do not do this, there will be terrible health consequences. On this broad sense of ‘desire’, if I inject myself, despite my deeply loathing to do so, it will be true that I *wanted* to inject myself just because the resulting bodily movement is a particular

²⁵Here I focus on the distinction as discussed in Nagel (1970, p. 29) and following him, Schueler (1995) and Schapiro (2014). Related distinctions are littered through the philosophical literature, where a broad sense of ‘desire’ is contrasted with some more narrowly circumscribed one. For example, Davis (1984) contrasts ‘volitive’ with ‘appetitive’ desires and Daveney (1961) contrasts ‘intentional’ wanting with ‘inclinational’ wanting. Compare also the list of items that fall under the category of ‘pro-attitude’ listed in Davidson (1963) which include desires.

form of *self*-movement: it is an *intentional action*. As Schapiro puts it, ‘desire’ in this broad or ‘placeholder’ sense is a “dummy concept” since it “simply takes the place of an explanation of how the agent, rather than something external to the agent, is its source” (Schapiro, 2014, p. 136).

Understood broadly then, a desire, as a matter of conceptual necessity, is always implicated in cases where one intentionally pursues a course of action or is motivated to do so. Notice that to ascribe someone a desire in this broad sense is not to ascribe any more determinate mental state than one that motivates the subject to act; it does not distinguish for instance between whether what motivated the subject was an emotion, an urge, an intention, a character trait or an otherwise more narrow notion of ‘desire’. To give an account of the sense in which I ‘want’ to inject myself, is simply to explain why the injecting event was an *intentional action*, rather than something that merely happened to me or ‘through’ me bypassing my agency, e.g. if my body was hijacked by a mad scientist in a sci-fi scenario, a reflex reaction and so on

We can contrast now a narrower sense of ‘desire’ on which it does *not* follow from the fact that some agent ϕ -ed that she *wanted* to ϕ . Schueler (1995) calls these ‘desires proper’. This can be discerned in the insulin case just mentioned. There is a sense of ‘want’ that we can intuitively latch onto in which, given my deep aversion to not receive injections, it is true that I injected myself despite really not *wanting* to.²⁶ It is this more narrow notion of desire that I am primarily interested in for the purposes of this thesis. Before we consider in more detail some of the important ways of demarcating this narrower sub-class of desire, I want to take the liberty to introduce a terminological stipulation. From now on, I will use the term ‘pro-attitude’ to refer to all and only the states that fall under the broad use of ‘desire’. I will reserve the term ‘desire’ to refer to all and only desires (proper) unless explicitly qualified.

²⁶One possible position here is to claim that the notion of a desire is such that it is never possible that one be motivated to ϕ without a corresponding desire to ϕ . This would amount to a rejection of the distinction between pro-attitudes and desires. However even Smith, an ardent supporter of a motivational account of desire, does not deny that there is a coherent notion of a desire that is more circumscribed than the very broad notion of a pro-attitude. For instance, Smith (1987, p. 55) at least entertains the possibility that other states such as wishes, hopes, emotions etc. might count as motivating states and so a ‘pro-attitude’ despite not being desires.

2.1.3 Functionalism About Desire

Having just outlined the distinction between pro-attitudes and desire proper, I want to consider now how these concepts can be understood within a functionalist framework. Very generally, functionalism is the view that mental states are functional kinds, states with a certain “functional organisation” (Putnam, 1960). The functionalist framework provides the materials for answering the central questions I posed at the start of this chapter, the Attitude and Content Question. It is important to emphasise at the start that we can conceptually distinguish functionalism as a thesis about *mental state individuation*, the view that a mental state is of a particular kind in virtue of its functional role, from functionalism as a thesis about *mental content*, the view that mental contents are determined by their functional roles (e.g. conceptual/inferential role semantics, cf. Greenberg and Harman, 2009). It is important to emphasise that it is conceptually possible to endorse a form of functionalism about what it is for a mental state to be a desire, whilst simultaneously endorsing a non-functionalist (e.g. representationalist) account of content. If however we treat mental states as having their content essentially, then these two aspects will have to be solved together (Piccinini, 2004, p. 375). Here I will primarily focus on the functionalist analysis of desire, the kind of attitude that desire is, backgrounding the issue about how the contents of those attitudes are simultaneously grounded within the functionalist analysis. Of course, as we will see, many of the functionalist accounts of desire we will consider often aim to tackle the second question too, so by considering these accounts, we will get a sense of how functionalism can avail us of a theory of content.

The most common functionalist theories hold that what it is for a state to have a certain functional role is for that state to play a certain *causal role*, where a causal role is specified by a set of causal relations between input states, output states (e.g. behaviour) and other internal states. Importantly, the nature of the *occupant* of the relevant functional role does not enter into the analysis of the mental state. In this way, the abstractly defined causal roles identified can be multiply realised by occupants of very different kinds provided that they can stand in the

requisite causal relations.²⁷ I will consider this simple form of functionalism to begin with, though in Chapter 3, I will introduce an alternative teleological way to individuate functional roles (cf. Sober, 1985).

The version of functionalism I will consider here, *analytic* functionalism, uses *a priori* information to identify the functional role of various mental states. This information about the causes and effects of a given mental state M is one taken to articulate the meaning of ' M '. On what is known as the 'Ramsey-Lewis' method, one begins by taking all the common sense judgements about M as amounting to a psychological *theory* over which we can construct a Ramsey sentence by replacing the relevant theoretical terms with variables bound by an existential quantifier (cf. Block, 1978, pp. 269–71; Lewis, 1994; Braddon-Mitchell and Jackson, 2006, Ch.3). Applying property abstraction to the Ramsey sentence, we can yield a property consisting in a set of causes and effects without reference to mental state terms, what Block (1978, pp. 269–71) calls the relevant 'Ramsey functional correlate' for a given mental state M . According to the analytic functionalist then, to be in a mental state, say a certain desire, is to instantiate the property expressed by the Ramsey functional correlate for that mental state.

2.1.4 Functionalism, Pro-Attitudes and Desire (Proper)

In providing a functionalist analysis of desire, we have to be clear whether or not we are simply aiming to provide an analysis of, as I distinguished, the idea of a pro-attitude more generally, or desire (proper). I want to begin first by discussing the kind of analysis that Michael Smith and Robert Stalnaker provides for desire, which they understand in the broad way synonymous with 'pro-attitude'.

This broad sense of 'desire' is taken up in a particular way given certain commitments about the nature of action. Take for instance the notion of 'desire' operative in the work of Michael Smith who understands the notion of 'desire' as intertwined with our understanding of intentional action.²⁸ Smith endorses a causal theory of action on which intentional

²⁷This form of functionalism, perhaps the most well-known, is a form of 'role' rather than 'realiser' or 'filler' functionalism, compare for instance Lewis (1980).

²⁸Here I describe Smith, as he himself does, as claiming that desires lay at the causal origin of all intentional action. Smith also writes interchangeably of desire as

actions are bodily movements caused in ‘the right kind of way’ by certain psychological antecedents. On his view, what distinguishes intentional action is that it is behaviour *controlled* by the subject in a certain way. This is what separates intentional action from mere bodily movements.²⁹ This is where the notion of a desire and a belief enters into the picture. As Smith understands it, belief and desire, in their essence, constitute the basic elements that form the “foundation of *agential control*”. As Smith understand it, belief and desires are behavioural dispositions with inter-defined functional roles which rationalise and explain the behaviour that constitute the manifestation of those dispositions.^{30,31} An influential skeletal statement of just such a functional role that Smith takes up for desire and belief is due to Robert Stalnaker who writes that:

To desire that p is to be disposed to act in ways that would tend to bring it about that p in a world in which one’s beliefs, whatever they are, were true. To believe that p is to be disposed to act in ways that would tend to satisfy one’s desires, whatever they are, in a world in which p (together with one’s other beliefs) were true (Stalnaker, 1984, p. 15).

The reason why beliefs and desires constitute the seat of agential control according to Smith is that they are dispositions to behave that

an essential part of *goal-directed* action. The functional role for desire given nicely lends itself to the second gloss on what is essential to desire. Whether this will be equivalent obviously turns on the question concerning whether all intentional action is goal-directed action.

²⁹For a recent statement of Smith’s position, see Smith (2012a). For more discussion about causal theories of action in general, see Aguilar and Buckareff (2010).

³⁰This inter-definition of belief and desire allows the functionalist account of pro-attitudes to avoid a problem that felled its behaviourist ancestors. Take a crude ‘dispositionalist’ accounts on which a desire that p is identified with sets of dispositions to perform certain actions. Suppose Tom wants to eat. What exactly are the kinds of things Tom is disposed to do which are necessary and sufficient for Tom to have the desire to eat? We very quickly run into trouble. Is it to have a disposition to walk to the fridge? Not necessarily. A disposition to drive downtown, where the restaurants are located? Not necessarily. Why? Because the exact courses of action one is disposed to perform depends on what one *believes* will satisfy one’s desire (Geach, 1957; Loar, 1981, p. 6; Grice, 1974). We cannot say in advance of knowing what Tom believes the kinds of things Tom will do as a result of his desire to eat; only if he believes that there is food in the kitchen will he be disposed to walk there. For discussion of inter-definition of functional role and the problem of circularity, see Braddon-Mitchell and Jackson (2006, pp. 55–9).

³¹For a given action, there is some belief-desire pair which explains and rationalises the action. This, as discussed in the previous chapter (§1.3.2) is what Smith, somewhat misleadingly, calls a ‘motivating’ reason.

will be differentially sensitive to changes in the agent’s beliefs and desires (Smith, 1994; Smith, 2012a; Smith, 2012b). Any difference in a content of the belief or desire will be reflected in some difference in how the subject will be disposed to act.³² In any case where we have intentional action then, we have the joint operation of belief and desire, making Stalnaker’s formula “all but analytic” (Smith, 2012b, p. 81).

Smith differentiates between desires and beliefs by their ‘direction of fit’. How exactly to understand the idea of direction of fit is notoriously difficult, but metaphorically, it is the idea that beliefs have a certain kind of direction of fit (world-to-mind direction of fit) such that beliefs are supposed to *fit the world*, that is, represent how the world is. Desires in contrast are states with which the *world* is, in some sense, supposed to fit (mind-to-world direction of fit). Smith aims to capture this idea in the articulation of the functional role of belief and desires by reference to a difference in how these states relate to a third psychological element, perception: “a belief that p tends to go out of existence in the presence of a perception with the content not p, whereas a desire that p tends to endure, disposing the subject in that state to bring it about that p” (Smith, 1994, p. 115).³³ This idea forms the heart of Smith’s defence of a ‘Humean’ theory of motivation on which beliefs alone are insufficient for motivation (see Smith, 1994, pp. 116–25 for this argument). This, Smith

³²A note on content: on Stalnaker’s causal-pragmatic analysis, content attributions are *not* grounded in the presence of language-like mental representations but in the subject instantiating states that bear certain law-like causal relations to their environment which also rationalise behaviour. Stalnaker view is that the causal-pragmatic analysis entails a coarse-grained account of belief and desire content modelled as functions from possible worlds to truth-values (Stalnaker, 1984, Ch. 1). For discussion, see Stanley (2010). As stated, this functional role is over-simplified since as Stalnaker notes, there is still substantial indeterminacy here as there are a whole range of different possible belief-desire assignments we can make that will be consistent with some disposition to behave a certain way. One way we can narrow the range of possible assignments is to appeal not just to the ideal *effects* of such states, but to the ideal *causes* of such states. Following Stampe (1987), Stalnaker thinks belief is an indicator state. In Chapter 5, we will consider extending this idea to the case of desire to help with some of the problems I will shortly raise for the motivational view. I set this complexity aside for now. For present purposes, I want to background this part of Stalnaker’s view from the aspect to do with the functional type-individuation of mental states and remain neutral on whether we should think of content along the lines that Stalnaker advocates, or along the lines of his opponents (cf. Field, 1978; Fodor, 1975).

³³For discussion of some of the problems that Smith’s functional analysis incurs and also the idea of ‘direction of fit’ analyses of belief and desire more generally, see Humberstone (1992), Frost (2014) and Sobel and Copp (2001).

argues, is explained by the fact that beliefs have world-to-mind direction of fit. Only in conjunction with a state with mind-to-world direction of fit can a belief motivate. Desires, Smith claims, are these general mental states marked out by their direction of fit.³⁴

It is clear why desire in the pro-attitude sense is understood as having an essential connection with motivation: it is simply understood as the most general motivating state, ‘a psychological state with which the world must fit’. However, as I have emphasised, there remains the possibility of a notion of desire on which it does not follow that if one is motivated to ϕ and intentionally ϕ -s, that one wanted to ϕ . There are a number of ways in which this possibility might be realised, here I want to focus on a division between accounts that construe desires as retaining an *essential* connection with motivation and those that do not.

2.1.5 The Motivational Account of Desire

Even once we distinguish between pro-attitudes and desire, the dominant view is that what is distinctive about desire is the role it plays in the regulation and production of action. This view explains why, as Anscombe (1957, p. 68) suggests, “the primitive sign of wanting is trying to get”. This class of accounts, we can call *motivational* accounts of desire.

According to motivational accounts of desire, desires are a subset of pro-attitudes, that is, they are essentially motivating mental states. Motivational accounts of desire can be framed within a functionalist framework and those attracted to this framework would do well to incorporate, as with the functionalist analysis of pro-attitudes, a functional role inter-defined with belief. Provided we have at least one type of motivating pro-attitude that is plausibly regarded not as desire (proper), then the notion of a desire would not collapse into the notion of a pro-attitude since it would then be possible that one be motivated to ϕ by this pro-attitude even though one would not desire to ϕ . For instance, if we take states such as hopes or certain emotions to have the power to motivate action, and if there is a notion of desire that can be distinguished from these states, then we would have a notion of a desire (proper) as essentially motivating despite the fact that it would not automatically follow

³⁴For critical discussion on this point, see Price (1989).

from the fact that someone was motivated to ϕ that she desired to ϕ , for she might have been motivated by one of these other pro-attitudes.

There are a number of ways in which we can develop an account of what is distinctive about desire whilst retaining a commitment to the idea that they are pro-attitudes, that is, essentially motivating mental states. Within the functionalist framework, we can appeal for instance to differences between the functional role of desires and these other states. For instance, we might appeal to ‘backward-looking’ differences pertaining to the kinds of *causes* that set desire apart from other states (cf. Stampe, 1987). Even within the class of desire, we can for instance draw distinctions between *kinds* of desires. For instance, we might think that *appetitive* desires are those characterised by a certain set of causes connected with a certain kind of input from sub-personal detectors that monitor homeostatic processes. Or we might appeal to ‘forward-looking’ differences pertaining to distinctive kinds of *effects* that set desire apart, for instance, the way in which desires interact with other internal mental states such as beliefs, perceptions and/or outputs.

Once we carefully distinguish between pro-attitudes and desires, we open up conceptual space for accounts of desire that *deny* that grounding some disposition to engage in action is necessary for a state to count as desire. Some philosophers have argued against the necessity and sufficiency of having the kind of motivational role of the kind sketched by Smith and Stalnaker for desire. As we will go on to see, these other accounts centralise the role of desire as dispositions to undergo certain kinds of positive affect states such as pleasure (Strawson, 1994), and others have emphasised the role of desire in reward-based learning (Schroeder, 2004). This again can be very neatly handled within the functionalist framework with some tweaking. It is important to emphasise at this point that we do not necessarily have to choose between motivational accounts at the expense of an account of desire that also incorporates the role of these other aspects such as affect. It would be useful then to draw a distinction between (i) *pure* motivational accounts on which the certain kind motivational role sketched by Smith and Stalnaker above is *necessary* and *sufficient* for desire from (ii) *impure* accounts which allow that though motivation is a necessary part of the functional role of desire,

other elements are typically required for sufficiency.³⁵

To summarise the results so far, I have distinguished between the notion of a pro-attitude and a desire (proper) and identified the central line of inquiry as focussed centrally on desire (proper) (§2.1.2). I explained how both notions are analysed within a functionalist analysis of mental states and explained, in particular, how the notion of a desire proper can be distinguished within this functionalist framework (§2.1.4). I then moved on to outlining the motivational account of desire, according to which desires are essentially states connected, in tandem with beliefs, with states of motivation (§2.1.5).

2.1.6 Basic Desires

Before we proceed to discuss the central problem I will be concerned with in this thesis, I want to end with one final section on desire, demarcating a specific class of desires that I will be primarily interested in. The focus of this thesis will be on a sub-class of desires proper that I call *basic desires*. I want to begin by providing some examples of basic desires before considering some marks that these desires have that we can use to demarcate the category of basic desire. In the course of clarifying the notion of a basic desire, we will consider a related distinction between ‘motivated’ and ‘unmotivated’ desires.

The paradigmatic case of basic desire are those related to bodily needs. Take for instance, the appetitive desire to eat or drink, or a desire to return from the cold into the warmth of a shelter. Closely related to this class of appetitive desires, are desires for things that address bodily needs though are more specific, for instance, the desire for foods high in sugar or salt, or foods with particular flavours or appearances. There are other basic desires that are not necessarily tied to bodily need, for instance desires for sex or social contact. These examples all seem connected by having objects that serve some direct biological function,

³⁵A further possibility is that the notion of a desire is a ‘cluster’ concept such that it is not possible to articulate any particular cause or effect of a state that is necessary or sufficient. Goldman (2017) for instance defends a form of functionalism which identifies a cluster of criterial properties, where “the properties or states in the cluster intervene between sensory inputs and behavioural outputs” (Goldman, 2017, p. 336) but importantly, the various “criterial effects are variable” (Goldman, 2017, p. 336). See Schroeder (2004, Ch.6) for a discussion of what he calls ‘messy’ theories of desire, in particular a version he calls the ‘mix-and-match’ theory.

however this need not be the case. Over the course of a lifetime, humans are capable of developing a range of idiosyncratic desires. For instance, some humans have basic desires to accumulate money, to consume drugs, to play certain games, to see their favourite sports team succeed, to undergo exhilarating and scary experiences, and so on.

Reference to the objects of basic desire can only take us so far; as we will see, it is not possible to demarcate the class of basic desire simply by reference to their objects. A more promising strategy will be to consider general features or ‘marks’ of basic desires. I will divide these features into two categories. The first category concerns *forward-looking* features of basic desires, features related to the role that basic desires play in guiding action and the way they interact with other mental states that the subject might be in. Within this class of properties, we will consider the more general distinction between intrinsic desires, of which basic desires are a subclass and instrumental or realiser desires. The second category concerns *backward-looking* features of basic desire, ones to do with how basic desires come to be, the kind of psychological antecedents that give rise to basic desires. These features will be generally sketched for now simply to aid us in acquiring a minimal grasp on the kind of desire we will be interested in; as the thesis progresses, so too will our understanding of these features and the category of basic desire.

Forward-Looking Features

First, basic desires give rise to episodes of psychological activity that, in tandem with other psychological states, motivate certain kinds of actions under certain conditions. These episodes characterise *occurrent* basic desires. When a basic desire is occurrent it becomes a “*potential influencer* of action” (Bratman, 1987, §2.2). In contrast to ‘pro-attitudes’ more generally, occurrent basic desires are “motivational states that exert an influence *on* the will” rather than simply being “states *of* the will” (Schapiro, 2009, p. 230). Take one of our paradigmatic basic desires: a desire to eat sweet things. A basic desire to eat sweet things can either refer to a standing state of the subject which we might describe colloquially as having a ‘sweet-tooth’, that is a ‘taste’ that grounds a disposition to seek out sweet foods. Under certain conditions, such as

when one is hungry, this standing state gives rise to a period in which one is motivated to acquire something sweet to eat, an *occurrent* desire to eat something sweet.

Of course, we need not act on occurrent basic desires. For instance, we might have a desire for outcomes which we do not believe we have a good chance of bringing about, desires that we believe we should not even try to satisfy (e.g. morally bad desires) or desires that are massively outweighed by other desires. Nevertheless, even when we do not act on them, as when we exercise self-control, they constitute one of the fundamental psychological elements that the process of practical reasoning operates on in determining what to do. As I will go on to develop, we can think of occurrent basic desires as constituting, as Schafer (2013, p. 264) puts it, the “standard inputs” into the process of determining what to do.

Second, basic desires when occurrent can persist despite rational cognitive changes, for instance the acquisition of new normative beliefs. They have the power to be recalcitrant to rational cognition. In this way, basic desires differ from what philosophers call *judgement* or *reasons* sensitive attitudes. Scanlon defines a *judgement-sensitive* attitude in terms of backward and forward-looking features as:

[An attitude] that an ideally rational person would come to have whenever that person judged there to be sufficient reasons for them and that would, in an ideally rational person, “extinguish” when that person judged them not to be supported by reasons of the appropriate kind. (Scanlon, 1998, p. 20)

In a similar vein, Richard Holton identifies what he calls ‘cognitive desires’:

In many cases, a desire is bound up with a reason or a justification: to want something is to want it for some reason. As one’s confidence in the reason diminishes, so the desire diminishes. (Dill and Holton, 2014, p. 4)

If we focus for now on just the forward-looking aspect, we can say that basic desires are not judgement-sensitive in the sense that they can persist despite the acquisition of beliefs about reasons that count against pursuing the course of action one has an occurrent basic desire to pursue.

The basic desires we have considered, such as a desire to eat something sweet, does not seem to to be judgement-sensitive in this way.³⁶

This point can be made with respect to both standing basic desires and their occurrent manifestations. For instance, a standing desire for peanuts can persist even if one comes to learn that one has a fatal peanut allergy. Moreover, when this desire becomes occurrent, one can be inclined to eat some peanuts despite this fact being salient to one. This point is brought out more vividly by certain kinds of desires for addictive substances like nicotine. One might know that there are very strong reasons not to smoke yet still be disposed to feel strong occurrent basic desires to smoke. The mere fact that one judges that one has no reason to eat peanuts or to smoke cannot immediately and directly ‘extinguish’ the occurrent desire or the standing psychological structures that give rise to those occurrent desires.³⁷

It is important as well to foreground a distinction often made between intrinsic and instrumental or realiser desires. Basic desires are a class of *intrinsic* desires. As standardly formulated, intrinsic desires are desires where the objects of the desire are wanted “for their own sake”. In contrast, *instrumental* desires are desires where the objects of the desire are wanted as a means to some distinct end.

It is important to highlight that it is possible that a subject possess both an intrinsic and an instrumental desire for one and the same *desideratum*.³⁸ For example, Uma might have an intrinsic desire to eat the delicious dish (her favourite) that her partner spent the evening slav-

³⁶This claim should not be confused with the thought that such desires do not have the function to track reasons for action. As I will go on to argue in Chapter 5, I hold that there is a way in which desires can be defective with respect to their proper function, as when they are systematically formed and maintained in a way that produces motivations for outcomes that are bad for the subject.

³⁷It might be objected that some basic desires *are* judgement-sensitive. Take a paradigm basic desire: an appetitive desire to eat some delicious looking slice of cake. If one were told that the cake in question was poisonous, one probably would not still want to eat the cake. Would such a desire then count as *non-basic* in the sense I am trying to articulate? This problem emerges due to a lack of clarity concerning the desires in question. In the cake case, we might imagine that the desire for that specific cake is non-basic but a *realiser* desire generated by a basic desire to eat cake (to be discussed shortly). But underlying that is a desire to eat cake (or more generally, the appetitive desire to eat) which can still persist even once one drops the desire for that specific cake. So it is important to distinguish the nature of the desire which I described as paradigmatically basic.

³⁸This point is made in Arpaly and Schroeder (2013, p. 7) and Schroeder (2004).

ing over the stove to prepare but also an instrumental desire to eat the meal to please her partner. We see then that we cannot identify whether a desire is intrinsic or instrumental simply by looking to the object of the desire. In tandem with the subject's beliefs, intrinsic desires typically give rise to instrumental desires for the means believed to conduce to the end intrinsically desired. For example, if Uma believes that she needs to lay out some cutlery in order to eat the delicious meal, she will form the instrumental desire to lay out some cutlery.

An important further category is *realiser* desires which are desires where the object is wanted because it realises or constitutes an end that is intrinsically desired. To illustrate, if Uma intrinsically desires the good health of her partner, she will also form desires for certain outcomes such as that he avoid catching the flu. Here not catching the flu is not a means to securing her partner's welfare but realises the outcome she intrinsically desires (Arpaly and Schroeder, 2013, pp. 7–8). Realiser desires such as these are also typically generated by intrinsic desires in tandem with the subject's beliefs about what states of affairs are realisations of other states of affairs.

The distinction between (i) judgement-sensitive vs non-judgement sensitive desires and (i) intrinsic vs. instrumental vs. realiser desires cuts across each other. Instrumental and realiser desires display judgement sensitivity. For instance, Uma's instrumental desire to lay out cutlery is motivated by an intrinsic desire to eat her partner's dish and a belief that she must do the former in order to do the latter. However, if, for whatever reason, she comes to learn that she does not need to do this (e.g. because this version of this dish must be eaten in the traditional way with your hands), then this would eliminate the instrumental desire to lay out cutlery. Likewise with realiser desires: suppose under the false belief that any bacteria in your gut is bad for your health, Uma forms the realiser desire that her partner have no bacteria in his gut (example from Arpaly and Schroeder, 2013). Were she to learn that certain gut bacteria is part of a healthy digestive system, this desire would similarly be eliminated.

In contrast, intrinsic desires are normally judgement-*insensitive*. Many of the examples of basic desires I have given such as appetitive desires for certain kinds of foods can often persist in spite of the acquisition of be-

liefs about the merits of pursuing the intrinsically desired end. As Arpaly and Schroeder (2013, p. 10) puts, it intrinsic desires have something of a “life of their own”.

Backward-Looking Features

To frame some of the backward-looking features of basic desires, I will consider a well-known distinction between ‘motivated’ and ‘unmotivated’ desires.³⁹ Thomas Nagel, who introduced the distinction, articulated it as follows:

It has been pointed out before that many desires, like many beliefs, are *arrived at* by decision and after deliberation. They need not simply assail us, though there are certain desires that do, like the appetites and in certain cases the emotions. The same is true of beliefs, for often, as when we simply perceive something, we acquire a belief without arriving at it by decision. The desires which simply come to us are unmotivated though they can be explained. Hunger is produced by lack of food, but it is not motivated thereby. A desire to shop for groceries, is on the other hand motivated by hunger. Rational or motivational explanation is just as much in order for that desire as for the action itself. (Nagel, 1970, p. 29)

Whilst basic desires share many of the features that typify unmotivated desires, contrasting with the category of motivated desires, these features need to be carefully handled in our understanding of basic desires. Here Nagel articulates three important features that allow us to demarcate between desires that are:

- (i) “arrived at by decision and after deliberation”,
- (ii) “simply assail us” or “simply come to us”, or
- (iii) amenable to “rational or motivational explanation”.

Let’s start with the idea that some desires “simply assail us”. As I have mentioned, under certain conditions, basic desires are active or occurrent and they incline us to pursue certain courses of action. In this way, *occurrent* basic desires simply assail us. This is one respect in which

³⁹It is made in Nagel (1970) and discussed in Schueler (1995) and Scanlon (1998).

we are typically said to be ‘passive’ with respect to our desires. However it would be inaccurate to say that standing basic desires assail us in the same way. The development of a standing desire is a process that often occurs outside of the subject’s awareness. Someone can have a standing basic desire to drink beer, a state that persists even when the subject is asleep or unconscious. Under certain conditions, such as when the subject is at the pub and is thirsty, this standing state can give rise to an occurrent desire for a pint of beer that then assails the subject unbidden.

Turning now to desires arrived at through decision or deliberation. Episodes of deliberation involve determining what normative reasons there are for one to act. Take a case in which one judges that there is decisive reason for one to ϕ and in light of this is motivated to ϕ . The fact that some desires are arrived at through decision or deliberation grounds the third feature, amenability to rational explanation. When we rationally explain the desire to ϕ , we can cite (apparent) reasons for which one forms the desire. We can sensibly ask *why* one wants to ϕ . In contrast, some of the desires that “simply assail us” are not amenable to rational explanation *because* they are not the products of deliberative processes and so not formed *for reasons*.⁴⁰

Finally, it is important to highlight again how the distinction between standing and occurrent basic desires relate to these backward-looking features to do with desires that are arrived at via deliberation or whether

⁴⁰Hunger, Nagel claims, is not like this: desires are not the products of deliberation and are not states we enter into for reasons. We have to however have a particular understanding of hunger for this point to go through. We can think of hunger as a complex state consisting in, inter alia, (i) a range of sub-personal homeostatic processes, like gut-brain signalling, as well as elements that are apparent at the personal level such as (ii) sensations like stomach pangs, feelings of weakness, (iii) cognitive components (e.g. intrusive thought) as well as what I have called an (iv) occurrent basic desire to eat. This will be discussed further in Chapter 3. Nagel is certainly right about hunger understood as consisting in (i)-(iii). We do not, indeed cannot, enter into this state as a result of deliberation and do not have reasons *for* entering into this stage. What about (iv)? Can’t we ask *why* one has this occurrent appetitive desire to eat? To this, it might seem that we often have an answer ready “Because I’m hungry”. But we have to be careful here to conclude that the desire itself, or the sensational concomitants of hunger, constitute one’s *reason* for having the occurrent appetitive desire to eat. See Alvarez (2010, Ch.4). It is plausible that the sense of the why-question is not asking for one’s reason to want to eat, but asking for a non-reason citing *explanation* of one’s desire to eat. It is very plausible that the occurrent desire to eat is triggered by a complex interaction of sub-personal processes (see Kringelbach, 2015) and not formed by taking any of the elements in (ii) or (iii) as a reason. When triggered in this way such desires qualify as *appetitive* desires to eat.

they are amenable to rational explanation. One complexity concerns the fact that some standing basic desires *can* be, in a sense, ‘arrived at’ via prior episodes of deliberation. Consider the following case. We can cultivate standing desires for foods that we are not presently disposed to seek out. Take the desire to have chocolate with high cocoa content, which might at first be too bitter to be enjoyable. In some cases of how desire cultivation occurs, it is true that, at some point in the past, one might, on reflection, have judged dark chocolate to be better. Through repeated exposure, one would start to form or strengthen a standing basic desire for dark chocolate. We can imagine that the sight of dark chocolate now sparks an occurrent desire for dark chocolate and that tasting the complex bitter flavours now gives rise to pleasure.

There are a few things we can say. Notice that the decision to instil a desire for dark chocolate and the simple form of deliberation that precedes the decision does not, in some sense, ‘directly’ give rise to the standing desire. In order to inculcate a desire, one needs to work to develop the desire through repeated association, pairing, in this case, consuming the dark chocolate with a certain kind of reward. This serves to contrast such a desire with the kind of desire with an existence that is, as Holton puts it, “bound up” with an apparent justification or reason. For example, consider Tom who decides that he wants to learn more about cinema. As a prank, Tom’s friend tells him that *The Room* is the pinnacle of cinematic achievement. With no reason to question the recommendation, Tom’s belief that *The Room* is an excellent film directly gives rise to, or perhaps even constitutes— if some forms of ‘pure cognitivism’ about motivation is right— a desire to see *The Room*. If Tom were now to learn that *The Room* is a terrible film, such a desire would immediately vanish, without residue, so to speak. This gives us a more refined understanding of this backward-looking aspect of basic desire. It is not that such desires, as a matter of conceptual necessity, *never* feature deliberation or decision in the list of causal antecedents to such desires, rather it is that they are not normally the ‘direct’ result of deliberation or decision.

I want to end by discussing the distinction between intrinsic and basic desires, in particular with respect to this backward-looking aspect. It might be thought at this point that intrinsic desires are roughly co-

extensive with what I have been calling basic desires in sharing many of the forward-looking aspects. One possible difference for which I think it is useful to introduce the notion of basic desire comes from the possibility that there can be intrinsic desires that are generated directly by decision or deliberation. What distinguishes basic desires from some intrinsic desires then is that basic desires are never the direct product of rational deliberation. An example of this might be a desire we form to support a certain sports team. Suppose you have to choose between two teams to support at a forthcoming match. Both teams are meritorious: Porto FC is a highly skilled and talented group, whilst Santa Clara FC, clearly inferior in skill nevertheless demonstrates a great deal of passion and perseverance. It is possible I think simply to decide to support Santa Clara FC and in so doing establish a desire for Santa Clara FC to win. In this case, the desire is intrinsic; for you want Santa Clara FC to win for its own sake (contrast for instance, a desire to support Santa Clara FC because you've placed a large bet on their winning). However such a desire is quite unlike the desire for dark chocolate that you cannot simply install through decision or deliberation but have to work to inculcate.

These sets of forward-looking and backward-looking features give us a reasonable articulation of some key features of basic desires, enough to move on to consider the main problem of this thesis. Ultimately, I think it is possible also to provide an explanation of these features of desire *architecturally*. As I will go on to develop, basic desires are those that are produced by dedicated sub-personal mechanisms that are distinct and insulated from the sub-personal mechanisms that underlie the production of other states like perceptions and beliefs and have their own 'rules' for producing novel basic desires.

2.2 The Problem of Conative Significance

In this section, I want to bring into view a problem with reconciling the intuition briefly alluded to in the introduction of this chapter concerning the seeming rational significance of desire with the account of desire discussed in the previous section. This section has three parts.

I will begin by clarifying the notion of rational support which will draw on material discussed in Chapter 1 (§2.2.1). Having outlined the

notion of rational support, I will motivate the claim that in many cases, desires seem to have the power to rationally support action and the formation of intentions, substantiating the intuition of the rational significance of desire (§2.2.2). In the following sections (§2.2.3), I develop the puzzle that will structure forthcoming chapters of this thesis. The problem I develop, The Problem of Conative Significance, is one vindicating the intuition of rational significance with a plausible understanding of the nature of desire. At first the target of the problem will be posed narrowly against the dominant account of desire discussed in the previous section, the pure motivational account of desire. After considering the problems that the motivational view faced, I generalise the problem and explain two constraints on answers to the Problem of Conative Significance.

2.2.1 Rational Significance

As discussed in Chapter 1, rationality is concerned with standards or norms that apply to transitions between, or sets of, mental states. Some of these standards or norms pertain to our apparent reasons; in these cases, rationality requires that we respond in certain ways to our apparent reasons for action. I want to begin by distinguishing between two kinds of rational norms that can apply to a subject in virtue of having certain attitudes.

Some attitudes, in virtue of their nature and content, introduce rational constraints on an agent's psychology. To illustrate, suppose there is an academic talk occurring later which would be interesting to attend, and your doing so does not conflict with any prior intention. Consequently, you form the intention to attend. In virtue of forming this intention, there is now a rational constraint on what other mental states you can then form. For instance, suppose on the way to the conference you spot a hill and it strikes you that it would be rather fun to do a roly poly down the hill, something you know will leave you muddy. Intending to make it to the conference which requires you look presentable means that you cannot, on pain of being less than ideally rational, form the intention to roll down the hill without first giving up your intention to arrive at the conference presentably. The rational constraint that you are subject to in virtue of your intending to make the meeting presentably

takes the form of a ‘wide-scope’ requirement: rationality requires you not to intend to ϕ and intend to ψ (where it would be impossible to ϕ and ψ). The rational norm you violate is ‘wide-scope’ in the sense that what is required is that *either* (i) you give up your intention to roll down the hill (if you intend to go to the conference) or (ii) you give up your intention to attend the conference presentably (if you intend to roll down the hill). You need to decide now which activity matters to you more: rolling down the hill or making the conference.

Some attitudes, in virtue of their nature and content, contribute rational support for various psychological transitions, for instance, the formation of certain other mental states. For example, suppose you perceive that p , absent any countervailing belief about the reliability of your perceptual faculties, it is rational for you to form the belief that p .⁴¹ Another example: suppose a sharp pain shoots up your leg. The mere presence of this unpleasant sensation makes it rational for you to do something to alleviate the pain, absent countervailing reasons you have not to do certain things (e.g. if taking pressure off your leg would actually bring about more pain). (Here we are explicitly considering, not instinctive reactions, but actions done in deliberate response to the sensation). The relevant mental states here contributes to the obtaining of a certain *narrow-scope* requirement: if you are in M_1 and given certain other conditions (e.g. there are no defeating conditions of which you are aware, etc.), rationality requires you to form M_2 . They are narrow-scope in the sense that they recommend particular responses. Failure to form the relevant mental state constitutes evidence of irrationality.

To sum, the contrast between wide-scope and narrow-scope norms show two ways that various mental states can be rationally significant.⁴² One way for a mental state M to be rationally significant is by being an attitude of a certain type with a certain content for which there is a covering wide-scope norm which then constrains what other mental states you can possess or form whilst remaining rational. Another way for a mental state M to be rationally significant is by being an attitude of a certain type with a certain content for which there is a covering

⁴¹It is important to emphasise the qualification that there is an absence of apparent reasons to doubt one’s senses; one would not be irrational to withhold a belief that p were this so.

⁴²For a related distinction, see Archer (2013).

narrow-scope norm which requires you to respond in a certain way.

I want to briefly connect this thought now to the previous discussion about reasons, specifically the latter idea that some mental states have a nature such that they are capable of contributing to the existence of a narrow-scope rational requirement. The fact that some mental states can contribute to the obtaining of a narrow-scope rational requirement is due to the fact that they are capable of ‘providing’ reasons. Take for instance, perception. It is highly controversial how we understand the reason giving role of perception. On some views of perceptual reasons, we should understand the reason that perception provides is simply the fact that p that one perceives, the fact which forms the basis of one’s belief that p (Schnee, 2016). Others think of one’s seeing that p as constituting the relevant perceptual reason to believe that p (McDowell, 1995). Alternatively, some think that a perception that p consists in a kind of seeming that can non-inferentially justify a belief that p (Huemer, 2006).⁴³ Nevertheless, what I want to call attention to is just the idea that these mental states, in some way, provide an apparent reason for a certain response either by constituting that reason of which one can be aware of as such on reflection, or by presenting a *fact* which constitutes the apparent reason for belief.

2.2.2 The Rational Significance of Desire: The Prima Facie Case

In this section, I want to make a preliminary case for the claim that desire is significant to practical rationality in the sense that desires can contribute rational support for certain courses of action, or for the adoption of certain intentions.

Recall the case of Uma discussed in the beginning of this chapter. At a certain point in time, Uma spontaneously forms an occurrent basic desire to eat and this desire rationally supports her pursuing certain courses of action. The desire emerges in a situation in which she has no competing pressures. When the occurrent desire assails her, her attention is drawn to the salient ways for her to satisfy this desire. In this case,

⁴³Some e.g. Burge (2003) who have a particular understanding of reasons, might want to resist the thought that perception providing reasons as opposed to grounding an entitlement for the corresponding perceptual belief.

it rationalises her decision to head to the kitchen. Reflection on this simple case provides *prima facie* reason to accept the following intuition of rational support:

Rational Support: *A*'s desiring to ϕ provides some degree of rational support for *A* to pursue, or intend to pursue, the end of ϕ -ing, absent countervailing judgements, competing desires, etc.

It is possible to see this intuition in clearer detail by considering how this case proceeds. As Uma reflects on her options, a more specific desire emerges: she has an occurrent desire for the sandwich as opposed to the soup, both of which she knows to be perfectly fine options. When this second desire emerges, it rationally supports a yet more specific course of action. Absent countervailing reasons she has to act against her desire for the sandwich option, her desire for the sandwich rationally supports a corresponding intention to have a sandwich rather than soup for lunch.

The reason why this second aspect of the case is interesting is that it seems to show that desires have some independent power to rationally support a choice between two rationally eligible options (cf. Chang, 2004). Once Uma's desire for one option over the other is on the scene, her practical situation is changed. It is not just that it would be rational for Uma to pursue the sandwich option, but that pursuing the soup option over the sandwich option seems, at first glance, an *irrational* course of action to pursue. Imagine for instance if Uma were to cast aside her desire and employ a random decision procedure such as flipping a coin. To do so would not just be perverse or unintelligible, but positively irrational.⁴⁴

Many philosophers have taken such cases to provide *prima facie* grounds to think desire is rationally significant (cf. Chang, 2004, Railton, 2012, Stampe, 1987, Stampe, 1986, Schafer, 2013). If this is right, then desires like these which pervade human life constitute an important part of human practical rationality. Nevertheless, it is of course possible to contest the rational significance of desire. For those who take desires to be rationally significant, the task set for us now will be to try to substantiate and explain *how* desires could be rationally significant in the

⁴⁴This specific point about the comparison between desires and coin flips is discussed in Chang (2004, pp. 83–5), Tanyi (2011) and Archer (2013, Ch.6).

way they seem to be.

2.2.3 The Problem, Narrow and General

In this final section, I want to consider the pure motivational view outlined in §2.1.5. We will discuss an argument famously advanced by Warren Quinn who argues that the pure motivational view of desire fails to explain the power of desire to rationalise choice and action. Through an examination of Quinn’s negative argument, I motivate two constraints on answers to the Problem of Conative Significance.

To throw this issue into stark relief, I turn to a well known example due to Warren Quinn. Quinn was concerned with the question of “whether pro- and con-attitudes conceived as functional states that dispose us to act have any power to rationalise those acts” (Quinn, 1994b, p. 32).⁴⁵ The lynchpin of Quinn (1994b) is an appeal to his famous ‘Radioman’ case, which he describes as follows:⁴⁶

Suppose I am in a strange functional state that disposes me to turn on radios that I see to be turned off. Given the perception that a radio in my vicinity is off, I try, all other things being equal, to get it turned on. Does this state rationalize my choices? Told nothing more than this, one may certainly doubt that it does. But in the case I am imagining, this is all there is to the state. I do not turn the radios on in order to hear music or get news. It is not that I have an inordinate appetite for entertainment or information. Indeed, I do not turn them on in order to hear anything. My disposition is, I am supposing, basic rather than instrumental. (Quinn, 1994b, pp.189-90)

I want to begin by clarifying the kind of state that we are to suppose Radioman is in. We are not to imagine the kind of forces to which Radioman is subjected to in virtue of this peculiar motivational state

⁴⁵Quinn’s purpose in appealing to this case was to motivate a view about rational agency according to which the agent must be responding to an evaluation of her desired ends as good, attractive or pleasant in some respect. I will set Quinn’s own views aside until Chapter 4; for now, I focus simply on establishing the insufficiency of the Pure Motivational view of desire to vindicate *Rational Support*.

⁴⁶As a testament to how influential Quinn’s case is, almost all philosophers who press similar arguments (e.g. Scanlon, 1998; Boyle, 2016; Boswell, 2018) rely on Quinn’s case. Much of this line of reasoning is echoed in Scanlon (1998) in the vocabulary of reasons.

along the lines of a reflex. We can suppose that the state that motivates him has the kind of functional role constitutive of the pure motivational view. The relevant motivational state here is one that possesses content and that motivates him to do what he believes will allow him to turn on radios by generating instrumental desires, e.g. to head to the electronics shop where he knows radios are located. As Quinn notes, like desires, they interface with his perceptual states and beliefs to conduce to action. Moreover, we can also allow that Radioman is able to exercise the capacity to weigh up this motivational states against other desires.⁴⁷ In this way, as Smith emphasises, this motivational state forms part of “a familiar capacity for control, the kind of control characteristic of those that people exercise when they act intentionally” (Smith, 2012a, p. 394)

There is an aspect of the motivational state that is underspecified by Quinn’s description of the case. Even granting the previous point that this motivational state is unlike a reflex or a twitch in standing in relations with Radioman’s beliefs and other desires to drive action, it remains unclear whether we are to imagine the relevant state as one which simply moves Radioman, bypassing his agency and capacity for choice altogether or not. On one interpretation, the motivational state can be understood as a kind of ‘intelligent reflex’, in the sense that when the appropriate conditions obtain (e.g. Radioman sees a radio he can turn on, and has no competing desire), his body begins to move. We might call this, to borrow a phrase from Wallace (1999), the ‘hydraulic’ interpretation of Radioman. On another interpretation, we are to understand the motivational state as having some phenomenological presence as a felt urge or impulse in response to which the subject is able to exercise an element of choice. It is important to stress how minimal this urge must be construed if not to undermine Quinn’s case: we are not even to imagine this urge as unpleasant in nature, for then we might begin to understand how it can rationalise doing what will relieve the unpleasant feeling.

Whichever way we interpret Quinn’s case, it is clear that Radioman is a subject who satisfies all the conditions imposed by the pure motiva-

⁴⁷As Smith notes “whether [Radioman’s] urge has any effect at all will depend on . . . whether he has stronger contrary dispositions: imagine telling him that if he turns on a radio you will shoot him in the head, so bringing his urge into conflict with his desire to preserve his life” (Smith, 2012b, p. 394).

tional view to count as having an occurrent basic desire. The problem that Quinn identified is that it is doubtful to say the least that a quasi-hydraulic force or a blank urge to turn on radios can rationally support or ‘rationalise’ a corresponding action or intention. There are, I think, two aspects to this claim that we can extract.

Brute Impulses and Rational Support

The first respect in which Radioman’s inclinations fail to rationalise his actions is that it fails to explain how the brute motivational states that Radioman possesses could *rationally support* pursuing the course of action he is disposed to bring about in the sense articulated in the previous section of this chapter.

We can highlight this point by contrasting the intuition canvassed earlier in the case of Uma. Intuitively, it seems that in having a desire for one lunch option over the other, it would be irrational for her to ignore the desire. Having a desire for one option over the other seems to add something of real normative significance which can inform choice and action, favouring a specific course of action. This is, at least in part, why it would be less than ideally rational for her, in this particular circumstance, to disregard the desire for no countervailing reason. In contrast, being subject to a brute ‘hydraulic’ force or assailed by a forceful impulse adds nothing of normative significance, nothing in the way of an apparent reason for Radioman to turn on radios. If we assume that Radioman’s impulse to turn on radios could be disregarded, or resisted as much as is possible, it is unclear why Radioman would be acting in a way that was less than ideally rational.

Brute Impulses and Intelligibility

The second respect in which Radioman’s ‘desire’ fails to rationalise his actions emphasises the rational intelligibility of his actions from Radioman’s perspective. This element is one that is emphasised by Quinn. An important class of rationalising explanations are *reasons*-citing action explanations. An agent’s (apparent) reason for action affords mutual and self-intelligibility of a course of action. To bring out the issue that the pure motivational view has with respect to the rational intelligibility of

action, let us consider the kinds of apparent reasons that we take to be operative in Radioman's case.

Let us start with a case in which Radioman is engaged in an action which is instrumental to his bringing about that he turns on radios. What is Radioman's reason for doing this? In this case, there is an instrumental reason that can be cited, namely that walking to the shop will allow him to turn on a radio. We have not however entirely rendered Radioman's actions intelligible, for the end of turning on radios is just as much indeed of rationalisation. Now suppose we ask Radioman why he is motivated to turn on radios in the first place. It is part of the case as described that Radioman's desires are intrinsic; he does not turn it on because he likes the sound of the switch, or that it gives him relief from his compulsion, etc. It is unclear how a brute urge to turn on radios avails Radioman of any answer to this question other than "I just do" or "I just want to".

There is an obvious parallel here to draw with Anscombe's discussion of desirability characterisations. Anscombe contends that to the extent that we can intelligibly understand a subject as being moved by a *desire* to ϕ , it should be possible that the subject provide an answer to the question *why* she wants to ϕ that avails us of a desirability characterisation of ϕ -ing as good or desirable. Here a desirability characterisation of a subject's desire to ϕ is an answer to the question of why the subject wants to ϕ , one that appeals to features of the object of desire, that renders the object of the desire attractive or desirable to the subject.⁴⁸ Desirability characterisations constitute reasons for which the subject desires the *desideratum*. This point applies to basic desires where *desiderata* are wanted for their own sake. Here the relevant desirability characterisation will relate the desideratum to some apparent good-making feature.

Anscombe motivates this thought with a case similar to Quinn's. Consider someone who avows a desire for a saucer of mud. When asked why, we are to imagine that we get the reply "I just do" (Anscombe, 1957, pp. 70–1). Anscombe's point is not that it is never intelligible for someone to form such a desire. After all, there are possible *instrumental* desires

⁴⁸For some more unusual objects of desire, it is insufficient simply to state that the desired outcome is good. Radioman is no more intelligible to us were he to express his belief that turning on radios is just plain good (cf. Copp and Sobel, 2002). Rather, what is needed is reference to some good-making or desirable-making feature of the object. For discussion of this point, see Setiya (2010) and Boswell (2018).

for saucers of mud (e.g. I want it to play a mean prank on someone). Nor is the point that it is impossible to have intrinsic desires for saucers of mud. Rather, it is that unless there is a desirability characterisation in the offing, she is not intelligible to us as even *wanting* the saucer of mud.

The point I want to focus on here is the idea that Radioman’s peculiar motivational state lacks features that avail him of a desirability characterisation of the outcome he possess an impulse to bring about. If we were to ask Radioman why he wants to turn on radios, we quickly bottom out at an answer that does not avail us of a desirability characterisation. Lacking this, Radioman is compelled to act by forces that fail to “make [his] acts sensible” (Quinn, 1994b, p. 237).⁴⁹ As Tamar Schapiro puts it, the desire “does not give the agent a perspective from which to see the action as worth doing, even in a minimal sense” (Schapiro, 2009, p. 235). A further problem then with the pure motivational view is that it fails to explain how desires are capable of rendering acts that they incline us to pursue intelligible; indeed, if Anscombe is right, it is hard to see Radioman as motivated by a *desire* much less rationally motivated by it.

The Problem Generalised

So far we have considered the ways in which the pure motivational view fails to explain how desires can rationally support action and the formation of intentions. This leaves us with the following question: how, in light of the failure of the pure motivational view, should we amend or

⁴⁹An exegetical note: on closer inspection of Quinn (1994b), Quinn seems to distinguish between motivational states that merely render courses of action *intelligible* as opposed to *rationalise* them. Quinn discusses the example of harmful sexual impulses such as incest or a pyromaniac’s desire to set houses alight. Quinn writes that these are pleasures which “no right-minded person... would suppose he had good reason to seek” and therefore do not *rationalise* action. Nevertheless, Quinn allows that “they make *intelligible*... a choice to pursue them” (Quinn, 1994b, p. 248). In denying the pyromaniac’s desire the capacity to rationalise, Quinn reveals commitment to a notion of rationalisation that is blocked when, as it is natural to assume in his examples, either one judges one has no good reason to ϕ (e.g. burn the barn) or that there is conclusive reason to *refrain* from ϕ -ing. This distinction is not appreciated by some commenters on Quinn’s paper, e.g. Boswell (2018). As I read Quinn, an inclination to act is rendered ‘sensible’ or ‘rationalised’ if it is made intelligible in a *specific way*; namely, in accordance with what Quinn takes to be missing from the pure motivational picture, namely as flowing from the subject’s assessment of her reasons for action.

extend our understanding of the nature of desire to vindicate the rational significance of desire? The challenge of providing an answer to this general question is the Problem of Conative Significance.

Reflection on the kinds of problems that Quinn raises for the pure motivational view suggest two constraints on answers to the Problem of Conative Significance. First, to the extent that the rational support intuition is compelling, an answer to the Problem of Conative Significance should explain how desires can rationally support action and the formation of intentions. Call this the *Rational Support Constraint*. The second constraint is that answers to the Problem of Conative Significance must explain how acting on, or forming intentions on the basis of, a basic desire is rationally intelligible from the subject's perspective. Call this the *Intelligibility Constraint*.

Conclusion

The aim of this chapter has been twofold. First, it has been to outline and clarify the notion of desire that will be focus of the thesis, basic desires. Basic desires are to be distinguished from pro-attitudes more generally in having the backward and forward-looking features I elaborated on in §2.1.6. Second, I outlined and motivated the central issue that will structure the following chapters, the Problem of Conative Significance, which consists of explaining how basic desires are capable of explaining some natural intuitions about the rational significance of desires which I motivated in §2.2.2.

The following chapter begins to tackle the Problem of Conative Significance. The chapter considers a class of views which aim to explain how desires can rationally support action by connecting desires with hedonic and affective elements. The accounts we will consider aim to explain how desires can provide reasons for action by appealing to elements like (prospective) pleasures, pains as well as non-pleasure based affective experiences.

Chapter 3

Desire, Pleasure and Affect

There is an intimate relationship between desire and affect. It is hard to deny that our ordinary experience of a great many of our basic desires are coloured by affect. It is natural to speak of the pain or discomfort of desire and yearning, the enticing, alluring way in which the objects of desire are presented to us, as well as the pleasure or joy that comes with their satisfaction. It is unsurprising then that despite the dominance of pure motivational accounts of desire, many theories of desires have been proposed that centralise, or at least, give pride of place to this aspect of desire. This chapter examines the connection between desire and affective phenomena, with a focus on clarifying how this connection can provide us with an understanding of the way in which desires can contribute to practical rationality.

The first part of this chapter considers what I will call *hedonic* accounts of the rational significance of desire (§3.1). These accounts all attempt to explain how desires contribute to what it is rational to do by centralising notions like pleasure or displeasure in our understanding of desire. I begin by clarifying different kinds of pleasure-related to psychological states and episodes (§3.1.1). I consider how such views are to be formulated and the kinds of desires for which their rational significance might plausibly be accounted for on these terms (§3.1.2-3.1.3). I will argue that hedonic explanations of the rational significance of desire face substantial problems and so calls into doubt its suitability as a general answer to the Problem of Conative Significance (§3.1.2, 3.1.4).

The second part of this chapter focusses on a different set of affective

components that are importantly distinct from pleasure and displeasure. I consider work by Ruth Chang, Declan Smithies and Jeremy Weiss that centralise a form of valenced experience in our understanding of desire (§3.2.1), which they claim explains, at least in a limited set of cases, how desires constitute apparent reasons for action (§3.2.2). I argue that such views fail to provide an adequate answer to the Problem of Conative Significance (§3.2.3). However, in the final section I consider a view proposed by Peter Railton and highlight areas in which his account is promising (§3.3).

3.1 Desire and Pleasure

The notion of an affective mental state or event includes a wide range of psychological states and events like emotions, feelings, moods, passions, pleasures, pains and so on (Mulligan and Scherer, 2012, p. 347). It is not clear if there is any single feature that is necessary and sufficient for membership in this nebulous class. I propose that we begin our exploration into the affective realm by highlighting several important sub-classes of the affective that will be especially relevant in our understanding of desire.

3.1.1 The Hedonic Realm

Let us start with the *hedonic*, psychological states or episodes that pertain to *pleasure* or what is pleasant.

Consider *sensory pleasures*; for instance, the pleasant feeling of a cooling breeze across one's skin, the delicious taste of chocolate, or the experience of orgasm. These are sensory states that are all pleasant or pleasurable to undergo. Correlatively, there are sensory *displeasures*; for instance, the unpleasant feeling of a cold slime, the bad taste of coriander, or the sound of nails scraping a chalkboard.⁵⁰

Sensory pleasure does not exhaust the class of hedonic phenomena. Consider the pleasure some take in meticulously polishing a prized vehi-

⁵⁰Notice here that I identify the correlative of pleasure as *displeasure* rather than pain. By 'pain', I specifically mean *physical pain* such as the pain upon stubbing one's toe. As Bain and Brady (2014) point out, there is a broad use of 'pain' to include all states of suffering, anxiety, depression, etc.

cle. Here what is pleasurable is an *action* or *activity*. That is not to say that the relevant activity does not have a sensuous nature; for instance, one can enjoy, or take pleasure in, an elaborate twelve course tasting menu, an activity which consists in various sensory pleasures. Correlatively, there are activities that can be displeasing to engage in. Again, some displeasurable activities like cleaning a rubbish bin can involve sensory displeasure, but need not, for instance waiting in line at the post office or filling out one's tax returns.

We can again distinguish sensory pleasures and pleasurable activities from what Feldman calls *propositional* pleasures since we ascribe such pleasurable states with reports of the form “*S* is pleased that *p*” and “*S* takes pleasure in the fact that *p*” (Feldman, 1988, p. 60). Take for instance being pleased that one passed the exam. Correlatively, there are *propositional displeasures*; for instance, being displeased that a friend has betrayed one. Propositional pleasures need not be ‘tied’ to particular sensory states. When one is pleased that the library is (finally) quiet, it need not be the case that silence is pleasurable. Nor does, as Anscombe (1967) pointed out, being pleased or enjoying the fact that one ϕ -ed entail that one found ϕ -ing pleasurable. To use her example, one might enjoy or be pleased by the fact that one went riding with Nina without finding the ride that one took with Nina pleasurable or enjoyable.

There are philosophically substantive questions about pleasure that will not be pursued here. For instance, what is it for a sensory state to be pleasant or unpleasant? Is there some feature in virtue of which all sensory pleasures are pleasurable?⁵¹ Is there some common factor that unites sensory pleasures, pleasurable activities and propositional pleasures? What is the relationship between sensory pleasures and propositional pleasures?⁵²

⁵¹A special difficulty arises when theorists try to articulate the sense in which a seemingly heterogenous class of sensory states, within and across sense modalities, can all qualify as *pleasant* (or *unpleasant*). Compare the pleasant bitter taste of dark chocolate and the pleasant cool feeling of a breeze; these are distinct types of kinds of sensory experiences, yet both count as pleasant. Is there some property in virtue of which these sensory experiences count as pleasant? This is one version of what is called the ‘heterogeneity problem’.

⁵²For discussion, see Feldman (1988).

Satisfaction and Frustration

What I wish to turn to now is the relationship between desire and the kind of phenomena I have identified as core sub-classes of the hedonic realm. In the introduction to this chapter, I referred to natural claims we make about desire that clearly refer to some kind of hedonic connection. We do speak of the pleasure of satisfying a desire, the discomfort or even pain of a frustrated or unfulfilled desire. What truth is there to these claims?

It is important to differentiate a purely *semantic* notion of a desire's being satisfied or frustrated from a *psychological* understanding of satisfaction and frustration (Stampe, 1986; Lycan, 2012). Many philosophical discussions employ the terms 'satisfaction' and 'frustration' in this purely semantic sense. On the purely semantic notion, a desire that p is 'satisfied' just in case p , otherwise it is 'frustrated'. In order to avoid confusion, I will instead use the terms 'fulfilled' and 'unfulfilled' as synonyms for 'satisfaction' and 'frustration' in the semantic sense. However we should not diminish the perfectly good psychological concepts that 'satisfaction' and 'frustration' can express. Fulfilling a desire, that is, getting what you want can be *satisfying*. An unfulfilled desire can be *frustrating*.

3.1.2 The Simple Hedonic View

Desire's connection with psychological concepts like satisfaction and frustration gives us a foothold on how desires can be connected with the hedonic realm. Before I consider how exactly to capture this relationship, it is worth examining how these claims about the connection between desire and pleasure can be incorporated into an account of desire.

Consider what Michael Smith calls a "strong phenomenological conception" of desire (Smith, 1994, p. 107) on which desires are *identified* with sensation-like states with a distinctively phenomenal character. A version of this view might identify desires with certain kinds of hedonic or algedonic states, for instance states of discomfort that can be alleviated by ϕ -ing. Smith rightly raises the problem that any such views face a serious obstacle of explaining how we can make sense of desire as having the content it does simply by appeal to the phenomenal character of discomfort. Desires are states with content. A desire, say, to eat a slice

of cake, can be understood as relating a subject to the representational content *that one eat a cake*.⁵³ The problem is that it is unclear first of all if sensations have representational contents.⁵⁴ Second, even if we grant a form of intentionalism about their phenomenal character, it would still not be clear if the kind of content employed in such explanations of the conscious character of discomfort constitute contents of the kind that we take desires to have, e.g. for the objects of desire (Smith, 1994, pp. 107–8).

If some essential connection to the hedonic realm is to be forged, it will be within the context of a *weak* phenomenological conception of desire. Weak phenomenological conceptions of desire do not identify desires with experiential states or events. Such views can thus piggyback off more plausible ways to ground the representational content of desires. The motivational functionalist account seems especially promising in this regard and the hedonic account can be formulated within this framework as an *expansion* of the functional role of desire to include hedonic and algedonic elements. (Though Smith would of course deny that these hedonic or affective elements are essential to desire. This is because for him, the notion of a desire is theoretically demarcated first and foremost via its role in action explanation.) Nevertheless, he is at least receptive to the prima facie plausible thought that the ordinary notion of desire is a complex state involving both motivational *and* affective components (Smith, 2012b, p. 82).

We can thus formulate a *partial* condition on desire, i.e. one that only states a necessary condition on desire:

The Simple Hedonic View: a desire to ϕ is a state which is such that:

(S1) its fulfilment is satisfying, a *pleasurable* psychological state or event, or

⁵³In claiming that desires have ‘representational’ content, this is not to say that desires necessarily represent the world as being a certain way, as in judgement or belief. They are not, as (Martin, 2002, p. 386) puts it, ‘stative’ or ‘assertive’. Nevertheless, desires have content in a broader sense common to non-assertive states like wishing or hoping “since they are all about (or of, or involve reference to) objects, properties and states of affairs” (Martin, 2002, p. 387).

⁵⁴For discussion, see Cutter (2017).

(S2) its being unfulfilled is frustrating or uncomfortable, a *displeasurable* psychological state or event.

The Simple Hedonic View promises to provide an elegant way to explain how desire can contribute to what it is rational to do. It does so given the plausible enough assumption that there is some defeasible reason for subjects to pursue pleasure, or alleviate their state of discomfort or displeasure. So if true, the fact that you have a desire that p grounds the existence of a reason there is for you to bring about that p . This explanation accommodates the two constraints discussed in Chapter 2. The Rational Support Constraint is accommodated by the natural enough thought that one's being in a state of discomfort constitutes a good basis to do what will alleviate the discomfort. The Intelligibility Constraint is accommodated by the fact that states of discomfort have a distinctive phenomenal character, their unpleasantness, which forms an adequate explanation of why it is intelligible for a subject to do what would alleviate the state of discomfort and *a fortiori*, satisfy the desire.

To put the Simple Hedonic view in the best position for success, I have formulated it as a disjunction of two conditions. Nevertheless, there are, as I will now argue, obvious problems with the Simple Hedonic View.

Problem: The Generality Problem

The Simple Hedonic View considered as a general claim about desire faces some insuperable problems. First and foremost, the mere fulfilment of a desire need not be satisfying. Clearly, getting what you want will be pleasurable if what you want consists in, or has as an effect, sensory pleasure or the doing of a pleasurable activity. But there are many basic desires that are not for objects that would result in sensory pleasure or a pleasant activity. The fulfilment of a desire is simply the obtaining of what the subject wants, and this can happen without the subject being aware of this. Satisfaction as understood by the Simple Hedonic view then is plausibly a psychological state or episode that depends on the subject's epistemic state. *A fortiori*, the mere fulfilment of a desire need not be, as (S1) suggests, pleasurable. So (S1) is false if intended as a general truth about desire. Even amending (S1) to refer to the *awareness* of the fulfilment of a desire need not guarantee any kind of pleasure for

the same reasons. Take the desire to clean the bathroom. It is easy to imagine a case in which the known fulfilment of that desire fails to yield any kind of tangible sensory pleasure.

Nevertheless, might one not be *pleased* at having fulfilled the desire? Expressions like ‘being pleased that *p*’ are used to express very general states of preference. Having fulfilled a desire, one might be pleased, other things being equal, that one got what one wanted. This is certainly true in many cases, but again is clearly false if intended as a general truth about desire. Whether one will be pleased at having fulfilled the desire depends in large part on the broader cognitive and conative state of the subject. My desire to strike my opponent in a game of tennis, when fulfilled, is likely only to produce self-disgust or regret. I certainly would not be pleased at having fulfilled my base desire. Less dramatic examples can be given without invoking desires that the subject would disavow. Thinking nothing about it, one might indulge a passing desire to have an espresso after dinner. Tossing and turning in bed later, one might then regret having indulged in that unnecessary shot of caffeine as one is kept awake.

The same problems afflict (S2); for an unfulfilled desire is not necessarily *frustrating*. A desire can be frustrating when attempts to fulfil the desire are unsuccessful, but this will depend largely on the strength and importance of the desire in question. A desire to order the sticky toffee pudding after dinner need not be frustrating when one is informed that the kitchen is all out of desserts; one might in fact be glad, thinking to oneself that one probably should not have dessert anyway given one’s increasing waistline.

These are compelling reasons to think that appealing to a connection between desire and the hedonic realm is not workable as a general claim about desire. Nevertheless, it is worth exploring if a robust connection can be sustained in the case of more specific desires.

3.1.3 Schiffer (1976)

In this section, I consider an account of desire proposed by Stephen Schiffer that centralises the role of hedonic phenomena in our understanding of a sub-class of desires. By being essentially connected in this way with

pleasure and displeasure, Schiffer argues that some desires rationalise action simply by constituting reasons for action.

Reason-Providing vs. Reason-Following Desires

Let us begin with an important distinction between what Schiffer calls ‘reason-providing’ (r-p) and ‘reason-following’ (r-f) desires.

Schiffer articulates this distinction by reference to a difference in the way that the relevant desires rationalise action. When one acts on an r-f desire to ϕ , one’s ϕ -ing is rationalised on account of the reason R for which one has the desire, where R is “entirely independent, logically, of the fact that one desires to ϕ ” (Schiffer, 1976, p. 197). Take for example acting on a desire to play the guitar in order to improve one’s skills. Here the fact that playing the guitar will contribute to the improvement of one’s skills is one’s reason for playing the guitar, rather than the fact that one wants to play the guitar.

In contrast, Schiffer tells us that in acting on an r-p desire to ϕ :

the reason for which one ϕ -s and, typically, the only reason one has to ϕ , is provided entirely by one’s desire to ϕ and... one’s reason for ϕ -ing just is that desire. (Schiffer, 1976, p. 198)

As examples, Schiffer primarily appeals to the “aroused bodily appetites” like those associated with hunger and thirst, as well as desires to scratch one’s nose (Schiffer, 1976, p. 202). We will explore these cases in more detail when we consider Schiffer’s explanation of *how* they qualify as reason-providing.

For now, it is worth unpacking this claim about r-p desires further, separating out a claim about normative reasons and one about motivating reasons. First, Schiffer claims that “typically the only reason one has to ϕ ... is provided entirely by one’s desire to ϕ ” (Schiffer, 1976, p. 198). This is naturally interpreted as the claim that having a desire to ϕ *constitutes* a normative reason to ϕ . Schiffer also claims that when one acts on a desire to ϕ , “the reason for which one ϕ -s... is provided entirely by one’s desire to ϕ ... one’s reason for ϕ -ing just is that desire” (Schiffer, 1976, p. 198). What it is to act *for* a reason is philosophically contentious, nevertheless it is relatively uncontroversial that when R is the reason for which one ϕ -s, it must be the case that R is an apparent reason for one

to ϕ . We can summarise both claims more perspicuously in the following form:

(RP1) when one is in a position to act on an r-p desire, the fact that one has an r-p desire to ϕ constitutes an apparent reason to ϕ .

In addition to this, Schiffer makes another striking claim about r-p desires:

R-p-desires also provide the reasons, the justifications, for themselves. There is never a reason for which one has an r-p desire, one never has an r-p desire because one thinks of what one desires as being desirable in a certain way; *one does, however, find desirable what one has an r-p desire for, and because what is desired is desirable in the way found desirable, one does have a reason to desire what one desires*, and one's desire is, in the absence of defeating considerations, justified. (Schiffer, 1976, 198, emphasis added)

Focus here on the italicised claim, namely that if one has an r-p desire to ϕ , then:

(RP2) The fact that you have an r-p desire to ϕ constitutes an apparent reason to *want* to ϕ in virtue of the *way* one finds ϕ -ing desirable.

Indeed, given that reasons for wanting to ϕ are typically just reasons to ϕ , the fact that desires to ϕ constitute apparent reasons to ϕ can be thought to *explain* their self-justification. It is worth pausing to admire just how neat a resolution to the Problem of Conative Significance Schiffer's view promises, at least for a proper subset of desires. If Schiffer's view is right, the rational significance of some desires are substantive rationalisers of action by constituting reasons for action, and simply in virtue of having the desire, their character *as* reasons for action would thereby be apparent to the subject (RP1). A lot then hinges on the central feature that Schiffer takes some desires to have, namely that they involve a *way* in which one finds what one desires desirable. This, Schiffer claims, is explained by the fact that some desires involve states of pleasure and discomfort to which we now turn.

Schifferian Desires

According to Schiffer, some desires are reason-providing, that is, satisfy (RP1) and (RP2) because they satisfy a certain condition:

[W]hen one acts on an r-p-desire one acts for the gain of pleasure and the relief of discomfort— usually both, always one or the other— that one’s action affords; it is for that one acts and that is why one acts, and because one’s action is made desirable by its connection with pleasure and discomfort, one’s desire is justified.

An r-p-desire is a self-referential desire for its own gratification; an r-p-desire to ϕ is a desire to ϕ to relieve the discomfort of that desire, a desire to ϕ for the pleasure of its own relief. So a thirst is a desire to drink, a discomforting desire to drink, a discomforting desire to drink which would be pleasurable to relieve, a desire to drink to gratify itself. In this way r-p-desires are self-referring. (Schiffer, 1976, p. 199)

Before we delve into Schiffer’s explanation of reason-providing property in terms of discomfort and pleasure, two clarifications are in order. First, a terminological clarification. Schiffer takes himself to be giving an account of r-p desires and for that reason refers to the class of desires he gives as ‘r-p desires’. In presenting his view, I wish to keep these two ideas apart. I reserve the term ‘r-p desire’ to refer to *desires that are not r-f desires and constitute reasons for action*; I use “Schiffer’s desires” or “Schifferian desire to ϕ ” to refer to the kind of desires that Schiffer thinks qualify as r-p desires. Second, Schiffer says that in desiring to ϕ , ϕ -ing “is made desirable by its connection with pleasure and discomfort” (Schiffer, 1976, p. 198). But in order to preserve the thought that Schifferian desires are not reason *following*, we cannot interpret this as claiming that one wants to ϕ *for the reason that* ϕ -ing would relieve discomfort or bring pleasure. Schiffer treads this fine line by claiming that the discomfort of desire and the pleasurable relief its fulfilment brings is not a contingent causal effect distinct from the desire itself (Schiffer, 1976, p. 199). Rather the thought is that “[i] the desire to ϕ , [ii] one’s desire to gain the pleasure of satisfying one’s desire to ϕ and [iii] one’s desire to relieve the discomfort of one’s desire to ϕ ... *are all one and the same desire*” (Schiffer, 1976, p. 199).

The rationalising component of Schifferian desires, the way in which what is desired is desirable, is disjunctive. It consists in either a discomfort based component or a pleasure based component. I will start by considering the discomfort based component.

Discomfort Component: some desires to ϕ are (i) inherently discomforting and (ii) are simultaneously desires to ϕ to relieve the discomfort of one's desire to ϕ .

On the face of it, this claim seems obscure. What is it for a desire to be *inherently* uncomfortable? To get a grip on this idea, I suggest we turn to the class of desires that Schiffer takes to be clear instances of Schifferian desires, “aroused bodily appetites”. Here I focus on appetitive desires to eat, as when one is hungry. The claim that some appetitive desires involve uncomfortable states is plausible, but it is worth examining the kind of discomfort that is involved.

What most people refer to as ‘the feeling of hunger’ should not be thought of as picking out a unique bodily sensation. Whilst many focus on the visceral sensation often called a *hunger pang*, it is widely accepted in the empirical research that human hunger involves a syndrome of bodily sensations and that hunger pangs are by no means a necessarily concomitant of hunger (cf. Monello and Mayer, 1967; Harris and Wardle, 1987).⁵⁵ In addition to this range of bodily sensations, the phenomenal character associated with appetitive desires also involves a range of other more psychological components. There is a connection between states of hunger and negative mood or emotion (MacCormack and Lindquist, 2019), captured by the colloquial expression ‘hangry’. Appetitive states can also lead to certain kinds of mind-wandering (Rummel and Nied, 2017), where the content of one's thoughts turn to concern what is desired. This is often combined with a perceptual-cum-attentional component: being hungry, one's attention might be exogenously captured by food-related cues. In addition, there is what might be described as a felt

⁵⁵Research has also shown that the feelings of hunger are strikingly diffuse and variable, interpersonally and temporally. Harris and Wardle (1987) found that the cluster of symptoms which best discriminated pre-meal and post-meal subjects yielded a “rather poor classification rate”, concluding that “[it] proved impossible to identify a specific subset of the constellation of hunger symptoms which were characteristically experienced by hungry people” (Harris and Wardle, 1987, p. 154).

urge. It is difficult to give linguistic expression to this aspect of a keenly felt appetitive desire to eat, other than with terms like *compulsion*, *drive* and so on.

What I want to emphasise here is that appetitive desires to eat can vary with intensity, where intensity is a multi-dimensional notion. A low intensity appetitive desire to eat might involve relatively minor unintrusive forms of mind-wandering, limited attentional effects and an imperceptible felt urge to eat. But over a certain intensity threshold, such desires can be uncomfortable in a range of ways: excessive mind-wandering can become a distraction, felt urges grow to uncomfortable levels, attention on strongly bodily sensations can also be an annoying distraction. So here we see that not all appetitive desires are discomforting, only those of a certain intensity, where intensity is a multidimensional notion involving other cognitive/sensory components.

In a critical discussion of Schiffer's view, Schueler (1995) argues that the characterisation of aroused bodily appetites as discomforting is phenomenologically inaccurate, writing that "[f]or most people, the mild thirst they experience at least a few times every day... has practically no phenomenological character at all, certainly nothing that could be called 'discomfort'" (Schueler, 1995, pp. 87–8). Though Schueler concedes that more intense appetite desires might have this character, he takes this to constitute a problem for Schiffer since the proposed explanation would only cover exceptional cases of appetitive desires. "[I]t would seem implausible", Schueler writes, "to claim that while intense thirst is a reason-providing desire, everyday thirst is not" (Schueler, 1995, p. 88).

There might be room for reasonable contestation given the phenomenological basis for Schueler's objection. Schuler's objection that Schifferian desires capture only exceptional cases relies on setting the threshold for discomfort relatively high. An obvious response for Schiffer to make would be to contest where the relevant intensity threshold lies. After all, Schueler does not claim that mild appetitive desires have *no* phenomenal character, only that it might accurately be described as uncomfortable. Here there is scope for considerable interpersonal variation. To take just one measure, it is plausible that different subjects can vary in their propensity to attend to the sensational concomitants of thirst and hunger

in a way that constitutes a distraction or annoyance. Even here, whether this aspect of mind-wandering counts as distracting or annoying depends on the temperament of the subject in question.

Let us now turn to the other component of Schifferian desires which Schiffer claims is sufficient for their being reason-providing (Schiffer, 1976, p. 199):

Pleasure Component: some desires to ϕ are (i) inherently pleasurable to fulfil and (ii) they are simultaneously a desire to ϕ to yield the pleasure of fulfilling one's desire to ϕ .

In order for the pleasure component to generalise even to the limited case of appetitive desires, there must be a very tight connection between the desire and the pleasure of its fulfilment. We know of many cases of basic desires, say to taste a durian fruit, where fulfilment would yield not pleasure but revulsion. For this pleasure component to hold in general, it must be instead that the desire is identical with a state which is inherently such that its fulfilment yields pleasure. And to reiterate an earlier point, in order that Schifferian desires do not collapse into r-f desires, it cannot be the case that the desire to ϕ is had for the reason that ϕ -ing will yield pleasure. But we cannot avoid this in the same way as with the discomfort component, since the pleasure in question is yet to be experienced.

There are two pressing worries. First, is it plausible that there are desires that are inherently pleasurable to fulfil? Second, even granting that there are such states, it is not clear to me that the pleasure component *is* sufficient to establish (RP1), that an r-p desire to ϕ constitutes an apparent reason for a subject to ϕ . But how can simply being in a state that would yield pleasure if one were to ϕ be *apparent* to one as a reason to ϕ ? It is possible to be in such a state yet not realise that one is in that state. This is a problem not faced by the discomfort aspect of Schifferian desires since the thought that the conscious character of discomforting states constitute apparent reasons to bring about their elimination is plausible.

The most plausible way to address these problems will be to make the pleasure component parasitic on the discomfort component. If relief from discomfort is *ipso facto* pleasurable, and intense appetitive desires

are inherently discomfoting states, then we can avail ourselves of the following explanation of the pleasure component: a Schifferian desire to ϕ is a state the fulfilment of which guarantees pleasure simply in virtue of the fact that relief from the discomfort of an inherently discomfoting Schifferian desire to ϕ is pleasurable. If this is right, then it seems that Schiffer will have to give up the *sufficiency* of the pleasure component for a desire's being reason-providing.

I wish to briefly summarise the results so far. Schiffer argues that some desires are reason-providing in the sense that they constitute reasons for action, such that if one is in a position to act on a desire to ϕ , then the fact that one has the desire is an apparent reason to ϕ (RP1). Moreover, in virtue of the way in which one desires to ϕ , the desire constitutes a justification for that desire (RP2). Some desires, chiefly appetitive ones, satisfy (RP1) and (RP2) *because* of their connections with the hedonic realm. In trying to clarify Schiffer's view, I have suggested that it is plausible that whether or not appetitive desires satisfy (RP1) is variable across subjects, so it is an open question whether (RP1) holds universally. Moreover, I have suggested that the discomfort component takes priority with respect to explaining (RP1) and (RP2), since it is unclear whether the pleasure component is sufficient to explain (RP1) and (RP2).

3.1.4 Problems for Schiffer's Account

In this section, I pose what I take to be the core objection to attempts to capture the rational significance of desire by appeal to reasons there are for the subject to alleviate discomfort. I will argue that even if some desires can rationalise action by providing discomfort-based reasons, intuitions about cases where it is possible to alleviate discomfort without fulfilling the desire show that the provision of such reasons cannot capture everything we want to say about why a basic desire to ϕ can rationalise one's ϕ -ing. Before this, I briefly raise a general worry about whether Schiffer's view that when one acts on the basis of an appetitive desire one acts for the reason that one has the appetitive desire in question.

Problem 1: Acting on Desire and Desires as Reasons

On Schiffer's view, when one acts on the basis of an appetitive desire, one acts for the reason *that one has the desire* in question (as per RP1) where the desire itself constitutes its own justification for having the desire (as per RP2). The aspect I wish to emphasise is that on this view, whenever we act on an appetitive desire such desires always lie at the 'foreground' of practical reasoning in the sense that the fact that one has the desire in question constitutes the normative reason for which one acts.⁵⁶

It is important to question if every case in which one responds to a basic desire has this character. Even focussing on appetitive desires, it is plausible that in many cases, mild appetitive desires can intelligibly guide action without the subject acting for the reason that one has that appetitive desire to drink. In such cases, an occurrent appetitive desire can render salient a known fact, that there's a bottle of water in one's bag, and one rationally treats *this* consideration as a reason to unzip the bag and rummage for the bottle. At no point need one reflect on the presence of the desire.⁵⁷

I will further develop cases like these in Chapter 6, explaining the characteristic activity of occurrent basic desires in guiding and rationalising action in the background. If the account of acting on an occurrent desire proposed is correct, then this should bolster scepticism about whether basic desires are necessarily 'foregrounded' when one acts on such desires. For now however, I wish to turn to a pressing problem of Schiffer's view.

Problem 2: The Rational Significance of Desire and Discomfort

Is it the case that the rational significance of appetitive desires can be accounted for entirely by the discomfort-based reasons that they constitute? The strong reason for thinking not comes by reflecting on cases in which it is possible to eliminate what we will grant is a discomforting desire to ϕ without ϕ -ing. Suppose one is enjoying some free time at home, relaxing without doing anything in particular and one begins to feel a

⁵⁶The term 'foreground' is one from Pettit and Smith (1990). Pettit and Smith's distinction will be discussed in Chapter 6.

⁵⁷Compare Alvarez (2010, pp. 100–18)'s discussion of appetitive desires.

strong desire to eat something. One attends to the gnawing sensation in the pit of one's stomach and one decides to do something about it. We can stipulate that one has two choices: one can easily go to the fridge and consume the sandwich, or one can go to the study and turn a switch to alter the quality of the sensation into a neutral one. We can stipulate here that neither option requires more effort than the other, and that consuming the sandwich will not yield any pleasure.

There are of course cases in which it would be rational for you to eliminate the effects that attend to appetitive desires rather than satisfy it, as when the costs of satisfying it are believed to far outweigh eliminating it. For instance, we might imagine that during an important exam in which the consumption of food is forbidden, the fact that you have a distracting appetitive desire to eat might rationalise your doing what you can to eliminate it in ways other than eating. However the case I am concerned with is one in which the cost of satisfying the desire is no greater than the cost of eliminating it, *vice versa*.

In such a case, it is intuitive that the desire rationalises doing what will result in its fulfilment, rather to do what eliminates the desire e.g. altering how it feels. The reason why it is plausible to think that it would be rational for one to procure the sandwich rather than eliminate his desire altogether stems from the thought that there are reasons why it is rational to satisfy a desire, ones that diverge from the reasons there are to seek relief from discomfort. In his discussion of Schiffer's view, Dennis Stampe makes the following point:

There is no denying that a desire often constitutes a reason to act in some such way to rid oneself of it. But even where this is feasible a desire remains a reason... to act so as to bring it about that the state of affairs one wants to obtain *does* obtain. And the view under discussion [Schiffer's view] cannot explain why that is so. It cannot explain why, even if one could extinguish the desire without fulfilling it, one's wanting the thing would still be a reason to try to get that thing and *fulfil* that desire. If for example I am hungry and want to eat, and I have a way of getting rid of that hunger and that desire without eating, it remains the case that I have, in fact that I want to eat, *a reason* to do exactly that. (Stampe, 1987, p. 350)

According to Stampe, Schiffer's view cannot explain why the desire is "still" or "remains" a reason to satisfy rather than merely eliminate the desire. This way of putting things is a bit unclear. After all, Schiffer claims that provided you still have the desire and so remain in an uncomfortable state, you still have a reason to satisfy it and the desire *remains* a reason to satisfy it. I think we can sharpen Stampe's objection by appealing to intuitions about the *balance of reasons*. Schiffer's view *identifies* the reasons that desire provides with reasons for discomfort alleviation. The problem is that, granting that some desires are inherently discomfoting states which generate reasons to do what will alleviate that discomfort, neither view can explain why it seems that the rational thing to do is to pursue a course of action wherein one eliminates the desire *by fulfilling it* rather than one which eliminates *without* fulfilling it.

What this objection relies on is the highly plausible thought that the kind of reason that appetitive desire typically provides is not just the ones that pertain to discomfort alleviation. Sometimes appetitive desires to eat are unpleasant to have: they can make one feel weak, they are distracting, and so on. In these cases, we certainly have some reason to terminate the desire. But intuitively, the reasons that such desires provide do not merely consists of such reasons. Appetitive desires to eat are (roughly) caused by a state of nutritional deficiency, a bad bodily state. If Schiffer's view were the *whole* story, then the only reasons we have to satisfy appetitive desires are the reasons we have for discomfort alleviation. The balance of reason would not decide between terminating the desire by consuming a desire-eliminating pill or terminating a desire by eating. But intuitively the reasons desire provides are not just those of discomfort-alleviation, they pertain to the role that appetitive desires play in motivating us to address bad bodily states. This claim is especially plausible in the appetitive case given the close link between appetitive states and biological need. In order to understand why *non-appetitive* desires can rationalise action it is necessary to provide an account of the role and function of desire more generally. As I argue in Chapters 5, desires are the products of mechanisms that have the function to produce good-conducive desires. It is in virtue of this feature that desires have the power to rationalise choice and action in ways that can explain where what it is rational to do in having a desire

diverge from what it is rational to do to alleviate discomfort.

3.2 Desire and Affect

We will turn away from pleasure and discomfort to explore different connections that desire might have to other affective phenomena. In the rest of this chapter, we will consider proposals developed by Ruth Chang, Declan Smithies and Jeremy Weiss (in joint work), and Peter Railton. All of these theorists have emphasised the role of a certain kind valenced affective state in our understanding of desire.

This section focusses primarily on the first two proposals in Chang (2004) and Smithies and Weiss (2019), leaving discussion of Railton’s view for the final section. The reason for this is that the accounts of Chang, Smithies and Weiss are claimed to provide similar solutions to the Problem of Conative Significance, whilst Railton’s view suggests an importantly different style of explanation. The first part of this section will be concerned with outlining and clarifying the kinds of affective experiences that Chang, Smithies and Weiss focus on in their accounts desire, or at least certain sub-classes of desire §3.2.1. The second part of this section will then consider and critically evaluate two specific proposals for how this affective aspect of desire can account for the rational significance of desire.

3.2.1 Affective Experiences and Desire

Chang (2004) argues that some desires are what she calls ‘affective desires’, “non-cognitive states essentially involving attraction to their objects without reference to any cognitive or quasi-cognitive element” (Chang, 2004, p. 68). Attractions are “intentional states: that is, they involve having an *attitude about some content*” (Chang, 2004, p. 69).

What is distinctive about these affective desires is that they “essentially involve some phenomenological feel” which she describes as involving finding what one desires appealing or attractive (Chang, 2004, p. 68). Chang differentiates between *feature-bound* and *feature-free* attractions. Feature-bound affective desires involve attractions to “particular features of an object; [for instance,] one is attracted to the creaminess of the ba-

nana” (Chang, 2004, p. 80). Feature-free affective desires “are desires for *the object itself* but not under any particular description” (Chang, 2004, p. 80). There is some unclarity in the intentional object of Chang’s feature-free affective desires. On the one hand, it is natural to understand the object of feature-free attractions as involving worldly objects such as apples and cars, but Chang gives a gloss of feature-free affective desires with act-types rather than worldly individuals as objects. For instance she writes:

[T]o distinguish feature-bound from feature-free affective desires, I will call the later “feelings like it’... If one ‘feels like’ wearing pink, there need be no particular feature of wearing pink that attracts one; one *just feels like* wearing pink. (Chang, 2004, p. 80)

Here *wearing pink* is the (intentional) object of the feature-free attraction, an act-type, but not a suitable worldly object like a pink sweater. In the examples she gives (Chang, 2004, pp. 80, 82), Chang seems to be assuming some connection between feature-free attractions and feeling like ϕ -ing. In discussing the case of Buridan’s ass, Chang considers the possibility of someone forming an attraction to one of two identical bales of hay (worldly object), which she claims gives rise to one’s feeling like *choosing* that bale of hay. Unfortunately, Chang does not spell out this connection.

Smithies and Weiss (2019) defend a theory targeted specifically at desires one has “not in virtue of having any other belief or desire” (Smithies and Weiss, 2019, p. 45). Call these *foundational* desires.⁵⁸ According to them, foundational desires are “dispositions to cause the phenomenal character of affective experiences of desire” (Smithies and Weiss, 2019, p. 49). Smithies and Weiss appeal to two classes of affective experience that they take to individuate such desires, “feelings of attraction or aversion” (Smithies and Weiss, 2019, p. 27). Like Chang’s affective desires, affective experiences of attraction or aversion are experiential episodes. Affective experiences of desire are a proper subset of affective experiences. Like affective experiences in general, they have three properties. First, affective experiences of desire are phenomenally conscious. Second, they

⁵⁸Smithies and Weiss calls these ‘basic desires’; I have suppressed this in order to avoid confusion with my notion of basic desire, which is clearly related.

are *intentional* in having the property of being about things. Third, and most importantly, they are *valenced* by which they mean that “[affective experiences] represent their intentional objects in a *positive or negative way*” (Smithies and Weiss, 2019, 29, emphasis added).

There are some differences between the two views. The first difference concerns how desires relate to these affective states or experiences. For Chang, some desires, viz. affective desires, simply *are* these episodes of attraction or ‘feelings like it’. In contrast, Smithies and Weiss do not make this identification, maintaining that basic desires are *dispositions* to have such affective experiences.

The second difference concerns the intentional objects of the kinds of affective experiences appealed to by both accounts. As we have seen, Chang claims that affective desires are directed at *features* of objects (e.g. the pinkness of a sweater), or objects themselves. But, as just pointed out, Chang seems also to talk about “feeling like it” which relate subjects to act-types. All her examples of ‘feelings like it’ seem grounded in attractions to objects and their features which are perceptually available to the subject. Smithies and Weiss however centrally appeal to affective experiences which involve valenced responses to “the prospect” of performing certain acts (e.g. one is attracted to the prospect of wearing a pink sweater). What does it mean to be attracted to the *prospect* of ϕ -ing? They seem to indicate that this is to be understood as a kind of response to conscious *thoughts* that concern the possibility ϕ -ing. For instance, one way for one to be attracted to the prospect of finally eating the simmering stew is for one to feel positive when thoughts of eating the stew occur to one, or when one reflects (Smithies and Weiss, 2019, p. 39).

Before considering arguments for the centrality of such affective experiences to desire’s rational role, it is important to distinguish these views from the two hedonic views we have considered. The two hedonic accounts considered implausibly suggested that to be in a state of desire is for one to be in a state that is occurrently pleasurable when fulfilled or displeasurable when unfulfilled. The affect-based views considered are not committed to either of these claims; for instance, it is possible on Smithies and Weiss’ view that one can be attracted to the prospect of ϕ -ing even if you would not find it pleasurable were you to ϕ . For example,

one might be attracted to the prospect of tasting a durian fruit even if one would find the flavour unpleasant (Smithies and Weiss, 2019, p. 44). Moreover, it is not the case that states of attraction are necessarily pleasurable or displeasurable states to be in. As Smithies and Weiss' point out, sometimes romantic attraction can be displeasurable states to endure (Smithies and Weiss, 2019, p. 44). Instead, to feel attracted to the prospect that P is a “distinct species of the genus” of which experiences of pleasure are also a sub-species.

3.2.2 Two Arguments for the Centrality of Affective Experience

Having just outlined the kinds of affective elements these theorists take to be crucial to explaining the rational significance of desire, I turn now to the kinds of arguments adduced to suggest that such experiences are needed to explain desire's rational significance.

The plan is as follows. I begin by outlining Smithies and Weiss' argument where I will isolate a key premise that will be subject to closer evaluation. Next, I turn to Chang's argument and identify the same premise at work. According to this premise, affective experiences of desire have introspectible phenomenological features which rationalises acting in a certain way. This, I argue, is far from clear and I provide a challenge that must be met for the premise to be accepted (§3.2.3).

Smithies and Weiss' Argument

Smithies and Weiss begin with the datum that some desires are rationally significant in guiding choice and action.⁵⁹ Their central claim is that the

⁵⁹They motivate this claim in a different way than I did in Chapter 2. Very roughly, their argument flows from an undefended variation of Lewis' analytic functionalism about desire, different in emphasising an undefended “folk theory of mind” containing *normative* elements (Smithies and Weiss, 2019, p. 45). On this view, what desire *is* is whatever it is that plays the theoretical role of our normative folk theory according to which, beliefs and desires provide reasons for action. This style of argument for their affective theory of basic desire is far from uncontroversial. One central difference is that whereas the explanatory task I am engaged in is to *vindicate* the rational significance of desire, leaving open the possibility that desires might lack rational significance altogether, Smithies and Weiss take it for granted. Rather than focus on this difference, I want to focus on assessing how strong their case is for their main claim that basic desires rationalise action in virtue of being connected with affective experiences.

rational significance of these desires lies in their being dispositions to undergo conscious affective experiences of desire. Such affective episodes rationalises choice and action *without standing in need of further rational support* (Smithies and Weiss, 2019, p. 37). Stronger still, they claim that all basic desires provide reasons for action *only insofar* as they involve affective experiences (Smithies and Weiss, 2019, p. 27).

Smithies and Weiss rely on motivating their positive claim through an extended negative argument, eliminating alternative non-affect-based explanations of the rational significance of foundational desires. Their argument has the following structure:

- P1. Some desires, foundational desires, can rationalise action without standing in need of further rational support.
- P2. “Affective experiences provide reasons for action without standing in need of rational support by further reasons” (Smithies and Weiss, 2019, p. 37)
- P3. No rival account of foundational desires can explain how they rationalise without standing in need of further support. [By Argument in Smithies and Weiss, 2019, pp. 37–40]
- P4. Some desires, namely foundational desires, involve dispositions to have affective experiences of attraction and aversion. [Datum]
- C. So basic desires rationalise action without standing in need of further rational support only insofar as they are dispositions to undergo affective experiences of desire. [By P1-4 and Elimination of Alternatives]

Here I will not rehearse all the details of this argument; in particular, I will not outline all their negative arguments to establish (P3). Instead, I will briefly consider two main options they reject and briefly sketch the kind of problem they raise. The first option they consider is whether a mere behavioural disposition can provide one with a reason for action, ruling out this option by appeal to Quinn’s Radioman case. They argue that Radioman fails to possess a reason to turn on radios because he fails to satisfy what they call the ‘Reflective Access Constraint’ on reason-possession according to which “one has a reason to ϕ only if it is

epistemically rational for one to believe on the basis of reflection that one has a reason to ϕ " (Smithies and Weiss, 2019, p. 34). It is not rational, they claim, for Radioman to believe simply on reflection that he has the behavioural disposition in question, nor would it be rational for him to take his having that disposition to constitute a reason to turn on radios. Another option they consider is whether an evaluative belief that ϕ -ing is good/bad might provide one with an apparent reason to ϕ . They rule out this proposal because rational evaluative beliefs must be *based* on reasons. But foundational desires provide reasons without being based on reasons and so such desires cannot rationalise in virtue of consisting in some evaluative belief (Smithies and Weiss, 2019, p. 42).

Smithies and Weiss do not give an argument for (P2), baldly asserting that "[it is] extremely plausible that some affective experiences are capable of providing us with reasons for action. If you feel attracted to turning on radios, just for its own sake, then you thereby have some reason for turning them on" (Smithies and Weiss, 2019, p. 37) Given the elimination of rival explanations, they conclude that only affective experiences have the power to provide reasons without being based on reasons. Smithies and Weiss suggest that what explains why foundational desires can provide reasons for action without standing in need of rational support is that they consist in dispositions to have such experiences. By being connected with affective experiences of this sort, a desire, say to eat tomatoes satisfies the reflective access constraint because "I can know by reflection how I'm disposed to feel about tasting tomatoes. For instance, I can just manifest the disposition by thinking about tasting tomatoes and considering how I feel about it." (Smithies and Weiss, 2019, p. 39).

I want to highlight (P2) which I will focus on in the critical evaluation to follow. Before this, I wish to turn to Ruth Chang's argument for a related but slightly weaker conclusion where I will show that a similar premise is in play. Once it is clear how pivotal this premise is for Smithies, Weiss and Chang, I will argue that we have good reason to be dissatisfied with appealing to affective experiences as the final and most fundamental explanation of the rational significance of desire.

Chang's Argument

Chang (2004) defends a weaker claim than Smithies and Weiss, namely that some desires, in virtue of being constituted by affective experiences, constitute reasons for action.⁶⁰ Chang's argument rests on intuitive judgements about the rationality of choice and action in 'tied' cases in which a subject faces a choice between two rationally eligible options, i.e. both of which are backed by sufficient but inconclusive reasons. We have already considered a case satisfying this structure where Uma is faced with a choice between soup or a sandwich for lunch (§2.2.2). However, Chang's argument focusses on a more specific version of such cases in which one is faced with two *identical* options, like Buridan's Ass who is faced with two equally sized, equally enjoyable bales of hay.⁶¹ Chang asks us to consider a variation of this case that differs only with respect to the fact that the subject now forms an affective desire for one option:

Now suppose that [Buridan's ass] 'feels like' the hay on the left, not because it is to the left or for any feature for it— he just wants that bale. If he is attracted to the bale on the left but goes for the one of the right, *surely he would not be doing what he has most reason to do. He has most reason to eat the bale on the left since all other reasons are evenly matched and he is attracted to that bale.* (Chang, 2004, p. 80)

For ease of reference, call these cases *putative tie-breaking* cases. Focus on the claim that in acting against an affective desire between two evenly matched options, one would “not be doing what [one] has most reason to do” (Chang, 2004, p. 80). This presupposes that if one has an affective desire to go for one option, then one's affective desire constitutes an apparent reason one has to go for that option. But in the first place, why should we think that acting against an affective desire involves acting against the balance of reasons for action? When we read further, we find a clue:

⁶⁰To see this, consider whether or not we should countenance the existence of non-affective reason-giving basic desires which would be inconsistent with Smithies and Weiss' position but not Chang's.

⁶¹I discuss this difference in further detail in Chapter 6 where I will argue that our intuitions about the role of desire for the rationality of choice vary between these cases.

If he goes left, his action can be rationalised by pointing out that he felt like having the bale on the left. If he goes right, *we would need some explanation for this puzzling act*; what reason does he have to go right given that he would enjoy each of the two identical bales equally and feels like having the one on the left? *What we would be looking for is a reason to go right that counteracts his reason to go left provided by the fact that he feels like it...* The fact that one is attracted to something can *per se* provide a reason to go for it when all other reasons are evenly matched. (Chang, 2004, p. 81)

The intuition being pumped relies on the thought that it is rational to pursue the option one “feels like” pursuing and puzzling not to. And if, as Chang assumes, what is needed to explain acting against one’s affective inclination is the recognition of a countervailing reason, we should accept that such affective orientations weigh on choice as a reason for action does. I want to highlight that this move is questionable. However I want to delay discussion of this point until the Chapter 6 where I will provide an alternative explanation of the intuitions that Chang adduces with respect to these putative tie-breaking cases. Here I focus on Chang’s claim that phenomenological attractions can “per se provide a reason” (Chang, 2004, p. 81) which I will argue against directly.

3.2.3 Can Affective Experience Explain the Rational Significance of Desire?

We have now identified two claims that I wish to evaluate. To recapitulate:

[W]hen one feels like something, *one has a definite phenomenological attraction to something that draws one’s attention...* [When the ass acts] *he does so in full awareness of his attraction*. He does not act for no reason; on the contrary, he acts for the reason that he feels like having it” (Chang, 2004, p. 82).

[It is] extremely plausible that some affective experiences are capable of providing us with reasons for action. If you feel attracted to turning on radios, just for its own sake, then you thereby have some reason for turning them on” (Smithies and Weiss, 2019, p. 37)

Chang, Smithies and Weiss claim that the affective experiences they consider have introspectible features in virtue of which they constitute apparent reason for action. What are these features?

Chang does not provide further phenomenological description of the features in question, referring simply to the phenomenology of being *attracted* to something that draws one's attention. Smithies and Weiss go further writing that an attraction to the prospect of ϕ -ing rationalises one's ϕ -ing because they have "introspectible features— namely, its *phenomenal force* and *phenomenal content*— that seem upon reflection to count in favor of [ϕ -ing]" (Smithies and Weiss, 2019, p. 38). "Phenomenal valence" recall is the valenced orientation of an affective experience that comprises the intrinsic feature of what it is like to have the affective experience (Smithies and Weiss, 2019, p. 30). Moreover, these intrinsic features of such experiences also "provide introspective reasons for belief about the phenomenal features in virtue of which it provides reasons for action — namely, its phenomenal valence and its phenomenal content" (Smithies and Weiss, 2019, p. 38). Thus, it would be epistemically rational, they claim, for such a subject to believe that they have a reason to ϕ on the basis of their attraction to the prospect of ϕ (Smithies and Weiss, 2019, p. 38). In this way, affective experiences of desire satisfy their 'reflective access' constraint.

Both theories seem to rest content with an explanation of how some desires provide rational support that ends here. I wish to make the case now that this is an inadequate place for explanation to stop. Why should we accept the claim that the affective experiences involved in having certain desires rationalise action? The main objection I wish to press against these accounts concerns the sufficiency of these phenomenological features of affective experiences for having a reason for action. The style of argument I wish to press is related to arguments made in the epistemology of perception against internalist (specifically, 'dogmatic') views that claim perceptual experiences immediately justify beliefs in virtue of possessing a certain kind of phenomenal character. These arguments appeal to the fact that the epistemic status of beliefs formed on the basis of perceptual experiences can be undermined or 'downgraded' if the perceptual experiences have certain kinds of epistemically compromising etiologies (cf. Markie, 2006, Siegel, 2013, Teng, 2018).

I want to begin with a question that Chang, Smithies and Weiss do not address. Why do we have affective experiences of the kind they identify? The kind of explanation I am after here is not causal; I am not interested in the genetic basis for the kinds of traits responsible for the neurophysiological structures on which affective states supervene. Instead, I am interested in the functional question: what role or function do such affective experiences serve? Do they serve to attune or orient us to perform good courses of action? Assuming that traits responsible for our being able to enjoy such affective experiences conferred some adaptive advantage, we can ask *how* they have historically operated so as to confer that advantage. If they do not in some way track the good, then why think that such experiences are significant from the perspective of practical rationality? Without some account of this, it remains something of a mystery how content-bearing valenced experiences can positively support choice by constituting a fact that favours engaging in certain actions.

Once we foreground this question, we can begin to see why it is doubtful that the kind of introspectible features *intrinsic* to such experiences are sufficient to ground their constituting reasons for action. To see this, consider a variation on Quinn's Radioman in which we expand Radioman's range of desires such that he now has basic attractions for a whole host of odd things: turning on radios, possessing saucers of mud and twigs of mountain ash, counting blades of grass, and so on. We are imagining here that the things he is attracted to do does not conflict with apparent reasons *not* to do those things (so we are discounting yens to drink paint).⁶²

On the views considered, we have a subject who is acting perfectly rationally. Yet, there is good reason to think that there is something awry about Radioman. I think it is intuitive that we would seek some sort of explanation of just *why* this person has the attractions he does. His attractions strike us as strange. Is this person of sound mind? Is he ill or malfunctioning in some way? The fact that being attracted to such things grounds the propriety of concerns about pathology or dysfunction

⁶²Now of course if these strange ends are all that the subject is attracted to, then we might seek alternative explanation about what is wrong with this subject. Here we should clarify that there are other more normal things that the subject is attracted to.

tells us something about the nature of attraction and aversion. A very plausible thought is that episodes of attraction and aversion are related to what is good or bad respectively for the subject and so certain kinds of goods make for normal objects of attraction whereas others, such as possessing saucers of mud or counting blades of grass require some level of explanation in order to connect it to what is good for the subject. It is reasonable to think that there is a way in which valenced affective experiences are caused, sustained and regulated which relate what a subject has affective attractions to do and what would be good for the subject to do.

We can further develop this argument by considering cases of attractions borne out of processes that put subjects at the risk of forming bad conducive attractions. Suppose we elaborate how Radioman comes to have such attractions. Perhaps, he forms such attractions in virtue of prior exposure to substances that substantially deform the mechanisms that naturally give rise to such affective experiences. If the affective theories considered are right, then the fact that a subject is attracted to possessing a saucer of mud should still be sufficient to make it rational for her to try to acquire a saucer of mud. It is rational, they claim, because the subject can come to be aware on reflection of features intrinsic to the phenomenal character of her affective experience: its valence and content. Yet this does not seem right. Whether or not Chang, Smithies and Weiss are right that attractions constitute reasons for action in the ‘normal’ case, the ‘checkered’ etiology of Radioman’s attractions discounts the attraction as a genuine reason for action.⁶³ We have no reason to doubt that the nature of the attraction Radioman has to saucers of mud differs in their intrinsic properties to an episode of attraction he might bear without a checkered etiology. These considerations then justify taking a sceptical attitude to whether features intrinsic to the phenomenal character of affective experiences like attractions and aversions are *sufficient* to ground a reason to pursue what one is attracted to, or avoid what one is averse to.

Nevertheless, might this version of Radioman be rational in treating his attraction to ϕ as a reason for ϕ -ing in virtue of its intrinsic phenomenal character? This will require that we address what I called the

⁶³Here I use the notion of a checkered etiology drawing on Siegel (2013).

functional question concerning the role or function these attractions play. Do they serve to track the goodness-conferring properties of the objects or acts they affectively orient us toward? I have argued that affective experiences are *insufficient* to provide a subject with a reason for action, but I have left it open whether some kind of affective experiences are necessary for a subject to rationally respond to her reasons for action. For instance, if affective experiences are the product of sub-personal mechanisms with the function of tracking the good, then perhaps the fact they have the conscious properties they do *enable* the subject to respond to the appraisals of these sub-personal mechanisms.⁶⁴ However, Chang, Smithies and Weiss do not provide any such account. Without such an account, it is obscure *why* attractions can rationalise even in cases where those attractions no longer connect with what would be good for him to do.⁶⁵ The argument I have provided gives us reason to think that appealing to intrinsic features of experience *per se* is inadequate without an answer to the functional question.

In the final section of this chapter, I will consider an argument which *does* integrate the affective experiences in a story that addresses the functional question. This account is proposed by Peter Railton, one in which affect plays a role in the regulation of motivation (§3.3). I provide reasons why this regulatory role that affect plays can contribute to an explanation of desire's rational significance even if conscious affective episodes do not constitute the full and final answer to the problem of conative significance. This regulatory role is developed further in Chapter 5.

⁶⁴Consider for instance certain 'appraisal' based theories of emotions on which certain emotions like fear are the products of sometimes low level processes in response to certain significant stimuli that constitute a kind of appraisal or evaluation. For discussion, see van Reekum and Scherer (1997), Scherer et al. (2001), Ellsworth and Scherer (2003).

⁶⁵As I will defend in Chapters 5, 6 and in the final section of this chapter, such an explanation will advert to the functions of the mechanisms that give rise to basic desire and the kinds of affective experiences that are natural concomitants. The extreme case of Radioman is interesting since the etiology of his attractions are the result of processes that have been deformed. I will also discuss neighbouring cases in which desires are the product of non-deformed mechanisms in environments for which such mechanisms were not evolved to cope with.

3.3 The Regulatory Role of Affect

Peter Railton has developed over the course of several articles an account of desire which incorporates the kind of affective phenomena that Chang, Smithies and Weiss centralise.⁶⁶ According to what Railton calls the ‘prospective model’, desire is a “*dynamic functional state*” (Railton, 2012, p. 36) involving motivational and affective components that relate to each other as follows:

Prospective Model of Desire: [To desire that P is to have] A degree of positive affect (attraction, liking) toward a representation p [that] functions to elicit and regulate a degree of positive expectation (affective forecast) and positive motivation (striving, wanting) toward maintaining or bringing about the act or state of affairs that p portrays; and this degree of positive affect is subsequently modulated by whether the actual experience of performing, realizing, or moving toward p is better than, worse than, or in conformity with, the affective expectation of it (Railton, 2012, p. 36)

Desire, according to Railton, is neither simply liking an idea or being disposed to bring it about, but a “structured relationship” (Railton, 2012, p. 28) between these two components.

One aspect of this is the way in which liking *elicits* wanting. Consider the way in which we can form motivations for certain ends in response to certain kinds of cues related to food, sex and so on. For Railton, in normal cases, the transition from disinterest to motivation is mediated by an affective response to representations stimulated by those cues. Seeing someone quench their thirst with a cold beer sets the end of drinking a cold beer before the mind, where this representation of a prospective state of affairs is the target of “positive affective interest (attraction, liking)” (Railton, 2012, p. 36). This affective component then causes the formation of a disposition to secure that end, or directly induces active occurrent strivings or efforts to secure that end, the *wanting* component of Railton’s model. Coming to desire then, we can think of as, “a transition from affect into action: from liking *X* to having a positive affective representation of *X*-ing, and thence to focused motivation to maintain *X* or bring it about” (Railton, 2012, p. 29).

⁶⁶See Railton (2010), Railton (2012), Railton (2017).

Another role that the affective component plays is the regulation of motivation by setting up what Railton calls a “positive expectation” or “affective forecast” (following terminology in Wilson and Gilbert, 2005). Liking an idea involves “attendant expectations of how things will be. It thus sets us up to notice discrepancies— negative or positive— between how things seemed in the favourable representation and how we actually found them” (Railton, 2012, p. 35). If the desire is fulfilled, a comparison is made determining “whether [the desired end] is better than, worse than, or in conformity with, the affective expectation of it” (Railton, 2012, p. 36). Subsequent propensity to form affective interest in that end is then modified if the end is worse than expected, and reinforced if better than expected. *Through* affect, subsequent motivation is regulated: learning that an end is worse than affectively expected thus will no longer lead to the elicitation of the motivational or wanting component of desire.

Railton’s view of desire is plausible, in particular, the kind of regulatory role that affect plays in eliciting and regulating subsequent affect and motivation. We will revisit this suggestion in Chapter 5 and 6. For now, I want to make two points about Railton’s account.

Railton and Berridge on Wanting and Liking

The first point concerns motivation and support for Railton’s prospective model. Railton claims to arrive at his account from the armchair via “speculative, *a priori* and phenomenological means” (Railton, 2012, p. 37), but he takes his view to be supported by, or at least consonant with, work in empirical psychology. For example, his view makes reference to two components, (i) affective interest or *liking* and (ii) motivation, or *wanting*. He refers to work by Kent Berridge and colleagues who have studied the neurophysiological bases that underlie motivation and reward.

Berridge identifies what he calls “liking” (scare quotes), a pleasure-related phenomenon that occurs in ‘hedonic hotspots’ present in parts of the brain like the nucleus accumbens and the ventral pallidum (Berridge, Robinson, and Aldridge, 2009) as well as in parts of the prefrontal cortex that comprise part of the limbic system, OFC and insula (Castro and

Berridge, 2014 referenced in Berridge and Kringelbach, 2015). “Liking” is operationalised by characteristic orofacial reactions shared by many animals (Berridge, 2009, 386, Fig. 3). “Liking” is to be distinguished from “wanting” or *incentive salience*, a “percept-bound” driver of motivation controlled by dopamine-based activity in the nucleus accumbens that “occur[s] as relatively brief peaks upon encountering a reward or a physical reminder of the reward (a cue)” (Berridge, 2009, p. 379).⁶⁷ Berridge and colleagues have shown through a variety of methods that “liking” and “wanting” can be dissociated.⁶⁸

Berridge is at pains to point out that that incentive salience does not correspond to the ordinary notion of wanting, which might involve cognitive components such as the explicit or conscious representation of an end (“cognitive desires” Berridge, 2009, p. 379). Berridge’s reason for this is that incentive salience is mediated by the activity of subcortical structures whereas more cognitive forms of desire requires higher cortex-based activity. Incentive salience, Berridge writes, “*does not require a clear cognition of what is wanted, and does not even need to be consciously experienced as a feeling of wanting*” (Berridge, 2009, p. 379). It is not clear then that the motivational component of Railton’s view as focussed on explicitly represented ends, or as Railton puts it, “motivational under a favourable idea” is *exactly* accounted for solely by incentive salience processes, even if it plausibly involves incentive salience processes. We should also be careful not to think that incentive salience desires are present to phenomenal consciousness. It is plausible, or at the very least an open question whether the processes that underlie “wanting” can occur in creatures without consciousness, or whether the effects of incentive salience processes can affect behaviour without accompanying subjective feelings (e.g. Anselme and Robinson (2016), for instance, interprets the findings of Fischman and Foltin (1992) in this way). Berridge however does

⁶⁷The function of the neurotransmitter dopamine is still a matter of considerable debate. For discussion, see Berridge and Robinson (1998), Holton and Berridge, 2016, Schultz, 2016.

⁶⁸For instance, rats with damaged dopamine neurons or administered with dopamine-blocking chemicals fail to be motivated to eat in the presence of food, even when hungry despite displaying many of the orofacial reactions when artificially fed suggesting that they nevertheless “liked” the food (Berridge and Robinson, 1998). In the opposite direction, “wanting” can be induced without “liking” by electrical stimulation of the lateral hypothalamus which causes increased motivation to eat with aversive orofacial reactions indicating “dislike” (Berridge and Valenstein, 1991).

suggest that, at least some of the time, incentive salience can “color conscious desires with motivational power, to make them compelling spurs to action” (Berridge, 2009, pp. 379–80).

The same can be said about “liking”. In Railton’s account, affective interest is something that is present to consciousness, it constitutes the way in which an idea is framed favourably. However Berridge suggests that it is plausible that the subcortical processes that underlie “liking” can occur below the level of consciousness though nevertheless measurably influencing “liking” reactions (see e.g. Winkielman and Berridge, 2004).

Railton on the Rational Role of Liking

The second point I want to raise concerns Railton’s account of how desires rationalise action. As we have seen, for Railton, desires are a complex involving a motivational and an affective liking component. Whilst I agree that desires can involve an affective component, I do not want to over-emphasise the role of liking in desire. I think we should resist making this affective component *necessary* for the rational intelligibility of acting on desire, and that we should only grant that it is sufficient given a certain understanding about the function or role of affect, and why the regulatory function of affect that Railton emphasises is rationally significant.

One reason against making affect necessary is phenomenological. I am willing to concede that conscious affect can enable and contribute toward an explanation of why normal cases of acting on desire can be rational from the subject’s perspective. However, there seems to be many cases of ordinary motivation by basic desires that are not caused or sustained by the kind of positive affective state that comprises Railton’s account of desire. Take for instance simple cases of basic appetitive desire which can motivate one to take a sip of a drink. Such desires seem to rationalise action, but they can do so even in the absence of affect.

In his discussion of Quinn’s Radioman case, Railton writes that:

[The account] suggest[s] an answer to the question why Radio Man’s behaviour seems so opaque or “strange”, as Quinn puts it. For his behaviour though certainly the product of perception and

motivation, and that extent explainable, is stipulated to involve no favourable representation of what he is doing, and so it be in that sense lacking in any rationale. . . if [the account] is right, this is not ground for discounting the rationalising capacity of desire, but for thinking that desire is the element missing in Radioman’s opaque behaviour.

If Radio Man’s motivation had instead stemmed from liking the idea of turning radios on. . . things would be different. We might not think much of the idea of turning radios on just because one likes to, and therefore as a favourable view of the prospect of doing so. But if this were the case for Radio Man, his actions would raise questions of taste, not intelligibility (Railton, 2012, p. 29)

But is it really right that things *would* be different if we added to Quinn’s case that Radioman just *likes* the thought of turning on radios? I find it hard to imagine that this would improve the rational intelligibility of Radioman. For surely we would wish to know why, or at least *how*, Radioman comes to like turning on radios. The point of caution I want to raise here reiterates the criticism I made of Chang, Smithies and Weiss’ views. Affect might be rationally significant given an account of the affective system as involved in sub-personal appraisal of what is good or bad for the subject where the conscious character enables the subject to respond to those appraisals. Matters would be improved for the rational intelligibility of Radioman if some background details about how his liking to turn on radios was part of the normal functioning of his affective system. However simply appealing to the mere phenomenological character of liking— the fact that he just finds turning on radios attractive or enticing— unmoored from function is rationally insignificant for the rational significance of *liking* is as much in need of explanation.⁶⁹

Finally, there is one further feature of Railton’s account that I want to flag which constitutes a point of agreement with Chang, Smithies and Weiss. As we outlined in §3.2.3, Chang, Smithies and Weiss claim that when an agent acts on a desire to ϕ , as when he chooses the affectively desired bale of hay, “he acts for the reason that he feels like having it” (Chang, 2004, p. 82). Here it is the desire of which one can be aware

⁶⁹Railton, I should add, does say something about how the mechanisms that regulate affect are attuned to what is good (Railton, 2012, pp. 34, 41). We will return to this in Chapter 5.

simply by undergoing the affective experience in question that constitutes one's reason for action. In the same vein, Railton claims that it is our *liking* an idea, of which we are aware, that constitutes one's reason for action. As Railton writes "distinctively human desire is possible because this normally subconscious affective coding can be brought to the surface, and self-consciously thought about" (Railton, 2010, p.34-5). In acting on a desire, "[t]he rational desirer *takes the fact of desiring that R or liking the idea of R (with a certain intensity) as a default, defeasible, reason to want (with proportionate intensity) to bring R about*" (Railton, 2010, p.38). We will return to this discuss this point in Chapter 6.

Conclusion

The aim of this chapter has been to explore the relationship between affect and desire. I considered a range of views and I wish to end this chapter by summarising the central conclusions and lessons we learnt by considering these views.

The first family of views, *hedonic* accounts of the rational significance of desire, centralised *pleasure* in our understanding of desire. Whilst it is no doubt true that the satisfaction of a great many of our basic desires can yield pleasure and that many of desires, especially appetitive desires, have displeasurable concomitants, it is unclear if this link to pleasure can provide the right kind of answer to the Problem of Conative Significance.

This is for two main reasons. First, many basic desires can rationalise choice and action despite the fact that they are not discomfoting states to endure, nor does their satisfaction guarantee pleasure. Even if we consider a proper subset of basic desires, such as the appetitive, where the hedonic accounts like the one proposed in Schiffer (1976) have the greatest chance of success, we have seen that the connection between pleasure and discomfort fails to generalise. Second, and most importantly, the rational role of basic desire cannot be captured solely by appeal to the reasons there are to alleviate discomfort or pursue pleasure. Basic desires rationalise actions that lead to their *fulfilment* whereas discomfort rationalises taking steps to eliminate discomfort. Cases where it is possible to eliminate a discomfoting desire without fulfilling it shows that the rational significance of desire may diverge from the rational significance

of discomfort.

The second family of views, *affective* accounts of the rational significance of desire, centralised a form of valenced experience in our understanding of desire. I outlined one variant of such views that argued that affective experiences of desire have intrinsic introspectible features that constitute apparent reasons for action. The central objection I raised against such views is that without an account of the way in which such affective episodes relate to what is good or bad for the subject, it remains obscure how such experiences can be significant from the perspective of practical rationality.

Finally, I considered a more sophisticated account of desire proposed by Peter Railton. Railton's proposal is important because it incorporates affectivity into an account of desire and provides an account of the rational role that affect plays in regulating subsequent affective states and motivation. I take this aspect of Railton's view to be credible and will explore it further in the final two chapters. I raised two points about Railton's view as stated. First, I discussed the relationship between the two components of Railton's view (positive affect and motivation) with work by Kent Berridge who similarly distinguishes between an affective component ("liking") and a motivational component ("wanting") involved in reward-based motivation. I suggested that these two pairs do not amount to the same thing. Second, I suggested that we should not over-emphasise the affective component of desire. Though affect is an important part of the normal operation of the conative system, it is not necessarily present and, echoing the lessons of the previous section, it is only rational significant to the extent that we have an account of its role in attuning us to what is good.

Chapter 4

Desire and the Good

This chapter outlines and evaluates a promising strategy for solving the Problem of Conative Significance. The strategy taken is embodied by a family of views known as the guise of the good. According to guise of the good theories, desires are constitutively related to representations of its object as normatively favoured in some respect, for instance, as *good*.

The plan for this chapter is as follows. I begin first by providing some background concerning the guise of the good, differentiating the different kinds of claims that have typically fallen under its banner (§4.1). I clarify which versions of the guise of the good theory I will be concerned with. In the following section, I outline and evaluate a *cognitivist* version of the guise of the good (§4.2). According to this view, desires are to be identified with a certain class of normative beliefs, beliefs about *reasons*. On this view, we should explain how desires contribute to what it is rational to do by reference to the introduction of propositions about the reasons we have to act. I raise two problems for this view which motivate an alternative approach which aims to develop the guise of the good by emphasising an analogy between desire and perception (§4.3). I suggest that as it stands, the approach is not adequately explanatory and outline how one might develop a more satisfactory explanation of the rational significance of desire (§4.3.4).

4.1 The Guise of the Good

Guise of the good theories concern the nature of human action. At the heart of these theories is the idea that intentional action is a *rational* activity. Guise of the good theories of action develop this idea by claiming that rational agency involves a certain kind of normative evaluation, such that what the subject does is a response to what is thought to be, in some respect, normatively favoured. Guise of the good theories have also been framed more specifically about the kinds of states capable of motivating subjects to act, for instance desires. There the thought is that being motivated by such mental states necessarily involves taking the object of the desire to be normatively favoured in some respect.

It is important to distinguish between the guise of the good and a position in moral psychology known as motivational internalism.⁷⁰ Motivational internalism is the view that judgements about what is morally good or of what there is moral reasons to do are sufficient to motivate action. As Tenenbaum (2013) points out, whereas motivational internalism is a claim about the sufficiency of moral judgements to motivate, the guise of the good, at least on its ‘broadest’ formulations (more in a moment) is a claim about what is necessary for intentional action. Further, the guise of the good certainly does not claim that intentional action involves taking the action to be *morally* good.

4.1.1 Three Formulations

We can distinguish at least three formulations of the guise of the good.

The first formulation articulates the guise of the good without explicit reference to desire, connecting intentional action as such with normative evaluation. On this view, if a rational agent intentionally ϕ -s, then she acts on the basis of a consideration she takes to be good in some respect, or as constituting a normative reason for action (cf. Raz, 1999; Williams, 1979; Velleman, 1992; Setiya, 2010). As Raz puts it, “choice and decision are subject to rules of rational constraint. . . one can only choose or decide for a reason, i.e. for what one takes to be a good reason for the option chosen” (Raz, 1999, p. 8).

⁷⁰See Björnsson et al. (2015).

This formulation of the guise of the good is sometimes framed as a ‘formal’ constraint on human agency, one with a neo-Aristotelian character. Boyle and Lavin (2010) have recently advocated this interpretation of the guise of the good as a claim about how we have to understand what is distinctive about human action, a form of *self*-movement distinct from non-agential events like the falling of a branch, or other kinds of teleologically structured movement like the growing of a plant. On this view, the guise of the good thesis is a thesis about what is distinctive about rational action where this is understood as involving the exercise of a subject’s capacity to act on the basis of her take on what is good (Boyle and Lavin, 2010, pp. 186–7). Insofar as the guise of the good applies to *desire*, it merely articulates a condition on the kind of contribution that desires must play in giving rise to rational action: desires must bear on this process of acting on the basis of one’s take on the value of acting.⁷¹

There are two alternative formulations that explicitly makes reference to *desires*.⁷²

The second of the three formulations is expressed by Kieran Setiya as follows “[i]f someone acts intentional in doing ϕ , she is acting on a desire, and desires represent their objects as good” (Setiya, 2010, p. 84). This formulation connects three things: intentional action, desires and normative evaluation. Here I want to highlight the notion of desire at play here, which is plausibly understood as expressing desire in the pro-attitude sense since it claims that necessarily one acts on a desire when one acts intentionally (cf. §2.1.2).⁷³

The last of the three formulations focuses squarely only what I have

⁷¹In Boyle (2016), we see him take this formal neo-Aristotelian constraint further in an account of desire in the rational agent on which desires must present their objects as *desirable* if desires can rationally motivate action.

⁷²For this reason, these two formulations cleaves closer to the Scholastic slogan “quidquid appetitur, appetitur sub specie boni”, whatever is desired is desired under the guise of the good.

⁷³Setiya (2010, p. 86) claims that the first formulation can be derived from the second. When one represents a desired object as good, one represents a respect in which that object is good: a good-making feature of that object. For example, when one intentionally takes out the rubbish, the claim is that one represents taking out the rubbish as say, decluttering, the respect in which it is good to take out the rubbish. This good-making feature of ϕ -ing amounts to a *reason* for ϕ -ing, indeed, it would be one’s reason for ϕ -ing. So we can derive the first ‘formal constraint’ formulation from this one via the principle that the pro-attitudes of a rational subject involve representing good-making features toward the act or outcome to which one has the pro-attitude.

reserved the term ‘desire’ for. This subclass of guise of the good theories shares the insight that desires are in some way connected with an agent’s evaluation of her actions as normatively favoured in some respect. However, in being focussed on ‘desire’ more narrowly construed, it is focus less on rational agency *per se* and more on the rational role of a restricted class of pro-attitudes.⁷⁴ It is this third formulation of the guise of the good theory that I want to focus on exclusively. I will call these *desire-based* guise of the good theories.

4.2 Belief as a Model for Desire

In this section, I outline and evaluate a *cognitivist* version of the desire-based guise of the good theory. The central commitment of the cognitivist about desire in general is that desires consists in a *truth-evaluable representation* of its objects as having a certain property. When the cognitivist position is incorporated into a desire-based guise of the good theory, it amounts to the claim that a desire to ϕ consists in a truth-evaluable representation of its objects as normatively favoured. For example, to have a basic desire to eat strawberries is to represent eating strawberries as normatively favoured in some respect. David Velleman describes such an account as follows:

Proponents of this . . . strategy portray motivation itself as an inference, governed in part by action-justifying content to be found in the motivating attitudes. To this end, they incorporate the valence of desire into its content, by describing desire, not as a favourable attitude toward the representation of some outcome, but rather as an attitude toward a favourable representation of the outcome.

Here, then, is one way in which rational agency comes to be conceived as a capacity for pursuing value. Desires are conceived as value judgments, with intrinsic justificatory force, so that the desire motivating the agent can be identified with the reason guiding him. The result is that all actions performed for reasons are conceived as arising from favourable value judgments, and hence as being aimed at the good. (Velleman, 1992, pp. 6–7)

⁷⁴For discussions of specifically desire-based guise of the good theories, see Hawkins (2008), Stampe (1987), Schafer (2013), Schapiro (2009), Gregory (2016), Boyle (2016), Boswell (2016).

As stated, there are two issues with respect to which this theory is underspecified. The first issue concerns clarifying *which* representational state we look to in the theoretical domain as our model of desire. The two main contenders in the literature are belief and perception, both arguably states that involve representing the world to be a certain way.⁷⁵ The second issue turns on the normative property to which one appeals in formulating the view. So far we have considered the view that desires involve taking a state of affairs or a course of action to be good. The threat of certain seeming counterexamples involving cases of bad, evil, perverse or whimsical desires (cf. Anscombe, 1957; Raz, 2010; Stocker, 1979; Hawkins, 2008) motivated the rise of alternatives that construe the relevant representation as one involving contents about what there is reason to do (Gregory, 2013; Gregory, 2016; Gregory, 2018) or what one ought to do (Lauria, 2017).⁷⁶

In order to have a fixed target for discussion, I want to consider a version of the cognitivist desire-based guise of the good theory as defended in the work of Alex Gregory. What is distinctive about Gregory's view is that he provides an account of desire as playing an analogous role to belief in practical reasoning and the rationalisation of action; indeed, his view is that desires just *are* a class of normative beliefs. I will begin by introducing the essentials of Gregory's view (§4.2.1) and elaborate on how the view explains the rational significance of desire in a way that satisfies the two constraints articulated in Chapter 2. I end by introducing two problems for the view which I think provide ample motivation to consider an alternative way of developing the guise of the good views (§4.2.2).

⁷⁵Within these views, there are differences concerning how to understand the relationship between desire and the attitude we take as our model. For instance, whilst views that look to belief have been tempted to *identify* desires with a certain kind of belief with normative content (e.g. Gregory, 2016), views that look to perceptions have instead emphasised a kind of *analogy* or *isomorphism* between desire and perception (Tenenbaum, 2007; Schafer, 2013).

⁷⁶Somewhat awkwardly I will retain the phrase 'guise of the good' to refer to all formulations that might differ on which normative property to invoke. Given that my emphasis will not be on this question, this should not matter too much.

4.2.1 The View

The rational role of belief is to introduce reasons into our psychology. When all goes well, our believing a certain proposition puts us in a position to act for reasons that would thereby rationalise our choices and actions. If we take belief as a model for desire in explaining the rational significance of desire, we will be led to think that in desiring that p , desires similarly introduce reasons that they represent into our psychology. But what do desires represent? The desire-based guise of the good theory claims that desires represent acts or states of affairs as *normatively favoured* in some respects, for instance, as *good*. If desiring to drink a beer involves representing one's drinking a beer as good in a certain respect, then we can begin to see how a desire would put us in a position to act on the basis of a reason, namely that respect in which drinking a beer is good.

Taking this approach, Alex Gregory defends the view that desires are to be identified with a certain class of normative belief.⁷⁷ As mentioned, there is some theoretical latitude in which normative property we identify as comprising part of the content of the normative beliefs in question. Here I will focus primarily on Gregory's account of desire being under the guise of *reasons* as opposed to the good, which Gregory has argued is the most promising way to proceed (Gregory, 2013). According to Gregory:

[**Desire as Belief (DAB)**] To desire to ϕ is to believe that you have normative reason to ϕ . (Gregory, 2016, p. 201)

According to Gregory, the DAB thesis explains the rational significance of desire in terms of the rationality of responding to our normative beliefs, i.e. beliefs about what reasons we have to act. It is plausible that rationality requires us, *inter alia*, to respond to our *de dicto* normative beliefs.⁷⁸ This requirement of rationality is typically captured by what is called an 'enkratic' principle. On Gregory (2018)'s favoured interpretation of the enkratic principle, "[r]ationality favours ϕ -ing to the

⁷⁷See Gregory (2016), Gregory (2018); cf. Campbell (2018)

⁷⁸Whether it is *exhausted* by responsiveness to de dicto normative belief is another matter. I briefly questioned this in Chapter 1.

extent that you believe you have reason to ϕ " (Gregory, 2018, p. 1072).⁷⁹ So if you believe that you have a reason to go to work, then rationality favours, to some extent, your going to work. According to Gregory, DAB can yield a neat explanation of the rational significance of desire by dovetailing with such a principle. Given DAB's identification of desire with normative belief, we can explain why desiring to ϕ can rationalise one's ϕ -ing in terms of the fact that rationality favours, to some degree, your ϕ -ing if you believe you have reason to ϕ .

4.2.2 Problems for Desire as Belief

In this section, I want to raise two problems for the DAB thesis. The first problem targets the problematic identification of all desires with a certain kind of normative belief. I will argue that features of certain desires, specifically occurrent basic desires, cannot be explained if desires are normative beliefs. This gives us reason to think that DAB fails to apply to such desires. The second problem is a worry that DAB distorts a natural understanding of the kinds of considerations in light of which we act when we act on a desire in a way that leads to an overly intellectualised account of the rationality of acting on a desire.

Problem 1: Basic Desires are not Normative Beliefs

The first problem I want to raise for DAB concerns whether it is able to provide an account of the kinds of desires I have been concerned with, basic desires. In treating basic desires as a class of *belief*, it mischaracterise basic desires by locating their source in our capacity to reflect on reasons for action. In order to build the case for this claim, I wish to revisit some of the differences discussed in Chapter 2 between basic desires and other psychological attitudes like belief and intention.

To start, I wish to outline some features of belief that I think are largely uncontroversial. Beliefs are what a subject holds to be true. Rational subjects form beliefs on the basis of evidence for or against. They are what Scanlon calls "judgment-sensitive attitudes" (Scanlon, 1998, p. 20): beliefs are regulated by the subject's take on the evidence

⁷⁹Compare Broome (2007) and see Gregory (2018, p. 1072) for an argument to motivate this version of the enkratic principle.

for or against the belief such that the judgements of reasons that count decisively against the belief would lead the rational subject to abandon the belief. Given that beliefs are judgment-sensitive, then, if Gregory's view is right, we should expect a rational subject to abandon her desires in response to judgements that count against her belief that she has a reason to ϕ .

Yet, as many have pointed out in the literature, many of our desires can persist despite countervailing judgements. There are a range of cases in the literature to illustrate the recalcitrance of desire to normative belief. For instance, take a dieter who desires to eat a cake against his better judgement that he should not eat the slice of cake offered to him (Schiffer, 1976, p. 195; Chang, 2004, p. 67). Or consider Scanlon's example of someone with a desire to buy the latest model of computer which can persist despite "believ[ing] (correctly, let us suppose) that the features of the newer models would be of no real benefit to me" (Scanlon, 1998, p. 43). We might thus try to mount the following argument against DAB:

The Argument from Recalcitrance

- P1. If DAB is true, then agent A wants to ϕ iff A has a belief that she has a reason to ϕ .
- P2. But there are cases in which A can want to ϕ despite lacking a belief that she has a reason to ϕ .
- C. So, DAB is false.

How might Gregory respond? The most promising route for Gregory to pursue would be to contest (P2) by denying that one lacks a belief that one has a reason to ϕ . Is it right that the dieter sees no good in eating the cake? It is a plausible interpretation of the case that the dieter believes that there are reasons that strongly, perhaps decisively, count against eating the cake. But this is consistent with the belief that there is *some* reason to eat the cake, perhaps that doing so will bring him pleasure. Similarly, Gregory can agree with Scanlon in his discussion of the computer case that he believes the features of newer models are of no real benefit, yet contest that this amounts to the belief that there is *no* reason to buy the new computer.

In order to defend the argument against DAB from this reply, it will be necessary to articulate a case that allows us to control for the contestable element. Consider the following case:

Cake (t_1): At time t_1 , Tom finds nuts extremely tasty and is presented with a cake. What strikes him is the cake's lavish peanut coating; the sponge interior is dry and uninspiring and Tom is indifferent to the cream topping, not being a lover of dairy products. Tom forms an occurrent desire to eat a slice of the peanut cake.

So far, there are no reasons to think that Gregory would deny that the desire in this case can be analysed according to DAB as involving a belief that there is a reason to eat the peanut cake, where the reason in question is that it is coated in peanuts. Now consider the following elaboration of the case:

Cake (t_2): At time t_2 , a scientist has altered Tom's physiology to induce an overwhelmingly unpleasant reaction upon contact with peanuts. The honest scientist informs Tom of his unfortunate condition and, taking him at his word, Tom thinks that the fact the cake is coated with peanuts now counts *against* eating the cake. Nevertheless, looking at the cake and the mouthwateringly delicious peanut coating, he still feels inclined to eat the cake.

This case is raises the same problems as the cases previously considered, but is stipulated to make it difficult to interpret the case in such a way that the subject retains the normative belief she had at time t_1 . It is plausible that at t_2 , Tom *abandons* his belief that there is a reason for him to eat the cake, a perfectly rational response to the scientist's testimony and his indifference to all other non-nut-related features of the cake. This case is problematic for DAB for two reasons. First, given the account DAB gives of Tom's desire at t_1 , DAB predicts that Tom would count as losing his desire to eat at t_2 . Yet it is plausible to suppose that Tom *retains* a desire to eat some of the peanut cake. Second, given the retention of Tom's desire, it cannot be *necessary* for Tom's desire for the cake that he believe there to be a reason to eat the cake. What we have here is exactly the case of a desire that is recalcitrant to shifts in the subject's assessment of reasons for action.

Is there a further response that Gregory can make to resist this counterexample? One possible response would be to deny that Tom has a desire to eat the cake. There are two ways of interpreting the case consistent with this strategy. The first way would be to deny that Tom has a desire to eat cake at both times t_1 and t_2 . The second option would be to deny that Tom has a desire at time t_2 after the normative belief is eliminated. But what considerations would motivate either interpretation? Here it would be instructive to look for precedent in how Gregory responds to structurally similar counterexamples. In Gregory (2016), Gregory considers a case in which a heroin addict wants to shoot up but believes she has little, if not no reason at all to shoot up. Here Gregory suggests a response like the one considered claiming that:

[T]o the extent that [the addict] is motivated to take heroin by a state of mind which is completely irrational and insensitive to facts about what she has reason to do, we might think that she is being motivated not by a desire, but instead by some more primitive compulsion or drive (Gregory, 2016, p. 210).

Notice that there is a danger of begging the question against those who think that the cake case constitutes a counterexample. Those who seek to deny that the appetitive inclination Tom feels is a desire cannot do so simply *because* it is recalcitrant to normative belief. So they will have to appeal to other features. Reflecting on the addiction case where Gregory opts for the desire-denying response is of limited help, for it is not credible to apply the same reasoning in the cake case given that strong desires for peanut cake are quite unlike addictive desires.⁸⁰ In such cases, we might grant Gregory's point that such motivational states might be experienced, not as desires, but as something else: unintelligible urges, cravings or compulsions.⁸¹ However, we need not understand

⁸⁰A plausible view of how substances like heroin work is by causing a powerful boost in dopamine levels that causes strong occurrent incentive salience-based motivation ("wanting") that drives the formation of disproportionately strong standing desires for the substance (cf. Berridge and Robinson, 2016; Holton and Berridge, 2016). Often such addictions induce motivational states that require substantial amounts of self-control to resist and have been substantially decoupled from pleasure and a subject's assessment of reasons for action.

⁸¹It is not obvious that we even should grant this point; to maintain this point, Gregory must provide a non-question begging argument to think that such urges or cravings are not just cases of powerful unusual cases of desire.

Tom's desire at t_1 - t_2 as sharing any of the exaggerated properties of addictive desires. Hence, it is, at the very least, inconclusive whether reflection on features of addictive desires license making the same response in the cake case.

Problem 2: Does the Account Over-Intellectualise Rational Support?

Gregory's account of how desires rationalise action rests on two claims. The first claim is his cognitivist account of desire; namely that a desire to ϕ rationalises action because a desire to ϕ is a belief that one has a reason to ϕ . The second claim extends this account of desire to explain how desires rationalise action: a belief that one has a reason to ϕ contributes to what it is rational for one to do because rationality favours a course of action to the extent that one believes oneself to have reason to pursue that course of action.

Several philosophers have put pressure on the first claim, arguing that to construe desires as normative beliefs would objectionably over-intellectualise desire (cf. Velleman, 1992; Hawkins, 2008; Saemi, 2014). As Velleman argues, such a view "implies that the capacity to desire requires the possession of evaluative concepts. Yet a young child can want things long before it has acquired the concept of their being worth wanting or desirable" (Velleman, 1992, p. 7). The objection I want to press targets the second claim. My worry is that Gregory's cognitivist explanation of how desires rationally support action is objectionably over-intellectualised.

I want to start by bringing out a puzzling feature of the proposed explanation, one concerning the rational grounding of the normative belief. What is the reason in question that one believes oneself to have? Presumably given that normative beliefs like these are formed for reasons, it is a precondition on one's wanting to ϕ that there is an apparent reason R which is the reason for one's belief that there is a reason to ϕ .⁸² Notice then that Gregory's view then introduces a separation between the agent's reason for acting when she acts on a desire to ϕ and

⁸²I am not clear what an alternative to this would be; denying this is unattractive since it is not as if one can rationally form the bare belief that one has some reason to ϕ , yet have no understanding or awareness of the reason in question.

that which explains the rational significance of the agent's desire. Take Uma's desire to eat some strawberries. On this story, if Uma were to act on such a desire, her reason for eating these strawberries would be some apparent reason she has to eat the strawberries, for instance their delicious flavour. What makes it rational for her to act on her desire is not the apparent reason she has, but the *belief* that she has a reason to pursue a certain course of action, viz. the strawberries having a delicious flavour.⁸³

The problem with this explanation is that it makes it a necessary condition for desires to rationally support choice and action that the subject have the capacity to form normative beliefs; that is, beliefs with the content that certain states of affairs constitute reasons for action. This would imply that only the desires of subjects with the capacity to form beliefs about normative reasons can rationally support choice and action.

I take this to be an unattractive consequence of the view because it introduces an unmotivated gulf between the desires of subjects with the sophisticated capacity to form beliefs about reasons from those that cannot. Many of the intuitions that led us to think that desires can rationally support action apply just as naturally to children without the capacity to form assessments about the kinds of considerations that justify particular courses of action. Take young Uma who forms an occurrent desire for the strawberry ice cream at the ice cream parlour. To the extent that we found it natural to think that such desires can rationally support the choice of a conceptually more sophisticated subject, it seems unmotivated to deny desires that power in the case of young Uma. As I will argue in Chapter 6, it is possible to explain the rational significance of desire without requiring that one form beliefs about one's reasons *qua reasons*.

⁸³This view is very close to Quinn's own view of "[the missing] element of desire that does the rationalising" (Quinn, 1994b, p. 246). Quinn's view is that rationalisation requires a certain kind of normative belief, "the thought that the direction in which I am psychologically pointed leads to something good (either in act or result) or takes me away from something bad" (Quinn, 1994b, p. 242). The present view thus approximates Quinn's preferred explanation in the language of reasons, rather than goodness. One point of difference is worth noting. Though Quinn (1994b) at points talks about this normative thought as what is missing *in* desire, his central message need not be read as making a claim about the nature of desire *per se* as opposed to the cognitive concomitants that must accompany one's pro-attitudes *if they are to rationalise*.

I will argue that the rational significance of desire can be captured by the role they play in allowing us to respond to a certain class of reasons without having to suppose that they consist in normative beliefs about reasons.

4.3 Perception as a Model for Desire

In light of the Argument from Recalcitrance, many have emphasised a similarity between desires and perceptual appearances. Consider for instance the Müller-Lyer illusion. Even after subjects come to learn that the two parallel lines are of equal length, it still appears to subjects that one of the lines is shorter than the other. The standard reason for why this is possible is that perceptions are to a sufficient degree ‘informationally encapsulated’ from, and cognitively impenetrable to, information in more cognitively ‘centralised’ systems.⁸⁴ This has led some attracted to the guise of the good to consider the view that a basic desire that p involves its appearing to one that p is F , where to be F is to be normatively favoured in some respect.

4.3.1 The Content View

Schroeder (2008b) clarifies two models that perceptual views can take. According to one model, which I will call the *Content View*,

[J]ust as the perceptual experience of grass represents it as green, thus prima facie licensing the belief that it *is* green, the desire to drink coffee represents drinking coffee as good thus prima facie licensing the judgement that drinking coffee is good (Schroeder, 2008b, p. 121).

Tim Scanlon develops a version of the Content View which takes the notion of a reason for action as the central normative notion. Scanlon introduces his account in order to explain how we can explain the recalcitrance of basic desires in a way that is consistent with denying that such desires are “special source[s] of motivation, independent of our seeing things as reasons” (Scanlon, 1998, p. 40). I will argue that although Scanlon’s view starts us down a promising path by emphasising

⁸⁴For discussion, see Robbins (2017) and Fodor (1983).

an analogy, not between desire and belief, but with perception, as stated, Scanlon's view remains susceptible to the problem raised for DAB.

Scanlon begins with a case of an appetitive desire to drink when thirsty which he claims begins with an unpleasant sensation of dryness in one's mouth or throat (Scanlon, 1998, p. 39). Being in such a state motivates one to seek out a drink, but it is a mistake, Scanlon claims, to think that these appetitive desires have a source independent from the capacity that undergirds our ability to form judgements about reasons for action. This is because being in such a state motivates by engaging our capacity to recognise considerations *as reasons for action* by directing our attention toward considerations (e.g. that a drink will bring relief or pleasure from discomfort) that *seem or appear to count in favour of* acquiring a drink (Scanlon, 1998, p. 65). These appearances are not evaluative judgements or beliefs (Scanlon, 1998, pp. 39, 65) but quasi-perceptual precursors to the formation of those judgements (Scanlon, 1998, p. 65). Unless one judges otherwise, one is in a position to form the *belief* that the future relief is a reason to acquire a drink which can then motivate. Thus, "the motivational work", Scanlon claims "seems to be done by [one's] taking this future pleasure to count in favor of drinking" as opposed to a motivational force—i.e an unmotivated desire—that operates independently of one's "tendency to see something good or desirable about [the object of desire]" (Scanlon, 1998, p. 38)

Scanlon thus defends the following variant of the Content View:

Content View (Reasons) to desire to ϕ is for it to appear to one that some property F of one's ϕ -ing constitutes a reason for one to ϕ .

How does Scanlon explain the kind of cases of recalcitrance we have considered? As we have seen, Scanlon thinks that desires involve dispositions for one's attention to be "directed insistently toward considerations that present themselves as counting in favor of P" (Scanlon, 1998, p. 39).⁸⁵ Scanlon distinguishes between *judging* reasons and *seeing* reasons. Seeing reasons involves a more or less insistently capturing of our

⁸⁵I want to flag one issue concerning the way Scanlon glosses such attention-directing desires, specifically, the kinds of things that having a desire directs your attention to. He claims that "A person has a desire in the directed-attention sense that p if the thought that p keeps occurring to him in a favourable light, that is to say, the person's attention is directed insistent toward considerations that present themselves as counting in favor of P" (Scanlon, 1998, p. 39). In the first part of this quotation, to

attention by considerations that “present themselves as” or ‘appear to be” reasons for action. According to Scanlon, in cases of recalcitrant desire, one *sees* certain features F as a reason for action contrary to a settled *judgement* that F is not a reason. To illustrate this, he considers the following case:

One can have a strong and recurrent tendency to see something as a reason for acting (under one’s present circumstances) even though one’s firm considered opinion is that it is not (under the circumstances) such a reason. This is clear not only in cases in which a person acts irrationally, but also in many other cases. Even if, for example, I have convinced myself that I should not be influenced by the approval or disapproval of a certain group, I may find myself wondering anxiously what they would think of something I am considering doing. When these thoughts occur, I may dismiss them immediately. Nonetheless, insofar as they involve (perhaps only momentarily) seeing something as a reason that I judge not to be one, they are instances of irrationality— a form of irrationality to which we are all subject from time to time. Even when desire in the directed- attention sense runs contrary to our reason (that is to say, our judgment) in this way, however, it remains true that the motivational force of these states lies in a tendency to see some consideration as a *reason*. (Scanlon, 1998, p. 40)

How compelling is Scanlon’s account when applied to basic desires? In the following section, I outline several pressing problems for Scanlon’s view concerning the substantivity of his distinction between judging and seeing reasons.

have a directed attention desire that P is for the state of affairs in which you get what you want (i.e. that P) to occur to you in a favourable light. However, he unpacks this metaphor by claiming that what your attention is directed “toward [are] considerations that present themselves as counting *in favor of P*”, so the variable P seems then to pick out propositions as well as *act-types* in order for this sentence to be well-formed. Later in discussion of his computer example, Scanlon focuses on *features* of what you want e.g. having a faster processor, which you might take, or indeed see, as a reason to do something, e.g. to buy the computer you judge you have no reason to buy. We can neaten the proposal as follows: A subject has a directed attention desire to ϕ if there is some F such that the fact that a is F presents itself to the subject as counting in favour of ϕ -ing.

Problems: Discontentment with the Content View

The first problem I wish to raise concerns whether or not Scanlon's account has done enough to rule out that the cases of desiderative recalcitrance considered can be subsumed as cases in which we *see* reasons where we know there are none. What Scanlon has given us is another way of interpreting what is going on in this case. However just because it is *possible* for there to be a version of the cake case as one in which the subject has an insistent tendency to 'see' the peanut coating of the cake as a reason does not rule out that there are more natural versions of the case in which the subject lacks a tendency to see reasons where there are none. Scanlon has given us no reasons to think that we cannot coherently interpret the subject as desiring to eat some of the peanut cake *without* this insistent tendency to see reasons. The worry then is whether or not we have simply changed the case we are considering (or at least, narrowed our focus to a less problematic form of the case consistent with the original description).

The second worry concerns how well motivated Scanlon's distinction between judging and seeing reasons is.⁸⁶ We have to be careful that we are not simply labelling the psychic surplus in the recalcitrance cases in such a way that avoids conceding that these are the products of a motivational source distinct from our capacity to form judgements about reasons, i.e. not as the influence of an occurrent desire, but as 'seeing reasons'. So for this not to be a superficial redescription of the cases described, Scanlon has to say something about the nature of seeing reasons, in particular how our capacity to see reasons relates to our capacity to form judgements about reasons. The problem is that Scanlon does not tell us very much about what it is to see something as a reason.

Finally, Scanlon's view still faces the worry concerning an over-intellectualised account of rational support. According to Scanlon, a desire to ϕ involves its seeming to one that certain considerations pertaining to the object of desire are reasons for action. As Scanlon is explicit about this, writing in his discussion of the desire for a new computer that "[such a desire] has *clear normative content*, since it involves a tendency to judge [/ or 'see'] *that I have reason to buy a new computer*" (Scanlon, 1998, p. 43). On

⁸⁶For more discussion, see Schapiro (2009).

this view, the rationalising force of desire flows from a subject's *endorsing* appearances to the effect that a certain feature purports to count in favour of his acting in a certain way.

I want to register a worry here that parallels Velleman's worry with the cognitivist account of desire. If the capacity to host an appearance as of a consideration's being a reason requires that the subject exercise the concept of a reason, then this would imply that subjects without the concept of a normative reason cannot desire. A similar problem arises with respect to how such normative appearances rationalise choice and action. If desires provide rational support only to the extent that they consist in normative appearances that one then endorses, then this would introduce a gulf between the desires of subjects who possess the conceptual sophistication to host normative appearances and those who cannot. This is problematic if, as I have suggested, a subject's decision to pursue a course of action can be rationally supported by a desire even if one was not motivated by an endorsement of the course of action as supported by reason-giving considerations.

4.3.2 The Attitude View

The problems raised for Scanlon's view exploits the fact that the Content View construes the relevant perceptual experience as one in which it appears to one that some course of action is normatively favoured in some respect. Many who are attracted to developing an account of desire along the model of perception have questioned whether the Content View is the best way to pursue the perceptual analogy.

Enter the Attitude View. The Attitude View is still a desire-based guise of the good theory and so maintains that desires involve a form of normative evaluation. The central feature that differentiates the Attitude View from the Content View is the insight that we should not construe the relevant normative evaluation as involving a representational attitude with respect to a *normative* content. Defenders of the Attitude View claim that this approach is more faithful to the analogy with perception since ordinary perceptions do not have contents of the form '*p* is true'; for instance, when one perceives that it is raining, the content of that perception is simply *that it is raining* and not *it is true that it is raining*

(Schafer, 2013, p. 269). Similarly, in desiring that it stops raining, the content one is related to is not the proposition *that it would be good if it stops raining* (assuming for the moment that *goodness* is the relevant normative property), but simply *that it stops raining*.

However, if the relevant normative property is not part of the content of the attitude, then what *is* the relation between desires and the relevant normative property? Advocates unite around the following analogy:⁸⁷

Guiding Analogy for the Attitude View: Desires involve a constitutive *good*-involving relation to their contents in the same way that belief or perceptions involve a *truth*- or *veridicality*-involving relation to their contents.

By making the relevant normative property part of the *relation* that the subject stands in to the content in question, we no longer have an account on which desires involve representations as of certain courses of action possessing normative properties. It is this feature that one of the problems raised earlier targets. The problem of over-intellectualisation for example exploits the fact that a subject can desire to ϕ even if she is incapable of entertaining thoughts involving normative concepts like REASON, GOODNESS, OUGHT and so on. Given that it is possible for one to perceive that p even if one lacks the concept of truth, the Attitude View claims to sidestep this problem (Schroeder, 2008b, p. 123).

4.3.3 Problem: The Attitude View and Explanatory Adequacy

By moving away from the Content View, some of the objections that posed deep problems for the Content View are defanged. But this departure has a significant cost: the Attitude View is no longer able to employ the same explanation of the rational significance of desire. This explanation is one that the Content View shared with DAB, according to which a desire to ϕ contributes to the rationality of ϕ because the desire involves either the belief that one has a reason to ϕ (DAB), or its seeming to one that one has a reason to ϕ (The Content View). The Attitude View will thus have to provide some alternative explanation.

⁸⁷In a critical discussion and interpretation of Tenenbaum (2007), Schroeder (2008b) calls this the “true” model of interpreting Tenenbaum’s view.

The central problem I want to raise is that extant versions of the Attitude View are simply too underdeveloped to provide a satisfactory alternative. To see this, I want to begin by considering how extant versions of the Attitude View explicate the relevant constitutive relation. What is the good-involving relation to a content that p constitutive of a desire that p ?

One idea might be that desires are mental states with a certain correctness condition. On this view, a desire that p is *correct* just in case it would be *good* if p . However, as Schroeder (2008b) and Baker (2014) point out, this suggestion is inadequate because being subject to a correctness condition is not sufficient as an explication of this constitutive relation that the defender of the Attitude View is after. After all, a Humean who, *contra* the Attitude View, denies that desires are under the guise of the good can coherently maintain that desires have correctness conditions. For instance, if one were to hold a view on which what is good is a function the desires of an idealised and fully informed rational agent, then a desire can be correct just in case their objects are consistent with those of an ideally rational desirer (Baker, 2014, p. 5).

This point is conceded by at least one advocate of the Attitude View (Tenenbaum, 2008). To address this problem, Tenenbaum gestures at a deeper connection between belief and truth. Tenenbaum asks us to consider Moore-paradoxical assertions of the form “I believe that p , but it is not true that p ”. Tenenbaum points out that

[T]he impossibility of forming beliefs that blatantly contradict what we take to be true, or of forming beliefs in light of overwhelming evidence to the falsity of their contents, suggests that truth is not just a norm of correctness for belief, but that believing p involves holding p to be true. (Tenenbaum, 2008, pp. 133–4).

Tenenbaum extends this thought to desire by claiming that an analogous relation holds between desire and goodness. There are several problems with this. One especially pressing problem is that the data supporting analogous Moorean absurdity in the conative case is considerably less compelling than in the doxastic case. Claims like “I want that p , but it does not seem good if it were the case that p ” seem far

more acceptable than their doxastic counterpart.⁸⁸ The issue I want to focus on however pursues the broad objection to the Attitude View I have been developing. It still leaves us without an *account* of this more ‘internal’ relation by which beliefs relate to truth, that is, of the obscure relation expressed by the placeholder construction, ‘holding-to-be-true’. Similarly, we lack an account of the analogous construction ‘holding-to-be-good’. Why is this a problem? It is a problem if we are aiming to provide an explanation of the rational significance of desire by appeal to these relations.

Karl Schafer pursues exactly this style of explanation, modulo the relevant normative property. Schafer rejects the view that desires rationalise action and choice because “when one desires something, one has a perceptual experience, part of the *content* of which is that the object of the desire ought to be” (Schafer, 2013, 269, emphasis added). A central problem with this is that it breaks with the analogy with perception. Schafer argues that perceptions rationalise beliefs not simply on account of the content they present to the subject, but rather, “because, in some sense, [perception] presents *p* to the subject as something that is true” (Schafer, 2013, p. 269). Instead, Schafer appeals to “the manner in which [desires, like perceptions,] present their content to the subject”, their Fregean ‘force’ (Schafer, 2013, p. 270). According to Schafer, we can explain why desires rationalise the formation of intentions by appeal to the fact that “when one desires to *A*, one is in a mental state that has the same force and content as an intention to *A*” (Schafer, 2013, p. 276). A desire “presents [a course of action] *A* with imperatival force [i.e.] presents *A* to me as something I ought to do” (Schafer, 2013, p. 276). The role of desire in practical reasoning then is to present actions as ones that one ought to do, where the subsequently rationalised “intention to *A* is simply an endorsement of the way [desire] already presents *A* to the subject” (Schafer, 2013, p. 275).

The problem however is that the explanatory power of this answer to the Problem of Conative Significance is limited. This is because the central explanans, “ways of presenting content”, is theoretically obscure and therefore ill-suited as the ultimate explanation of the rational significance of desire. To see this, we can probe at how the view draws the

⁸⁸For more discussion, see Baker (2014).

distinction between force and content. What is the difference between standing in the presenting-as-true relation to a content *that ϕ -ing is good* and standing in the presenting-as-ought-to-be-done relation to the content ϕ ? Extant attitudinal theories do not equip us with the resources to explain this difference because they fail to provide an account of the force-content distinction. This is no trivial matter. Being able to draw a principled distinction between these two cases is the central explanation of the rational role of desire. We have no clearer grip on the notion of “ways of presenting content” than of how desires rationalise intentions. We cannot then hope to explain the latter in terms of the former.

4.3.4 The Functional Isomorphism with Perception

In this section, I want to briefly introduce the core idea that will inform the broad contours of the account I aim to develop in the following chapters. I will start by revisiting the functionalist construal of belief. The idea I want to pursue here is that the sense in which beliefs involve a ‘representing-as-true’ is determined by the functional role that belief representations play. To simplify matters, I will consider a hybrid framework that combines functionalism with representationalism. According to this view, a belief that p and a desire that p are distinct relations to a token mental representation that means that p , where what makes the mental representation that p either a belief or desire will be determined by the *functional role* that the representation plays in the psychological economy of the subject.

Coordination of Representational Function

What aspects of the functional role of a belief qualifies it as a *representing-as-true*? There are two features of the functional role of belief that we could appeal to. First, there are what we can think of as the ‘forward-facing’ features of the functional role. According to the motivational functionalist, beliefs play an essential role in guiding action. A belief that p is a state in which the information that p guides action, disposing the subject to behave in ways that would satisfy her total pro-attitude state in a world in which p , that is to say, it is to act *as if P*. Second, there are ‘backward-facing’ features of the functional role. Most functionalists

treat beliefs as a type of *indicator* state in the sense that in ideal conditions, a subject believes that p only if p (e.g. Stalnaker, 1984; Dretske, 1988; Stampe, 1987). One's beliefs come to carry the information they do as a result of dovetailing with one's perceptual faculties which are apt, in some sense to be elaborated, to serve belief-forming processes with information about the world.

I wish to focus on the second of these two features, in particular on what it means to say that perceptual states are 'apt' to to serve belief-forming processes. The relevant notion of aptness here is not *merely* causal, in the sense that perceptual states tend, as a matter of fact, to *cause* one to form certain beliefs. Intuitively, the nature of belief and perception are more intertwined than that. Perceptions are, in their nature, inputs into a process in which they constitute an explanatory element of the 'right kind' of way for one to form beliefs. They are parts of a subject's naturally endowed system for forming beliefs that guide action.

How do we explicate the loose talk of perceptions being one of the 'right ways' to form beliefs? It is plausible that perceptions are apt to serve belief-forming processes because they share a common core *function*: to represent veridically. Following Burge (2003), call this their *representational function*. Here is where I think we can make progress on talk of 'sameness of force' (Schafer) or 'sameness of ways of presenting content' (Tenenbaum) which in my view we can fruitfully cash out in terms of *coordination of function*.

For this, we need to adopt a view on which mental state types are individuated, not by their causal roles, that is, what they are caused by and disposed to cause, but by their teleo-functions, that is, what they have the function to be caused by and to cause. A perception that p constitutes a good basis for the formation of a belief that p . This is because, as with belief, perceptual states both have the function to represent veridically.⁸⁹

⁸⁹This is not to say of course that a perception that p *inexorably* causes the formation of a belief that p . This is because the process by which a perception that p leads to the adoption of the belief that p (or the cessation of the belief that not- p) involves the broader employment of rational faculties, e.g. rational sensitivities to defeaters and so on.

Coordination of Practical Functions

Within this framework, an attractive solution emerges which I will roughly sketch. We should think of desires as having a *functional role* that is isomorphic to the functional role of perception. Specifically on the teleo-functional view advocated, we will pursue the idea that our desire-producing mechanisms have the function to produce desires for ends conducive to the good of the organism. (Further discussion to follow in Chapter 5 and 6.) Here we capture the isomorphism between desire and perception by appeal to coordination of these practical functions. As I will shortly elaborate, by discharging its practical functions, desires put our systems for planning and intention in the best position to discharge their function to attain practical goods for the subject. Desires promote the fulfilment of practical functions in virtue of the fact that (i) a system of desire is reliably conducive to the production of good desires and (ii) it is effective in causing the formation of an intention (in the absence of salient countervailing reasons to doubt refrain from intention-formation).

I will make the case that this idea can be developed into an account of desire to elaborate the analogy between desire and perception in a way that avoids the problems of over-intellectualism that plagued views discussed earlier. Moreover, this account constitutes an advancement of the idea that desires represent their objects ‘as good’ in ways that are more informative than the views discussed in §4.3.3. The resulting account will underwrite my final account of the rational significance of desire. On my view, the explanation of how basic desire rationally supports intention formation is fundamentally *etiological*: desire owes its rational significance to the fact that they are caused and regulated by a mechanism that have certain practical functions.⁹⁰

The Plan

This is the idea in a nutshell. In order to develop it in more detail, there are core concepts that need to be explicated. The first part of the following chapter will address the following questions:

- What does it mean to say that something has a function? (§5.1)

⁹⁰This view is related to an influential proposal by Stampe (1987) which I will discuss in the following chapter.

- What is the nature of our desire-producing mechanisms? What do we know about their operation? (§5.2, 5.3.2)

Settling these questions allows an important intermediate claim to be established, namely that our desire producing mechanisms have a certain proper function: to produce good-conducive desires. With this claim established, I provide an account of the way in which desires are functionally isomorphic with perceptions. To do this, I motivate the view that this function is one shared by our intention-forming systems in a way that mirrors the coordination of function we see with respect to perception and belief.

Chapter 5

The Function of Desire

The purpose of this chapter is to introduce the approach that will be taken to address the Problem of Conative Significance. The central idea is that an explanation of the rational significance of desire should begin with the biological function that basic desires have: to motivate the subject to attain biological goods. More accurately, it should begin with the biological function of our *desire-producing mechanisms* (DPMs) which have evolved to produce desires for the kinds of goods needed for survival and reproduction.

Clearly, we are capable of forming desires for an extremely diverse range of things that are not obviously connected with survival and reproduction. So simply appealing to the fact that our DPMs motivate us to pursue ends conducive to survival and reproduction cannot provide a *general* account of the rationality of desire. Nevertheless, the central thought here is that the starting point for an account for the rational significance of desire should begin with an understanding of how desires, despite their diversity, stem from a capacity that has evolved to be sensitive to the kinds of things that is good for the kind of living thing that she is.

As I will explain, DPMs do not just produce desires that motivate us to pursue ends that have been historically good for the kind of individual that she is. They also produce novel desires resulting from interactions between her conative system and the employment of representational capacities, both of her environment and of hypothetical outcomes. By explaining how DPMs operate, it will be clear how despite the diverse

objects of our desires, desires nevertheless have features that connect them to what is historically good for the kinds of creatures we are. By understanding how DPMs work as a *productive* capacity against this core biological function, we have the beginnings of an answer to the Problem of Conative Significance.

The plan for this chapter is as follows. In §5.1, I introduce a framework for understanding function ascription which will be employed to state the central claim of this chapter, an etiological account of functions drawing on the work of Ruth Millikan. In §5.2, I begin by refocussing the discussion on DPMs and introduce a hypothesis about their core biological function which dovetails with the etiological account. I clarify the notion of core biological functions with respect to DPMs and desires (§5.2.2, 5.2.1) and introduce how the account developed will provide the foundation for an explanation of the rational significance of desire (§5.2.3). Throughout the course of this section, I motivate the need for such accounts to explain how this style of explanation can be extended to account for desire in all its diversity. To resolve this question, in §5.3, I turn to the question of how DPMs operate. By drawing on an empirically-informed account of desire, I isolate several features about how DPMs produce novel desires in ways that retain a connection with the core function of DPMs. This will address the worry developed in §5.2 in a way that preserves the explanatory strategy to be pursued in the following chapter.

5.1 Etiology and Function

One of the question to be pursued in this chapter concerns how our DPMs and the system in which these mechanisms are embedded works.

It is important to differentiate two senses in which a mechanism or system can be said to ‘work’. As Millikan (1986) emphasises, one sense of ‘work’ does not presuppose that there is a way for the mechanism or system in question to work *correctly* or to *fail* to work. For instance, in studying how a molecule works, one might study the forces that hold the system together as well as how the system interacts with other molecules. It does not make sense to claim that in observing physical change that the molecule might undergo, that the molecule is failing to work. The

other sense of ‘work’ does presuppose the applicability of such notions as failing to work well, or working as it is supposed to. Items that fall within biological categories are like this. Take for example the heart for which certain of its changes or effects constitute its failing to work well or working as it should.

5.1.1 Function Ascription

We can make sense of there being certain ways that a device is supposed to work by reference to certain *functions* that the device has. Function ascription is regularly employed across the biological sciences as well as in discourse about artefacts. Examples include:

4. The function of a heart is to pump blood.
5. The function of the fan is to cool its surroundings.

There are two key features of function ascriptions. The first feature is that the fact that something has a function seems to impose some normative standard: the fact that the heart has the function it does constitutes the standard against which the heart’s producing a certain effect counts as a malfunction or failure. This feature of function ascription allows us to frame questions about how certain devices ‘work’, in the sense that presupposes the possibility that certain things that the device could do count as its malfunctioning. The second feature is that certain ascriptions of functions to items have the power to explain the presence of those items. For instance, by ascribing to the heart in my chest the function of pumping blood, we can explain why it exists there in my chest. Function ascriptions thus constitute a form of teleological explanation of the existence of certain entities.

These two features of function ascription require explanation. Take the idea that function ascription allows us to make sense of an item as operating dysfunctionally. Some explanations intuitively apply to artefacts where it is the *designer’s* intentions about how the device should work that constitute the standards relative to which dysfunction is articulated. But such an explanation could only be used to explain the standards that apply to biological items like hearts if we abandoned a broadly naturalistic worldview. The second feature needs explaining too.

Function ascriptions tell us about the kind of effect that something is supposed to have. But how can the kind of effect that something is supposed to have, one that the token item in question need not have actually ever produced, enter into an explanation of the existence of that token?

5.1.2 Proper Functions

A dominant ‘etiological’ approach to function ascription appeals to something in a device’s *past* that grounds the function ascription without recourse to the intentions of designers. Roughly, etiological approaches analyse what it is for a token x to have a function to ϕ in terms of the fact that prior items of the kind X have previously ϕ -ed, and it is in virtue of X s having ϕ -ed that explains why the token x now exists. For example, an etiological account of the function of Tom’s heart analyses Tom’s heart as having the function to pump blood as opposed to cause a rhythmic sound because pumping blood is the type of activity that ancestral hearts have engaged in that explains why Tom’s heart exists.

When it comes to biological phenomena, the best known explication of the part of the explanation connecting ancestral X s to the presence of current X s appeals to the process of natural selection (cf. Wright, 1973; Millikan, 1984; Godfrey-Smith, 1994). It is the heart’s capacity *to pump blood* that explains the presence of hearts in members of a species since those with organs with this effect possessed superior relative fitness, explaining why over time the traits responsible for this effect have proliferated throughout that population.

We can sharpen this thought by drawing on the work of Ruth Millikan, in particular some of the theoretical concepts she uses to articulate her sophisticated theory of mental content. For biological items like hearts, what makes something a heart is that it was produced via mechanisms in accordance with an explanation of the production of ancestral hearts that, having done certain things (e.g. pump blood), have aided the prospects of that organism’s survival and reproduction (Millikan, 1986, p. 51). This constitutes what Millikan calls a ‘Normal’ (capitalised) explanation of heart production, where ‘Normal’ refers to a non-statistical etiological-cum-normative notion of normality (Millikan, 1984, pp. 33–34).

The production of these effects are what Millikan calls the *proper function* of hearts (Millikan, 1986, pp. 28–9). More precisely: a proper function of a token X is a function, the performance of which by ancestral X s explains the proliferation of the genes responsible for it.⁹¹ Relative to a proper function that an item has, we can identify a Normal explanation for its producing the effect ϕ responsible for its proliferation. Normal explanations include some manner of operation that constitutes how that item Normally functions (i.e. how it operated to produce the effect it was selected for), as well as Normal conditions for the item's ϕ -ing (i.e. conditions that obtain for Normal functioning). Relative to an item with a proper function, we can identify a *norm*, a standard the fulfilment of which articulates its functioning Normally. It is by reference to this norm that we can make sense of the items working as it should, or malfunctioning.

The notion of a proper function allows us to make sense of function ascription in cases when certain items are unlikely to, or cannot due to defect, produce effects in accordance with their proper function (Millikan, 1986, p. 53). An individual sperm cell has the proper function to fertilise an ovum even though the average sperm cell is unlikely ever to succeed in fertilising an ovum. And a heart so defective as to never have pumped blood counts as a heart and has the proper function to pump blood in virtue of having been produced by mechanisms in accordance with a Normal explanation of hearts.

Before we turn to desires and the mechanisms that produce them, I want to introduce an important distinction between *direct* and *derived* proper functions. We have just claimed that one way for a token X to acquire a proper function is for it to be a replication of an ancestral X in line with a Normal explanation. When this is the case, the token has a *direct* proper function. But there is another way to acquire a proper function: by being the product of an item with a proper function. We can illustrate this with Millikan's example of a chameleon's pigment-rearranging mechanism. This mechanism has the direct proper function to change its colour to match its immediate environment thereby camouflaging it

⁹¹For a more precise and thorough definition than is strictly needed for present purposes, see Millikan (1984, Ch. 1).

from predators.⁹² Suppose a chameleon finds itself on a highly specific paisley pattern * that it has never been produced before and its pigment mechanisms change so as to match *. Here this particular *-patterning has a *derived* rather than direct proper function to camouflage it from predators since the pigment mechanisms have never produced this * pattern before.⁹³ Derived proper functions then are those acquired in virtue of being the product of a mechanism with a certain proper function (Millikan, 1984, pp. 41–2). In contrast with the ‘direct’ proper functions of their producers, it is possible for such devices to have derived proper functions that have never been possessed by ancestors.

For present purposes, these are the only notions I want to use from Millikan’s work; since my aim is not to produce a teleosemantic account of desire content, I will not discuss Millikan’s complex consumer-based account of content-determination.⁹⁴ I want to emphasise here that although I will be making use of a teleofunctionalist account of mental state individuation, I will not be defending, or indeed assuming, a teleofunctionalist account of mental representation, that is, a *teleosemantics* for the content of mental states. It will suffice for present purposes, simply to refer to mental representations leaving aside substantive issues of how their content is fixed.

⁹²Millikan calls these *relational* proper functions which is a function something has in virtue of its standing in certain relations (Millikan, 1984, p. 39).

⁹³The resulting skin’s having that * pattern is what Millikan calls an ‘adapted’ device (Millikan, 1984, p. 40).

⁹⁴Very roughly, alongside the notion of a producer, Millikan has the notion of the consumer of a representation. Consumers are entities that ‘use’ a representation *R* in order to perform their relational function (with respect to *R*). For Millikan, it is by reference to *consumers* and their functions that one determines the content of representational states. For a consumer of a belief representation to perform its function in response to *R*, *R* must be embedded within a system of representations that systematically map onto states of affairs, one captured by a mapping rule. For Millikan, the content of a belief is what must obtain in the least detailed Normal explanation of how the *consumer* performed its proper function where this will determine the mapping rule that maps the belief onto the content-determining state of affairs. Similarly, there is a mapping between desires and the state of affairs that are the effects of those desires. The consumers of desiderative representations are those involved in what we can think of as the process of ‘practical inference’ that underwrites action. The content of a desire is determined by the least detailed Normal explanation of how the consumer of the representation performed its proper function of causing behaviours, one that with the relevant belief, leads to some content-determining state of affairs determined by the relevant mapping.

5.2 The Guise of the Biological Good

Let us now return to the topic of basic desires; specifically the mechanisms that produce them. What I wish to do here is to propose a hypothesis about their proper function. This will be important in order to frame the following discussion about how DPMs work. Indeed a plausible hypothesis about the proper function of the mechanism and of the system in which it is embedded, is needed in order to even know *when it is working* in the sense that presupposes that there are ways for it to working well, or malfunctioning (Millikan, 1986, p. 56). Against this hypothesis, we can begin to articulate explanations of their Normal constitution, why they are constituted as they are, and demarcate Normal or abNormal ways in which it operates.

5.2.1 Biological Good as the Function of DPMs

What then is a plausible (set of) function(s) ‘proper’ to DPMs?

It is highly plausible that DPMs having in the past produced desires that motivate subjects to do what is conducive to survival and reproduction explains the existence of our DPMs currently operating as they do. One trivial function of course is to produce basic desires. As I will explain, given the plausibility of some (impure) motivational account of desire, their minimal function is to produce states that dispose the subject in tandem with beliefs to engage in certain behaviours. Consonant with the approach taken to study the proper function of desires, we should emphasise the fact that DPMs are the products of natural selection.

We can sharpen this thought in two ways by drawing a distinction between different targets for function ascription. So far we have considered function ascription applying to components of systems, such as the heart which has a function within the circulatory system. Whole systems are functionally individuated; for instance, we can speak of the function of the circulatory system. Finally, we can speak of functions that apply to the organism as a whole.⁹⁵ Within the category of functions that apply to individual organisms, we can identify what we can call *core biologi-*

⁹⁵For a discussion of whole animal functions in the context of articulating primitive forms of agency, see Burge (2009).

cal functions. Examples include eating, drinking, locomoting, foraging, fleeing, mating, rearing offspring and so on. These core biological functions contribute to an explain of why a certain class of organisms have successfully reproduced.

A plausible hypothesis then is that mechanisms that produce basic desires have the direct proper function to produce desires for outcomes that, should they obtain, conduce to the fulfilment of core biological functions of the animal of which they are a part. I will refer to these outcomes as *biological goods*. The fulfilment of core biological functions constitutes a central norm for such mechanisms. Given that fulfilling a biological norm is good for an organism, the desires that are produced by our DPMs when functioning normally are good-conducive. It is in virtue of ancestral DPMs having produced good-conducive standing desires that explains why our DPMs contributed to our survival and reproduction that led to the proliferation of traits responsible for our DPMs. Producing good-conducive standing desires is part of Normal functioning for DPMs.

5.2.2 The Proper Function of Desire

What then of the proper functions of desire?

Let us take as our starting point Millikan's claim that "the most obvious proper function of every desire. . . is to help cause its own fulfilment" (Millikan, 1986, p. 63).⁹⁶ How does desire Normally contribute to the cause of its own fulfilment? Clearly not just any state of an organism that conduces to behaviour is a desire. Many simple homeostatic drives of simple organisms simply lack features contained in our ordinary notion of desire. These motivational states are often brought about by homeostatic imbalance that activate fixed patterns of behaviour.⁹⁷

⁹⁶It is important to point out that within the discussion in which Millikan makes this claim, she is explicitly employing the sense of 'desire' in accordance with the broader notion of a 'pro-attitude' discussed in Chapter 2 of which the kinds of desires we are interested in are a proper subset (Millikan calls these 'yearning desires').

⁹⁷Take the feeding behaviour of the blowfly which employs a simple set-point system. A blowfly has chemoreceptors on its legs which are stimulated when it steps on liquid containing sugar. If its chemoreceptors are stimulated, then if its stomach is empty, the fly will extend its proboscis and consume the fluid. Neural signals transmitted from the stomach via a recurrent nerve inhibit the consumption behaviour as its stomach begins to get full. Such motivational states lack the hallmarks of desire. Though it

As claimed in Chapter 2, it is plausible that desires are states that operate with a system of beliefs, in particular beliefs about means-ends relations, to produce action. One plausible central function then is to serve as inputs alongside beliefs into systems that lead to the production of actions or the formation of intentions. As Millikan puts it, desire's function is to "participate in processes that ultimately effect their own fulfilment" (Millikan, 1986, p. 63). Much more will be said about this process in Chapter 6, in particular how this process works in such a way that can address the Problem of Conative Significance.

It is equally important to emphasise the derived proper functions of desires owing to their being the products of DPMs. DPMs have core biological functions which they perform by producing desires that conduce to the fulfilment of these functions. Desires then, it is equally plausible to suppose, have the proper function to conduce to the attainment of biological goods. Desires do this by participating in processes that lead to actions, e.g. initiating processes that lead to the formation of intentions. (For simplicity, this glosses over details that will be expanded on later; specifically the distinction between standing and occurrent desires.)

Both of these claims about the function of self-fulfilment and the function to secure biological goods are plausible. How then, we might wonder, does the function of self-fulfilment relate to their biological function? We can articulate this question in another way that I think draws out something crucial that needs explaining.

There is an evident diversity in the kinds of desires that a mature human can form compared to more representationally limited organisms. Each of these desires aims at their own satisfaction even if extrinsic conditions mean that they never end up getting satisfied, as when they are outweighed by other desires, or set aside owing to the low probability of their satisfaction. Yet, the mechanisms that produce desires have proliferated because they produced desires that conduce to the fulfilment of core biological functions. This gives rise to the possibility that desires

is true that various states of the blowfly might be said to carry information about its internal or external states, this falls short of constituting beliefs. The blowfly does not 'believe' that it is currently stepping in sugary water, or has an empty stomach. *A fortiori*, its behaviour is not produced by a 'desire' to extend its proboscis in the belief that this will allow it to consume the sugary fluid. For discussion, see Dretske (1988, p. 124).

can have conflicting proper functions. A desire can have the proper function to conduce to an organism's good, yet simultaneously conflict with the function to work with belief to cause their fulfilment. The fulfilment of one proper function means the frustration of another when what is wanted is bad.⁹⁸ Why aren't such conflicts utterly pervasive? What connects the desires that are produced that have no obvious bearing on the attainment of core biological goods and the ones that do? What is needed then is an explanation of some explanation of the Normal operation of our DPMs, their Normal constitution (i.e. how it was constituted to produce the effect it was selected for). In order to better characterise the workings of our DPMs, I want consider an empirically informed account of desires, which gives the evidential basis for the characterisation of the operations of the psychological systems that underlie motivation. This will be conducted in §5.3.1.

Before this, however, I want to consider an account of desire proposed by Dennis Stampe in his influential "The Authority of Desire" (Stampe, 1987), which also draws a connection between desire and biological goods. There are two reasons for this other than to do due diligence to the literature. First, it is possible to raise a related explanatory worry concerning his account as has just been discussed, which will help clarify the issue I have suggested in this section needs explaining. Second, I wish to start explaining how all of this— the account of desires, DPMs and their core biological function— provides the foundation for a resolution to the Problem of Conative Significance. My explanation will employ the same foundations Stampe used to explain what he calls the 'authority of desire'.

5.2.3 Stampe and the Authority of Desire

Stampe defends a version of the motivational functionalist account (see Ch. 2) on which desires, in tandem with beliefs, are behavioural dispositions to do what will satisfy one's desire in a world in which one's beliefs, whatever they are, are true. The central difference is that according to Stampe, we should think of desires as akin to perceptions, in particular, with respect to their having certain 'ideal' causes:

⁹⁸Millikan makes a related point in her discussion of the example of a maladapted honey worker bee display, for more discussion Millikan (1984, p. 43).

Desire is a kind of perception. One who wants it to be the case that p perceives something that makes it seem to that person as if it would be good were it to be the case that p , and seem so in a way that is characteristic of perception. To desire something is to be in a kind of perceptual state, in which that thing seems good. . .

[Desire's] epistemic role will be construed on the model of that of perception. I do not say that desire is a mode of sensuous perception, involving a distinctive sensation. . . For the essential connection between perceiving and sensing is not the matter of having sensations, but that of being *sensitive*. . .

Sensitivity, I suggest is a matter of being affected by and only by the presence of specific properties. And the sensitivity involved in desire is a matter of such states being produced by a mental mechanism that is activated, ideally, by and only by the apparent goodness of a state of affairs. (Stampe, 1987, pp. 359–60)

Stampe's proposal, specifically the part concerning mental mechanisms i.e. DPMS, is clearly related to the account articulated in this chapter of DPMS having the core biological functions. When he first introduces the perceptual analogy (Stampe, 1987, p. 361), Stampe claims that a desire to ϕ is a state 'ideally' caused by the apparent goodness of the state of affairs in which one ϕ -s, in the same way that a perception that p is a state caused by a mechanism 'ideally' activated by the state of affairs in which p . Stampe however does not define what makes a cause ideal rather than non-ideal. An obvious candidate would be to give an account of ideality by reference to functions, so that a desire that p is a state caused by a mechanism that counts as *functioning Normally* when it is activated by the apparent goodness of a state of affairs.⁹⁹

Stampe appeals to this to explain what he calls the authority of desire:

The fact that I want something, in and of itself, is ordinarily a reason for me to act accordingly. The desire itself therefore makes the act a rational one, in the sense of one I have a reason to perform—not, perhaps a decisive one, perhaps not even a good one, but still *a* reason. (Stampe, 1987, pp. 342–3)

Though I think Stampe's strategy is the key to solving the Problem

⁹⁹Stampe was one of the first philosophers to suggest that the notion of functions can fruitfully supplement the development of a causal theories of content. See Stampe (1977).

of Conative Significance, I am not sure that all of Stampe's claims about the authority of desire are right.¹⁰⁰ For now, I want to isolate his claim that a desire can "make an act a rational one", what I have called the intuition of Rational Support.

Stampe's account lays a foundational element in a promising answer to the Problem of Conative Significance. In particular, it will allow us to provide an explanation of desire's rational significance that can satisfy the rational support constraint. We can meet this constraint if, as Stampe proposes, desires are *reliable indicators* of the apparent good toward which they motivate us (Stampe, 1987, p. 364). For this to hold, we first need to make the case that there is, in the first instance, a connection between desire and what would be good for the subject to do. This is something that both Stampe's account and the proper function account of DPMs do by connecting desires, through their causes, with the apparent good. In the latter case, desires are produced by mechanisms which Normally produce desires that conduce to the fulfilment of core biological goods.

I want to highlight a lacuna in the proper function account of DPMs given so far, one that bears on its deployment within the kind of solution I wish to give to the Problem of Conative Significance. It is important in order to satisfy the rational support constraint that DPMs *reliably* produce good-conducive desires. This reliability condition will fail to hold if, as per the worry raised in §5.2.2, DPMs were inefficient, producing many desires that motivate the subject to do what often has no bearing on what is good for the subject (or worse, what is *bad* for the subject). This is an issue of which I will delay discussion until the end of §5.3. I will argue that features of how DPMs work mitigates this worry. However before this, I want to take more time to consider Stampe's own account,

¹⁰⁰We can isolate three distinct claims. First, there is what we can call the *rationality* claim, namely that a desire to ϕ can make that action a rational one. This is interwoven with a second claim that we can call the *reason-constitution* claim according to which a desire to ϕ is a reason to ϕ . Finally, there is what we can call the *possession* claim on which the desire to ϕ itself "would be sufficient... for one's being in cognitive possession of that reason— that is, one's "having" a reason to do those acts" (Stampe, 1986, p. 158) These claims, and their interrelations, need to be clearly articulated. For instance, is the rationality claim to be explained by the possession claim? If so how does the reason that one is thereby in possession of relate to the reason-constitution claim? Is the reason one possess the fact that one wants to ϕ ? These issues will taken up in Chapter 6.

in particular, his discussion of what he takes to be the ‘ideal cause’ of desire.

The Ideal Cause of Desire

Stampe introduces the notion of a ‘ Σ state’, a term used to refer to the ideal causes of DPM activation. Stampe defines a Σ state as (i) “some state of the desirer himself” that (ii) “causes it to seem that some state of affairs would be good”, which is (iii) distinct from the desire itself and (iv) “[to be] distinguished from the ‘objective’ state of affairs in which that [desired] thing would be good for the subject” (Stampe, 1987, p. 372). This is a somewhat abstract proposal so it will help to consider two instances that Stampe takes to be Σ states.

The first kind of candidate for states that qualify as Σ are physical states of an organism, those which partially ground the *needs* of an organism:

It may be, however, that Σ is a plain bodily state— for instance, a state in which the body is depleted of air or water where one’s being in that state makes it seem to one as if it would be good to breathe or drink. (And in this case, the subject’s state may be identified with the objective state of affairs which is such that it would be good for him to breathe or drink.) That state, then, will be the object of nonepistemic perception.

... Σ may be a state such that the subject *needs* something, as one depleted of water may need water. In this way it may be that a desire is a perception of a need. The causal set-up here is roughly as follows: the organism’s being depleted of water causes it to be the case that it needs water, *and*, its being depleted of water also causes certain bodily sensations which cause the desire for water. The desire for water is the perception of the state of depletion, thus of a state *in which* one needs water, thus of a state such that it would be good if one had water. (Stampe, 1987, pp. 373–4)

The second candidate for states that count as Σ are the subject’s background of beliefs and desires:

I suggest that Σ may also be comprise the cognitive and larger conative state of one who has a certain belief, aims and other desires, where one’s being that state makes it seem to be as if something would be good— as it *may* do. Here too, I suggest, in wanting

that thing one is subject to a certain kind of nonepistemic awareness of one's own state, even one's own total mental state. (Stampe, 1987, p. 373)

Stampe provides only a sketch of the first candidate and there is almost no guidance on how to understand the second candidate other than the two sentence proposal quoted above.

A worry for Stampe's proposal taken at face value concerns whether his account is able to provide the etiology-based explanation that will cover every desire that we take to be rationally significant. In a nutshell, the problem concerns how we are to find Σ states of the right kind to provide a plausible etiology for a range of desires. Whilst it is clear in the case of appetitive desires, say to drink, what the relevant Σ state would be, e.g. states of water deprivation, it is not so clear for most other desires.

Consider the basic desire to eat ice cream. We face problems when we try to explain the generation of such desires by reference to the two central examples of Σ states that Stampe provides. It is unclear how a basic desire to eat ice cream can be understood on the model of the first option that Stampe considers, "plain bodily states" (e.g. being dehydrated) or the state of a subject's needing to consume water. It is implausible to think that there is some physiological state, one corresponding to a need for ice cream *per se*. *A fortiori*, it is unclear what can serve as the initial cause of the causal set up that Stampe sketches in the quotation given earlier. So that leaves the second option, the larger cognitive or conative state of the subject. There are unresolved questions about how this is supposed to work. Presumably, given that the desire in question is basic, we should not think of the larger conative-cognitive state of the subject as figuring in the cause of the desire on the model of how *instrumental* desires are generated. In the case I am envisaging, one is simply "assailed" by a desire for ice cream. How then, if we pursue this idea, do we explain the generation of basic desires?

To summarise the broader worry illustrated with this example, we can ask: how is the kind of etiological explanation we wish to give of desire's rational significance going to account for desire's rational significance in all of its diversity? We can think of this worry as the same one raised in §5.2.2 but from the opposite direction. There we wondered

how the diversity of desire relates to the biological function of DPMs; here the issue is about how to explain the diversity of desire given the underspecification of the way in which DPMs operate.

5.3 Normal Functioning and Normal Constitution

In this section, I aim to address these outstanding issues by enriching our account of the nature of DPMs and desire by drawing on an empirically informed account of desire proposed by Tim Schroeder. I begin by outlining Schroeder's account (§5.3.1). As we will see, Schroeder's account of desire is one with enough structure to allow us to articulate how our DPMs produce a diverse range of basic desires in ways that connect their production with our DPM's proper function to produce good-conducive desires in a way that allays the worries discussed in the previous section (§5.3.2). This will amount to an elaboration of the Normal constitution of DPMs and how they produce desires in line with a Normal explanation of the performance of their proper functions.

5.3.1 Schroeder's Reward Theory of Desire

In this section, I consider an account of desire proposed by Timothy Schroeder (Schroeder, 2004; Arpaly and Schroeder, 2013). Schroeder's account is unique in centralising not motivation, affect or pleasure, but *reward* or *punishment* as essential to desire. According to Schroeder, desire is connected to motivation and pleasure, not by consisting in behavioural or hedonic dispositions, but by being the contingent *cause* of these effects.¹⁰¹ Schroeder's account distinguishes between desires of which reward is the essence, and aversions of which punishment is the essence. Schroeder succinctly formulates his account of desire as follows:

To have an intrinsic (positive) desire that P is to use the capacity to perceptually or cognitively represent that P to constitute P as a reward. To be averse to it being the case that P is to use the capacity to perceptually or cognitively represent that P to constitute P as a punishment. (Schroeder, 2004, p. 131)

¹⁰¹For discussion, see Schroeder (2004, Ch.3-4).

Before we unpack Schroeder's account, it will be necessary to outline some features of our ordinary concept of reward and punishment. Reward or punishment are diverse: whether something is a reward for someone will depend on features of that person's psychological makeup; for example, food pellets for a rat constitute a reward, but a punishment for a gourmand. Rewards and punishment have effects on stable behavioural dispositions. Schroeder focuses on two such dispositions: (i) changes in behaviour that proceed via memory, planning and deliberation (Schroeder, 2004, p. 41), e.g. a child might learn that visiting her grandmother typically yields pocket money that can be spent on sweets and so does what she can to visit; and (ii) rewards can also lead to changes in habits, e.g. a child might 'learn' to sit still at the dinner table through repeated reward or punishment. Rewards and punishments can also have psychological effects, in particular, the generation and elimination of certain intrinsic desires. For instance, a child rewarded by its parents for kind behaviour and empathetic concern toward his sister, and punished for selfish and mean behaviour, can lead to the formation of intrinsic desires for his sister's well-being (Schroeder, 2004, p. 42).

Schroeder aims to provide us with an account of what it is for an event to constitute a reward or punishment in a way that *explains* these effects.¹⁰² Here I will focus first on reward for which Schroeder provides a more detailed account before turning to punishment which receives comparatively less attention. Schroeder's view is that rewards should be understood within the context of a *learning* system:

Contingency-based Learning Theory of Reward: For an

¹⁰²It is important to separate the notion of a reward from the technical notion of a *reinforcer*, which, as Gottfried (2011, p. 45) notes, is often referred to by behavioural scientists as 'reward'. A textbook definition of a reinforcer is an event that increases the probability or rate of a certain response or behaviour when the event is presented after the behaviour on which it is contingent (Pearce, 2013, p. 93). Failure to keep these two notions clearly in view can lead to problems. Schroeder accuses the 'pure' behaviourist of making this mistake. As Schroeder (2004, p. 57) notes, we want a notion of reward that can *explain* the effects we observe in operant conditioning experiments. Rewards, on Schroeder's view, *causes* rather than simply *consists in* changes in behavioural disposition. We can thus explain why a rat has a strengthened disposition to press a lever by appeal to the timely administration of reward. But if 'reward' is simply defined as any event that increases the probability of a certain response, then we could not appeal to reward—its features and how its representation integrates with the psychological makeup of the subject—to *explain* why the subject develops strengthened dispositions to respond in a certain way (Schroeder, 2004, p. 57).

event to be a reward for an organism is for representations of that event to tend to contribute to the production of a reinforcement signal in the organism, in the sense made clear by computational theories of what is called ‘reinforcement learning’. (Schroeder, 2004, p. 66).

To unpack the contingency-based learning theory, it will be necessary to sketch a system employing a simple form of reinforcement learning, one that Schroeder takes to be implemented by the neural structures that embody the biological reward and punishment system. The reward system is one that can be employed whenever a certain mental state M_1 participates in causing another mental state M_2 . The core function of the system is to produce a ‘learning’ signal that modifies the likelihood that the occurrence of M_1 causes M_2 . In this case, we will focus on the role that reinforcement can play in increasing the likelihood that the representation of a state of affairs in perception or thought (e.g. receiving a food stuff) will cause a mental state that participates in bringing about a certain behavioural response ϕ (e.g. pulling a lever). Whenever M_1 causes M_2 , the reward system performs a calculation which then leads to the emission of either a *positive*, *negative* or *neutral* signal. A positive signal produces a change that *increases* the probability that the presence of M_1 causes M_2 . A negative signal *decreases* the probability that the presence of M_1 causes M_2 . A neutral signal brings about no change. For instance, after representing the delivery of a food item, the organism pulls a lever, the system can produce a positive, negative or neutral signal which increases, decreases or leaves unchanged respectively the likelihood that the subject will pull a lever on represented receipt of the food item. On Schroeder’s view, what it is for something to be a *reward* is for the representation of it to produce a *positive* signal.

Whether the reward system produces a positive, negative or neutral signal is the result of a sub-personal calculation that yields “the difference at each moment between the expected amount of net reward. . . in the world, and the actual amount of reward” (Arpaly and Schroeder, 2013, p. 133). In order to perform this calculation, the subject employs mechanisms that carry information about *actual* net reward at time t and *expected* net reward at t . When actual net reward is greater than expected net reward, a positive signal is produced. This signal represents

what is called a positive ‘reward prediction error’. When actual net reward is lower than expected, a negative signal is produced, representing negative prediction error. Otherwise when actual net reward is exactly as predicted a neutral signal is produced.¹⁰³

Schroeder proposes that a similar system underwrites aversions, although his account of aversion is less developed than for reward. For a subject to be averse to its being the case that p is for a representation that p to constitute a punishment for the subject. To constitute something as a punishment is for representations of it to produce a negatively valenced punishment signal as a result of a similar comparison, one that leads to changes decreasing the likelihood that representations of its being the case that p lead to some other mental state, for instance, one responsible for a certain kind of action. Correlatively, the punishment system generates a positively valenced signal (a ‘negative reward’) which indicates *absence* of punishment (Schroeder, 2004, p. 132; Arpaly and Schroeder, 2013, p. 131).

Schroeder suggests that there is evidence supporting a picture of the biological reward system as operating in roughly this way. The task of detecting levels of reward is discharged by areas including the sensory and association cortex, where sensory and cognitive representations occur, and the hypothalamus, where monitoring of certain metabolic processes occur. Schroeder speculates that structures such as the *nucleus accumbens* and the *caudate nucleus* are involved in keeping track of expected reward and reward comparison calculations (Schroeder, 2004, p. 53), as well as the *orbitofrontal cortex* (OFC) which is connected to input structures involved in categorising perceptual representations as involving rewards or punishments (Schroeder, 2004, p. 53). Schroeder suggests that two structures called the *ventral tegmental area* (VTA) and the *pars compacta* of the *substantia nigra* (SNpc) constitute the *output* structures of the reward system. Schroeder cites work by Wolfram Schultz and colleagues showing that the neurons of the VTA/SNpc release dopamine, relaying an output to other areas of the brain which have been found to contain dopamine receptors. Schultz et al. found that the

¹⁰³For an introduction to reward prediction error and the role of dopamine in signalling reward prediction error, see Schultz (2016). For more general coverage, Sutton, Barto, and Bach (2018) is an extensive and wide-ranging study of the kind of learning in question.

dopamine neurons of the VTA/SNpc displayed patterns of activity that can be interpreted as carrying information about reward prediction error.¹⁰⁴ Schroeder suggests that the behavioural effects of our ordinary notion of reward could be in principle be explained in the way he suggests given that the biological reward system sends output to neural structures such as the motor striatum which integrates a range of information needed for the production of habitual behaviour (Schroeder, 2004, p. 119) and goal-directed behaviour (Schroeder, 2004, p. 129).

Schroeder is more circumspect with respect to details concerning the neural basis of the punishment system. Schroeder speculates that the punishment centre of the brain is located in the *dorsal raphe nucleus* (DRN) (Schroeder, 2004, p. 54). The DRN receives information from the OFC, hypothalamus and the striatum just as with the VTA/SNpc. Schroeder also hypothesises that serotonin carries the relevant punishment signal, the “counterpart to the VTA/SNpc reward signal” (Schroeder, 2004, p. 54). However as Schroeder himself is keen to emphasise the state of research concerning the neural realisation of punishment is substantially less developed in comparison with that of the neural underpinning of reward. In particular, the role that serotonin plays in punishment based learning remains far from clear and there are no parallel findings to that found in the work of Schultz et al. (for discussion, see Schroeder, 2004, pp. 55–7).

5.3.2 Diversity, Reliability and the Good

My reason for outlining Schroeder’s theory is that it contains enough structure to allow us to articulate the way in which the constitution and operation of our conative system allows for the production of a diverse range of basic desires in ways that connect their production with our DPM’s proper function to produce good-conducive desires. To this end, I want to emphasis some general features about the mechanisms that underlie the formation and regulation of our desires and explain how these features are incorporated on Schroeder’s theory.

¹⁰⁴The role of the dopamine remains highly contested. An alternative view is one on which dopamine plays a role in the generation of episodes of “wanting” discussed earlier. For discussion, see Berridge (2007).

Relating Desire Formation and Maintenance with the Proper Function of Desire

The capacity to develop novel intrinsic desires for objects that we could not possibly have been innately endowed with is highly beneficial since it allows us to respond flexibly to a changing environment to maximise on the acquisition of goods needed for survival and reproduction. Obviously, this capacity would be all for nothing if our desire-producing mechanisms formed standing desires in a random fashion. Instead, our DPMs produce desire by means of a process that is subject to constraints to ensure that our DPMs produce desires that are reliably good-conducive.

The first of these constraints is that desire-formation is always constrained by a set of background desires. Roughly, this set of background desires are what I will call, following Railton (2010), our *conative priors*. Some of our conative priors are *innate*: these are desires set down over the course of our evolutionary history in order to ensure that we form desires for things that reliably tend to be good-conducive in the environments in which our DPMs have evolved e.g. food, water, warmth, social contact, and so on. Given the centrality of these goods for survival and reproduction, it is entirely unsurprising that we find any organism with the capacity to desire similarly configured. This idea is accommodated on Schroeder's Reward Theory because whether an outcome is desired is a matter of whether representations of that outcome lead to the production of reward signal, and whether a representation leads to the production of a reward signal depends on the subject's background desires. These innate conative priors comprise the Normal constitution of our DPMs.

Now, in order for DPMs to produce new desires for outcomes that are reliably good, it needs to be the case that there is some operation by which the subject is able to move from representations of a type of outcome that is previously undesired to the formation of a desire for that outcome. The Reward Theory fastens onto the role that repeated association of a represented outcome with a positive reward signal plays in the formation of desire. Intrinsic desires are formed when perceptual or cognitive representations which do not yet contribute to reward signalling are "tokened in a way that coincides with other representations that are contributing to reward signals" (Schroeder, 2004, p. 147). Here

the role of our conative priors enters as a constraint in what other desires we can form. Repeated coincidence over time ensures that “the connections between the representations and the reward system will become strengthened, until the connections are efficacious at producing reward signals on their own” (Schroeder, 2004, p. 147). Schroeder gives as an example the formation by an infant of an intrinsic desire for its mother’s presence by a process in which its innate desires for food, warmth and social contact are, being satisfied, lead to the production of reward signals at the same time as visual representations of its mother, which then lead to representations of the mother themselves being capable of causing reward signalling.¹⁰⁵ Though temporal contiguity is a fairly crude way to form intrinsic desires, it is an economical and reliable way to form intrinsic desires that have a connection with the satisfaction of one’s basic needs. This new intrinsic desire can then participate in the process by which yet more intrinsic desires are formed. It is plausible then that the formation of intrinsic desires via this associative process can, over an individual’s lifetime, explain the diversity of the objects of her desires, yet preserve a connection, however indirect, with their goodness for the subject.

These points also apply to the way in which we develop novel aversions. Just as there are innate conative priors, states of affairs that are naturally rewarding, there are innate aversions, state of affairs that naturally drive punishment signals. These obviously include the representations of bodily damage, the ingestion of various substances, and so on. Similarly, there are processes that lead to a represented state of affairs becoming associated with a positive punishment signal.

There is one final feature concerned with the regulation or maintenance of our stock of basic desires. Our existing desires are regulated in ways that maintain their capacity to motivate one toward the good. When our environment changes so that a type of state of affairs that was once constituted as a reward instead yields a punishment, our conative system exploits reciprocal connections between the reward system and the punishment system to weaken or eliminate certain basic desires.

¹⁰⁵A process like this very plausibly underlies ‘secondary reinforcement’ when stimuli that are not natural rewards for the subject (e.g. the ringing of a bell) reinforces actions when associated with other stimuli that is naturally rewarding (primary reinforcer).

Schroeder writes that this can be accounted for on the Reward Theory: we come to lose our intrinsic desires if “it regularly happens that when the representation [i.e. one that the subject constitutes as a reward] is tokened, other representations are also tokened that cause a punishment, rather than reward, signal to be released” (Schroeder, 2004, p. 147). Over time, Schroeder suggests that, “the connection between representation and reward becomes weakened, until eventually no efficacious connection is left at all” (Schroeder, 2004, p. 147).

In the discussion of Railton’s account, it was suggested our affective system responsible for producing pleasurable and displeasurable experiences could play a role in regulating desires. Railton’s suggestion was that liking the prospect that p elicits and regulates an expectation of how good or pleasurable it would be if p . This expectation is then compared with good or pleasurable the fulfilment of the desire was. When a fulfilled desire yields a better than expected outcome, this reinforces a subsequent propensity to have occurrent episodes of prospective liking, which can then elicits yet more motivation.¹⁰⁶ Railton’s account makes episodes of liking the triggering and sustaining cause of occurrent desires. Even if we don’t go along with this suggestion, it is nonetheless plausible that our affective experiences have some role in regulating our standing desires.¹⁰⁷

Although here I have focussed on explicating these features on Schroeder’s view, I do not want to endorse it in all its details, in particular Schroeder’s claims about the neurological underpinnings of both the reward and punishment systems. I discuss it only because it has enough structure and complexity to capture how the operation of our DPMS also tend to satisfy their function to produce good-conducive desires. There are other rival empirically-informed views of desire containing enough structure

¹⁰⁶Relevant here is Berridge’s suggestion that stimulation of the same structures that contain hedonic hotspots can give rise to or enhance “wanting” (Berridge, 2009, p. 385).

¹⁰⁷There is some important experimental work about the way in which affective states regulate our sensitivities to the reward value of certain outcomes. Balleine and Dickinson (1991) for instance finds that rats who had undergone aversion conditioning to change its affective reactions to a reinforcer needed to re-experience its altered affective reactions in order to drive changes to sensitivities to incentive value. Dickinson (2008) contains some interesting hypotheses about the function of affective experience in interfacing the cruder system controlling affect and the system we have for goal-directed behavioural control.

on which the same features may be observed. One such view, influenced by Schroeder's approach, has been developed by Patrick Butlin (and in joint work with David Papineau) which demarcates desires by the role they play in the capacity for specifically goal-directed control of behaviour (Butlin, 2016; Butlin, 2017; Butlin and Papineau, 2016). Goal-directed behaviour is conceptually distinguished by employing dual representations of the reward or incentive value of certain outcomes as well as representations of instrumental contingencies to drive behaviour (cf. Dickinson, 2011).¹⁰⁸ Butlin identifies desires as states that have the function of representing the incentive value of outcomes. Such a view diverges in important ways from the Reward Theory.¹⁰⁹ A full discussions of Butlin's view would take us too far away; for present purposes, I want to suggest that the features I have highlighted are effectively accommodated in modified form.¹¹⁰

Let us now return to the worry that was raised in §5.3.2. I suggested that it is a plausible hypothesis that the proper function of DPMs is to produce desires that conduce to the fulfilment of core biological functions. If this is right, then every basic desire produced by DPMs has the fulfilment of core biological goods as a proper function. But each desire, it is equally plausible to suppose, has as its function to bring about its *own* satisfaction by participating in processes that lead to action or the formation of intention. How then do we explain how the function of self-fulfilment relates to its core biological function? These features I have outlined using Schroeder's account as a framework show that our desire-producing mechanisms constitute a *productive* faculty, one operating within *constraints* and capable of adapting in response to environmental changes, thereby ensuring that the basic desires we form retain some connection with the subject's good, however indirect. Thus, we have an explanation of how diverse basic desires retain some *indirect*

¹⁰⁸Correlatively, there are two paradigms used to assess the presence of the capacity for goal-directed behavioural control. Outcome devaluation experiments test for sensitivity to manipulations of the reward value of certain outcomes (e.g. Adams, 1982) whilst instrumental contingency degradation tests for sensitivity to manipulations in instrumental contingency (e.g. Dickinson, 1998).

¹⁰⁹These differences are discussed in Butlin (2017).

¹¹⁰For example, see discussion of desire updating on Butlin's view (Butlin, 2017, §3) and standing basic drives as psychological primitives which captures the idea of conative priors (Butlin, 2017, §4, 6).

connection to core biological functions, mitigating the kinds of conflict we worried would be possible between the function of self-fulfilment and their core biological functions. These same features provide a bridging explanation of the kind that will help Stampe explain how we yield desires for objects that bear a less obvious connection with those states of affairs that constitute the ‘ideal cause’ of desire.

Reliability

I wish to end this section by considering one final issue concerning the reliability of DPMs to generate good-conducive desires, which in §5.2.3, I suggested, following Stampe, provides a way to pursue an explanation of desire’s rational significance in a way that satisfies the rational support constraint. I want to spend some time clarifying this thought.

Some preliminaries on reliability. First, the reliability of the mechanisms that comprise the conative system of which basic desires are a part turns on the degree to which these mechanisms yield good-conducive desires and minimises or eliminates bad-conducive desires. Second, whether a desire-forming process is good-conducive is multidimensional and a matter of degree. It is multidimensional because one process can reliably bring about desires that conduce to the fulfilment of a proper subset of biological functions and no others. And it is a matter of degree because what matters is conduciveness to function fulfilment, and different desires can be more or less conducive to core biological function fulfilment. Third, reliability of a desire-forming process is always relative to certain domain or environment, the contingencies of which determine the degree to which it is reliable; no process of desire formation can be good-conducive in *all* possible worlds.

The strategy we are pursuing is one which aims to explain how desires rationally support actions and intention-formation by appeal to their being the products of mechanisms that reliably produce good-conducive desires. What is the kind of environment reliability relative to which matters for such an explanation?

One view is that it is reliability relative to one’s *current* environment that is key to understanding rational assessment.¹¹¹ On this view, our

¹¹¹Compare Stampe (1987, 366, esp. fn.22).

DPMs and desires contribute to the provision of rational support only if they yield a sufficiently high degree of good-conducive desires in the agent's current environment. One problem with this is that it would make rational support hostage to changes in the agent's environment in ways that do not track our intuitions about when desires can rationally support choice.¹¹² For example, suppose unbeknownst to everyone, a cosmic ray lethally irradiates all the nutritive items we have natural appetites for. In this scenario, our DPMs will be highly unreliable. Nevertheless, it is plausible that someone unaware would be rational to act on the basis of her occurrent desire to drink water, even doing so would be fatal.

An alternative view ties reliability to normal functioning, drawing on work by Peter Graham (Graham, 2012; Graham, 2014). Let us start by returning to the example of sperm which is an item that gets replicated because of an effect it rarely produces. In contrast, some items get replicated because they *reliably* produce an effect. Compare for instance, a heart that fulfils its function *reliably* under Normal conditions. Were our DPMs selected for as a result of having *reliably* produced good-conducive desires? This is a delicate question, but I think it is plausible that it is in virtue of having, in the past, *reliably* produced desires for biologically good outcomes that our DPMs were selected for. When we consider the myriad, changeable and reoccurring needs of human beings in the kinds of dynamic, perilous environments that characterised the conditions in which our DPMs were selected for, it is plausible that mechanisms that only motivated us once in a blue moon to seek out biological goods on which survival depends will not do. Some degree of reliability is necessary.

If this is right, then reliably producing good-conducive desires comprises the proper function of our DPMs. On this view, a critical factor in an explanation of how desires can rationally support action will turn on whether they are produced by DPMs that are *functioning Normally*. When DPMs function Normally under Normal conditions, they reliably produce desires for outcomes that retain a more or less direct connection with what is good for the organism. This avoids the problem that faced

¹¹²This is analogous to the concerns some have pressed against certain process reliabilist accounts of epistemic warrant, e.g. the 'new evil demon' objection (Lehrer and Cohen, 1983).

the current-environment reliability proposal since it allows us to preserve the thought that the subject in the irradiated environment would remain rational were she to respond in a certain way in response to an occurrent desire, say to drink some water.

5.4 ‘Representing as Good’

Now is the time to return to the proposal sketched towards the end of Chapter 4 concerning how to interpret the thought that desires represent their objects as good. A brief refresher is in order. In §4.3 (“Perception as a Model for Desire”), we considered attempts in the literature to develop a parallelism between desire and perception to explain the rational contribution of desire to rational agency. These accounts latch on to the idea that a perception involves presenting or representing its objects *as true*, where what it is to present or represent p as true is *not* to be understood as representing the content *that P is true*, but simply to stand in the relation of *representing-as-true* to the content that p . Similarly on Schafer’s view, the idea was that a desire to ϕ “presents [a course of action ϕ] with imperatival force [i.e.] presents [ϕ -ing] to me as something I ought to do” (Schafer, 2013, p. 276), where for one to be presented with ϕ -ing as something I ought to do is part of “the manner in which [desires] present this content to the subject” (Schafer, 2013, p. 274)

The notion of attitudinal force is then employed to explain the rational significance of desire on the model of perception. What explains why perceptions can rationalise a belief that p is that there is an element of ‘fit’ between both states: the former presents its contents as true (with ‘assertoric’ force) and the latter is, as Schafer puts it, “an endorsement of the way the perceptual experience as of p makes things seem to the subject” (Schafer, 2013, p. 275). Similarly, a desire to ϕ can rationalise an intention to ϕ simply because desires present a course of action as something I ought to do and “the intention to [ϕ] is simply an endorsement of the way that such a mental state already presents [ϕ]-ing to the subject” (Schafer, 2013, p. 275).

I argued that this proposal is unacceptably obscure. To remedy this, I sketched a view on which we should analyse what it is for desire to represent its objects as good by the kind of functional role they have

in leading to action and the formation of intentions. We are now in a position to develop this in greater detail.

Desire as Good-Conducive Normal Initiators of Action

DPMs are psychological structures that have the direct proper function to produce desires that satisfy core biological functions. DPMs perform their function in the Normal way by operating alongside one's cognitive or perceptual capacity to represent the environment in order to produce novel desires as a function of the agent's conative priors and a reliably good-conducive update rule. So far, in order to simplify matters, we have suppressed mention of the distinction between standing and occurrent desires. I will begin to reintegrate this distinction here, developing it further in the next chapter.

Standing desires are states that have the direct proper function to cause the formation of intentions or action and they have the proper function to initiate the formation of intentions and actions that satisfy core biological functions. Standing desires perform their function in the Normal way by interfacing with the subject's cognitive and perceptual capacities to form intentions for outcomes that are believed to lead to the satisfaction of the respective standing desire, or to cause the formation of occurrent desires. Occurrent desires are desires that are active over a period of time in contributing to action selection by interfacing with the subject's cognitive and perceptual capacities. (More on occurrent desires to follow.)

The sense in which a standing desire represents its objects as good can be explicated in two parts. First, both standing desires and the intention-forming mechanism share common biological functions the performance of which conduce to the attainment of biological goods. Thus, both components of the system have a shared or coordinated function. Second, standing desires are one of the Normal *initiators* of action or intention-formation (absent countervailing considerations to refrain from acting or forming the relevant intention). In what we can call the 'default' case, it is part of *Normal functioning* for standing desires to initiate action or intention formation. Thus, the cash value of talk about desires presenting its objects *as good* is that standing desires are produced and regulated by

mechanisms to ensure that they Normally initiate good-conducive action and intention-formation. It is in this respect that desires are analogous to perceptions; like perceptions and beliefs, desires and the mechanisms for intention-formation have coordinated functions (truth-aiming and good-aiming respectively) where both perceptions and desires serve as elements in the Normal initiation of belief-formation and intention-formation respectively.

Conclusion

So far the view I have articulated only specifies that desires serve as reliably good-conducive Normal initiators of intention-formation and action. An important feature of the account is that it provides an analysis of what it is for desires to count as representations of courses of action as good *without* construing desires as personal-level judgements that a certain course of action is good. Though this feature allows us to avoid the problems that faced the desire-as-belief (DAB) model (§4.2), it means that we cannot appeal to the same answer to the Problem of Conative Significance relied on by DAB-style views to explain how desires can be rationally *intelligible* bases for action (§2.2). Without further elaboration of *how* desires exert their influence on the will, a satisfactory resolution to the problem remains out of reach. In the following chapter, an account of the relationship between desires and our capacity to act for reasons will be developed.

Outstanding Issues

I have relied on the idea that desires ‘initiate’ action and intention formation; this will be further elaborated. For now, it should be emphasised that initiating does not consist in simply *causing* actions and intentions. There are two main reasons for this.

First, this would fail to respond to the problem raised by Quinn against the Pure Motivational views of desire. For on such a view, desires would be functional states that merely ‘push and pull’ the subject, even if those brute pushes tend reliably to direct her toward the good. Second, this account would fail to respect some important differences between

the way in which desires give rise to action and intention and perception to belief. One key difference is that in most cases, a perception that p gives rise to a perceptual belief that p without first being evaluated or ratified.^{113,114}

However, intention-formation on the basis of desire cannot always be so direct (though this is not to say that they need *always* be ratified by an act of the will before being acted on, or initiate intention formation). Intentions are elements within larger structures, *plans*, that are involved in the trans-temporal coordination of action (Bratman, 1987, Ch.2-3). If, as default, desires simply gave rise to actions or the formation of intentions, agents would be constantly violating norms of planning rationality (Bratman, 2009). Desires then would unacceptably *disrupt* rational agency. Whether an action or intention is rationally formed when based on a desire will depend on whether some competent sensitivity was exercised to incorporate it within one's overall plan.

The central aim of the following chapter will be to explain how the teleo-functional account of desire outlined here can fruitfully supplement an account of the rationality of acting on desire. On the way to this account, it will be necessary to back up and revisit some fundamental questions concerning practical rationality that was foregrounded in Chapter 1. The following question will be pursued:

- How does acting rationally relate to our capacity to *act for reasons*?
What is it for one to act in light of an apparent reason for action?
(§6.2)

Tackling this question will put us in a position to see how various mistaken views about what acting for reasons must amount to have shaped predominant accounts of the rational significance of desire, e.g. the DAB account of Chapter 4. I motivate an alternative view which emphasise the role of *reasons-sensitive* competences in explicating what it is for one to act in light of an apparent reason (§6.2.3). This will put us in a position to articulate the rational role of desire in *facilitating* our capacity

¹¹³This kind of process is one that psychologist Daniel Gilbert calls 'Spinozan', see Gilbert, Tatarodi, and Malone (1993).

¹¹⁴This is not to say that one *must* form the belief that p in the face of perception. This is illustrated by cases of disbelief in perception. For discussion, see Martin (1993, p. 86).

to respond to reasons (§6.3).

Chapter 6

The Rational Role of Desire

In this chapter, I develop an explanation of the rational significance of desire, one that draws together lessons learnt in criticising various extant accounts of the rational significance of desire. The explanation builds on the work of the previous chapter, establishing an account of the nature and function of our desire-producing mechanisms and of our basic desires. At the heart of this explanation is the idea that we should think of a particular system of conation, our basic desires and the mechanisms that create and regulate them, as comprising a competence to respond non-deliberatively to some of our reasons for action. Thus, we can think of our desires as playing a role in *attuning* us to reasons. The rational significance of an occurrent basic desire can thus be understood as constituting manifestations of a capacity we have to recognise and respond to our reasons.

The plan for this chapter is as follows. I will begin by revisiting the Problem of Conative Significance, summarising what we have learnt from the various accounts we have considered, and distilling an important question that will frame the discussion to follow (§6.1). Given the plausible connection between acting rationally and acting for reasons, the question is one of understanding what a subject's reason is when she acts on a basic desire. I call this the Reasons-Location problem. The following section considers the question of what it is for something to constitute an apparent reason for action and motivates a proposal on which apparent reasons for action are a function of a set of reason-sensitive competences (§6.2). Building on work in Chapter 5, I will argue that we should under-

stand our system of basic desires as one such reason-sensitive competence and that occurrent basic desires constitute manifestations of this capacity to recognise and respond to our reasons (§6.3). This account is put to work to tackle the Reasons-Location problem. Finally, I will tackle an outstanding issue concerning the rational role of desire in certain cases of choice between rationally eligible options (§6.3.3).

6.1 The Road Behind, The Road Ahead

The Problem of Conative Significance is one of explaining how basic desires have the power to rationalise choice and action. What do we have to add to an otherwise plausible account of desires as inputs into practical reasoning in order to explain how they can not only bring about action or intention, but do so in a way that *rationally supports* actions and intentions, doing so in a way that is *intelligible* from the subject's perspective?

It is worth explaining once again the two constraints I suggested should apply to any proposed explanation. According to the *Rational Support Constraint*, answers to the problem must explain how desires can be of genuine normative significance in guiding choice. It seems that the presence of a desire can independently rationalise the choice of the desired option over another eligible option in cases where both options seem 'tied' (Chapter 3). According to the *Intelligibility Constraint*, answers to the problem must explain how desires can guide action and choice in a way that is intelligible from the subject's perspective. As rational creatures, we act for apparent reasons. These reasons, even if merely apparent, constitute the central element in the mutual- and self-intelligibility of actions.

6.1.1 The Reasons-Location Problem

Rather than repeat the various accounts and the respective objections we raised, I think it is possible to distill down a general problem that in one version or another has been a reoccurring theme throughout the thesis. The problem concerns how we should relate two elements in an explanation of the rational significance of desire: the basic desire

which rationalises a relevant action, and the agent's reason to which she responds in ϕ -ing.

Some of the accounts we have considered relate these two elements by identifying the reason grasped for ϕ -ing with the basic desire to ϕ . Chang's account of affective desires discussed in Chapter 5, for example, claims that an occurrent basic desire consists in a certain kind of affective experience which has a phenomenological quality that is taken as a reason to pursue a certain course of action. Whilst this account has the advantage of providing us with an account where desires *per se* play a role in the rationalisation of action, it remains unclear how affective experiences can genuinely count in favour of a certain course of action. Other accounts such as the 'desire as belief' account discussed in Chapter 4 locates the reason grasped for acting as inhering in some objective good-making or reason-giving feature believed to be possessed by the desired course of action. Whilst these accounts have no problem identifying a plausible apparent reason for action grasped as favouring the action, such accounts implausibly makes the rational significance of desire conditional on the subject's believing herself to possess a normative reason for action.

The Reasons-Location Problem is one of trying to understand how to relate the kind of rational support that desires seem capable of providing with the kind of consideration grasped by the subject as a reason for her to pursue the course of action she desires. Through tackling this problem, the account of the rational significance of desire I wish to defend will emerge. The Reasons-Location problem will require that we back up and revisit the notion of an apparent reason that was discussed in Chapter 1.

6.2 Competences and our Apparent Reasons

There is a close connection between practical rationality and apparent reasons. We act rationally when we act in light of an apparent reason. Apparent reasons constitute a rational basis for the formation of an intention or the choice to pursue a certain course of action. So far, we have glossed the notion of an apparent reason as follows:

Apparent Reasons (Schema) R is an apparent reason R for agent A

to ϕ iff there is some state of affairs R such that one stands in the appropriate epistemic relation to R and A ‘takes’ R to count in favour of ϕ -ing.

Despite the abstract and somewhat schematic nature of this formulation, we can identify two components. First, there is some kind of epistemic relation that the agent must stand in to R that puts the subject in a position to respond to that reason. Second, there is some kind of additional relation that the subject has to stand in to R in virtue of which R constitutes an apparent reason for her to respond in a certain way. In Chapter 1, we introduced the notion of ‘taking’ R to be a reason. But what is it to ‘take’ a state of affairs *as* a reason to pursue a course of action?

6.2.1 The De Re View

To get clearer on what role this placeholder notion is playing, it is important to rehearse an objection I considered to a certain account of apparent reasons considered in Chapter 1. In his gloss on apparent reasons, Parfit writes that “we have some apparent reason when we have some belief whose truth would give us that reason” (Parfit, 2001, p. 25). Here Parfit makes it a precondition on some R ’s being an apparent reason that the subject believe that R obtains. We can generalise slightly to include, as other philosophers have supposed, that presentational states like visual perception can also form the relevant epistemic relation. Following Sylvan (2014), we can codify this claim as follows:

De Re View R is an apparent reason for agent A to ϕ iff it appears to A that R , and R ’s truth is a reason to ϕ . (Sylvan, 2014, p. 591)

The De Re View is simply a claim about what is necessary and sufficient for something to be an apparent reason for action. It does not yet say anything about *responding* to one’s apparent reasons, as when we choose to act for that reason, or form an intention on the basis of that reason. According to Parfit, we respond to these apparent reasons if the appearance of these reasons “lead[s] us to do, or try to do, what we have reason to do” (Parfit, 2001, p. 31). When a subject responds to this apparent reason, the relevant epistemic state will enter into an

explanation of her actions, one that rationalises her action. If Parfit is right that a belief that R is the requisite epistemic condition for R to constitute an apparent reason for action, then what explains the agent's ϕ -ing will be, in part, her belief that R .

The problem we raised for The De Re View is that even with these epistemic conditions satisfied, being 'led' to act on the basis of one's belief that R obtains is *insufficient* for that way of responding to be *rational*. This is true even if you ϕ in the belief that R , and R is a reason to ϕ . We considered two examples, one practical and the other epistemic, to show this. The practical example involves a gormless child who believes there is a bear in front of her might be led to do something that is favoured by the state of affairs she correctly believes to obtain. She might for instance make herself look as big as possible in trying to hug the bear, and successfully scare it off. Yet, because what she does is not appropriately sensitive to the *way* in which what she believes favours her action, she does not count as responding to an apparent reason to make herself look as big as possible. The epistemic example involves someone of poor mathematical ability, Tom, who believes that p , where p is some true mathematical proposition. Suppose p entails Q , where Q is some arcane mathematical proposition that Tom is in no position to deduce. Though it is true that the fact that p is a consideration that Tom believes, one whose truth *is* a reason to infer that Q , it simply would not be rational for Tom to infer that Q on the basis that p .¹¹⁵

The role that reason-takings play then is to capture the idea that for a response ϕ to a consideration R to count as a response to an *apparent reason*, the subject must, in some sense, appreciate or understand how considerations of type R favour responses of type ϕ , and be moved to respond in that way *because* of this appreciation. How are we then to capture this idea of a subject's being able to appreciate something as being a reason for action?

6.2.2 The De Dicto View

In reaction to cases that posed a problem for the De Re view, we find in the literature an account which aims to capture this idea by appeal to a

¹¹⁵For further discussion, see Sylvan (2014, §3).

certain kind of normative belief or appearance. Following Sylvan (2014), we can call these *de dicto* accounts of apparent reasons:

De Dicto View R is an apparent reason for agent A to ϕ iff it appears to A that R is a reason to ϕ .

The De Dicto View adds a further necessary condition for a consideration R to be an apparent reason for action; namely that R must appear to the subject *to be a reason*. But notice that this does not yet explain how this differentiating condition figures into an account of acting rationally in a way that will allay the problem facing the De Re view. For this, we can appeal directly to this further condition in explaining proper responsiveness to reasons. The idea will be that when one ϕ -s for the apparent reason R , it is not just in response to the apparent reason R that one ϕ -s, but to an awareness of R 's being a reason for one to ϕ .

We find this view in the work of Joseph Raz. According to Raz, acting for reasons requires that subjects act on the basis of an appreciation that the considerations for which they act are *reasons*. "Agents are", as Raz puts it, "led to awareness of the facts that are reasons *qua reasons*" (Raz, 2011, p. 35). Acting rationally is simply "rational reactions to this awareness" (Raz, 2011, p. 35). How does this awareness figure in an explanation of acting rationally? Raz tells us that rational agency must be such that the notion of a reason "itself figures in an explanation [of the action]" in such a way that "one is motivated by the fact that is the reason, *through the mediating belief*, recognizing it as such" (Raz, 2011, 29, emphasis added). When an agent A ϕ -s in response to an apparent reason R for A to ϕ , what explains A 's ϕ -ing is, in part, A 's belief that R is a reason for her to ϕ .

The De Dicto view constitutes an important advancement over the De Re view. It puts flesh on an important insight: that acting for apparent reasons requires sensitivity to your apparent reason for action's status *as a reason*. However the most persistent objection to the De Dicto view is the charge that it *over-intellectualises* rational agency. (Lavin, 2011; Sylvan, 2014; Saemi, 2014). This objection has been framed in a range of ways. Parfit (2001, p. 118) for instance argues that the De Dicto view would implausibly put the possibility of responding to apparent reasons out of reach of subjects such as certain animals and infants who lack the

concept of a reason. Though I think the spirit of this objection is on the right track, framing it in this way pulls focus away from the main problem, dragging us into contentious issues surrounding the ascription of conceptual cognition in animals and infants.¹¹⁶

A better response would be to motivate the possibility of responsiveness to apparent reasons without positing appearances with the content that one's grounds for action are reasons for action. Competent or skilful action has this structure. Imagine an experienced mover trying to remove a large bed through a narrow door. Seeing that the bed is the size that it is, she might think "Right, I better unscrew that part and angle the left corner like so". Cases like these are examples of an agent's being motivated by apparent reasons. She counts as being motivated by an apparent reason because she responds to the facts as they seem to her, in a way that manifests an underlying competence to treat certain features of the situation as she perceives it as reasons to respond in certain ways. The competence in question reliably results in the subject coming to a view about what one is to do given her take on the facts.

In the case above, the mover relies on her expertise to guide her conduct. It does not seem to proceed on the basis of an evaluation of

¹¹⁶Consider for instance two responses open to Raz. One response claims that such subjects *do* possess the concept of a reason, e.g. by appealing to some less demanding standard for the ascription of conceptual thought, for instance the ability to "[react] to [would-be reasons] appropriately and trying to adjust... responses when becoming aware that they are inappropriate" (Raz, 2011, p. 32). This response, as it stands is unappealing. First, it is unclear if this move jeopardises what is distinctive about the De Dicto view. Second, this suggestion would require substantial work to motivate. A relatively common assumption is that contents are conceptual to the extent that it exhibits sentence-like structure, that is, compositional structures formed from intra-propositional semantic elements like predicates and subjects. The bar for ascribing conceptual thought will depend on (i) how much compositional structure is necessary for conceptuality, (ii) whether sub-propositional units need to be fully systematic, recombining freely and generally, and relatedly (iii) whether conceptual capacities must undergird the ability to produce an unbounded number of thoughts. To the extent that this is a necessary condition for conceptual thought, it is clear then that a defender of the De Dicto view does incur a substantial explanatory burden since it is unclear that the kinds of dispositions Raz appeals to could satisfy even the lowest bar for conceptuality in thought (cf. Sylvan, 2014). A second response would be to *deny* that lower animals can respond to apparent reasons in the way that mature human subjects can. On such a view, lower animals who fail to respond to appearances of their grounds as reasons can only act *as if* for apparent reasons and thus merely approximate full-blooded rational sensitivity to reasons. This strikes me as unattractive, for it is plausible that certain animals *can* respond to facts in a way that manifests a sensitivity to favouring relations. Why deny that they can respond to apparent reasons?

the position of the bed-frame as justifying her angle of approach. To see this, it need not be the case that the subject is willing, or indeed, able to defend her conduct by reference to a reason for action. Yet this is what we would expect if her action is done in the belief that her grounds for action are reasons for the approach she takes. To insist that her response must be mediated by reflection about those facts constituting *justifications* for her response would be to freight her psychology with an unnecessarily reflective component.

The De Dicto theorist is right to emphasise the fact that rational agency is motivated by a sensitivity to one's reasons for action. However, the problem is that sensitivity to reasons need not take the form of a belief about what is or is not a reason for action. In the following section, I develop a framework for thinking about rational sensitivities to reasons. The account draws on some ideas from virtue epistemology in making use of the notion of *reason-sensitive competencies*. The result will be an account of one's apparent reasons and the rational significance of these apparent reasons for rational agency that avoids some of the problems that faced the De Dicto view.

6.2.3 The Competence View

I want to begin with the notion of competence, in particular, the notion of competence that features prominently in the virtue-theoretic epistemology developed by Ernest Sosa.¹¹⁷ Let us start with Sosa (2007, p. 22)'s well-known example of the archer which he uses to illustrate the kinds of normative assessment we use to evaluate performances. We can isolate three levels of performance normativity. When an archer aims to hit the bullseye, the archer's performance is *correct* or *accurate* if the arrow hits the bullseye. We can say that the archer is *competent* or *adroit* (relative to the activity of archery) if, in taking the shot, the archer manifests a disposition to succeed in hitting the bullseye. Combining these two evaluations, we can say that the shot is *apt* if the archer's shot hits the bullseye *because* it manifests the archer's competence.

Here I want to draw a parallel between (i) the relationship between competence and correctness and (ii) the relationship between rationality

¹¹⁷See Sosa (2001), Sosa (2007), Sosa (2010) and Sosa (2015).

and correctness. The archer can manifest his competence at archery yet still fail to achieve correctness, as when a gust of wind blows his arrow off course. Similarly, a subject can manifest rationality in acting yet still fail to achieve correctness, as when one drinks what takes to be a gin and tonic when it is in fact poison. Rationality, like competences in general, involves a *fallible* disposition to succeed in a certain respect; that is, the manifestation of these dispositions does not entail success.

An attractive view of rationality develops this parallel by construing rationality as a function of a subject's competency to respond to her normative reasons for action (cf. Sylvan, 2014; Kauppinen, forth.) On this framework, we see rational agency as an activity in which one aims to act on the basis of sufficient reason to ϕ . We can replicate the same three-level evaluative structure. A subject, in acting in way ϕ , acts *correctly* if she ϕ -s in accordance with the balance of all normative reasons there are for A . She acts *competently* in ϕ -ing if she, in ϕ -ing, manifests a disposition to succeed in the activity of acting on the basis of sufficient reason. Call this her reasons-sensitive competence (we will discuss in more detail shortly). And she acts *aptly* if she ϕ -s correctly because, in ϕ -ing, she manifests her reasons-sensitive competence.

Several general features of rational competence with reasons are worth pointing out. First, since competences are dispositions to succeed, reasons-sensitive competences ground a disposition "to treat considerations like objective reasons to do ϕ -like things only if they are objective reasons to do ϕ -like things" (Sylvan, 2014, p. 604). Thus, given that we are aiming to build an account of rationality by reference to such fallible competences, rationality retains an indirect connection to objective normative reasons for action without collapsing into either correctness or coherence (Sylvan, 2014, p. 605). Second, reasons have weights and stand in dynamic relations to each other. For example, (i) some considerations count in favour of taking courses of action that are vastly outweighed by other considerations, (ii) some consideration strengthen, weaken or disable the favouring strength of other considerations and so on. Correspondingly, our reason-sensitive competencies will also have to ground dispositions to respond in different ways that reflect these different weights and the way would-be reasons dynamically interact with each other (Kauppinen, forth, p. 25). Finally, I think it is plausible that we should think of our

overall competence to respond to reasons *as reasons* in terms of a *network* of local competences (Kauppinen, forth, pp. 26–7). Consider the mathematical example earlier. Mathematical competence with respect to developing complex proofs is something that one develops over the course of one’s lifetime, something which consists in the acquisition of a set of skills distinct from the skills operative in other domains where one can exercise reasons-sensitive competencies (e.g. one’s underlying perceptual competences used to detect flavour notes in wine).

We can appeal to the notion of a reasons-sensitive competence to good effect to account for the distinction between merely acting in accordance with one’s reasons and responding to one’s apparent reasons. Recall the case of the mover where a natural description of her motivation is one of seeing certain features of her situation as ‘calling for’ responses. Seeing that the bed frame is of a certain dimension (bigger than usual), she is inclined to take a certain kind of approach (removing certain awkward parts and tilting it a certain way). From the perspective of the practically-engaged agent, certain considerations R form the object of an attraction to treat R as a reason to pursue a particular course of action (cf. Sylvan, 2014, pp. 598–9). What makes this way of responding to reasons *rational* is the fact that this attraction to treat certain considerations as reasons is the manifestation of an underlying fallible competence to treat certain considerations as reasons for action only if they *are* reasons for action.

We can analyse competency-manifesting *attractions to treat* in functionalist terms. Reasons-sensitive competences ground dispositions to be *occasionally motivated* to treat how things seem to the subject as reasons for action. How things seem to the subject is, in part, a product of the subject’s perceptual apparatus, e.g. how things look, smell, sound, and so. It can also include cognitive seemings, e.g. that p seems to follow from Q . Thus, we can think of reasons-sensitive competencies as taking as operating on perceptual and cognitive representations and producing as output inclinations to treat what is represented as favouring acting in certain ways.

This view strikes the right balance between the De Re view and the De Dicto view without their shortcomings. The De Re view is right that when we rationally act for reasons, we respond to the facts, or at least,

what seems to us to be facts in representing certain states of affairs as obtaining. But for the response in question to be rational, the agent must be sensitive to the *way* in which those facts favour certain actions. The De Dicto view tries to remedy this by appealing to the mediation of *beliefs* about how those facts relate to what reasons we have to act. But being motivated by normative beliefs is just one of many ways we have to be sensitive to which facts favour which responses.

On the Competence View, this sensitivity is built up of a network of local reasons-sensitive competencies which the agent relies on in guiding action. One key benefit of the Competence View is that it clearly avoids the issue of over-intellectualisation. On this view, we can act rationally in response to our apparent reasons without first coming to a view that certain grounds constitute reasons for action. This feature nicely accommodates the respect in which certain pre-conceptual subjects might respond to apparent reasons, whatever we think about their capacity to entertain beliefs with conceptually structured contents. It is necessary only that their responses are mediated by competence-manifesting attractions to treat those considerations as reasons. Provided the disposition to treat certain considerations as reasons is a manifestation of a reasons-sensitive competence of the kind discussed, then there is no obstacle to granting the resulting actions of lower animals the status ‘rational’.

Before moving on, I want to clarify that though I take the Competence View to be highly credible, strictly speaking, I need only maintain a weaker claim entailed by such views for present purposes, *viz.* that *some* apparent reasons are a function of reasons-sensitive competences.

6.3 The Rationality of Acting on Desire

In this section, I explain how the teleo-functional account of the nature of desire developed in Chapter 5 can be marshalled to provide an account of the rationality of acting on desire. The account to be developed begins by explaining how the conative system responsible for the formation and regulation of our basic desires grounds a competence to respond to a certain class of reasons. Next, I suggest that occurrent basic desires are a manifestation of this reasons-sensitive competence and explain how occurrent basic desires operate in such a way as to ground

reasons-responsive action. This account will explain the role that desires play in rationally supporting intention-formation and action as well as resolving the Reasons-Location problem discussed at the outset of this chapter.

6.3.1 Desires and Reasons-Sensitive Competence

We can split the exposition of how the conative system to which our basic desires belong contribute to our capacity to respond to reasons into two parts: (i) the underlying constitution of this system, the ‘seat’ of this competence as Sosa puts it, and (ii) the operation or activity of occurrent desires, which I will claim should be understood as a manifestation of this competence.

The Seat of the Competence

Competences are fallible dispositions to succeed in a certain kind of activity. Sosa identifies several levels to a competence, one of which he calls the ‘seat’ of this disposition to succeed. For example, someone who is competent at driving has a disposition to succeed that is in part due to an “innermost driving competence that is seated in one’s brain, nervous system, and body” (Sosa, 2017, p. 191)

I want to suggest here that we should think of the conative system to which basic desires belong as constituting the seat of a fallible disposition to succeed in responding to a certain class of reasons for action, ones connected with biological goods. This system comprises our innermost competence to respond to such reasons. On this front, we can draw on the work developed in Chapter 5. Here I claimed that our mechanisms for desire production are regulated in ways that ensure they are reliably conducive to the satisfaction of core biological norms. Their function or aim is to produce standing desires for outcomes that are biologically good. In turn, the function of these standing desires are to motivate the subject to act in ways that are normally good for the subject. Since outcomes that would be good for the subject to pursue are those that the subject has some reason to pursue, we can think of our system of basic desires as a reliable disposition to motivate the subject to act in ways that are responsive to these reasons.

The Activity of an Occurrent Desire

How does this system of conation allow us to respond to such reasons? In order to answer this question, it will be necessary to consider the activity of occurrent basic desires. We can illustrate this by considering a simple case. Suppose you are at the greengrocer and you are looking for a small snack to eat. You look across a vast array of fruit; the punnet of delicious bright red plump strawberries catches your eye and you form an inclination or attraction to eat those strawberries. We can highlight two important stages in this process.

The first stage involves the activation of a standing desire which gives rise to an occurrent basic desire for those strawberries. Standing desires are psychological structures we retain even through long periods of inactivity, depression, or loss of consciousness. A subject can have a standing desire to eat strawberries even if he is in a deep depression and spends all his time sleeping. Under certain conditions, standing desires give rise to *occurrent* desires. When we perceive the plump red flesh of the strawberries, we represent a certain kind of outcome (eating strawberries) as available (cf. Butlin, 2017, pp. 619–22). This combines, under certain conditions, with someone's standing desires for eating strawberries, thereby generating an occurrent desire to eat those strawberries. Occurrent desires exist over periods of time during which they are active in influencing various processes that lead to action.

The second stage, the one that I want to focus on, concerns the way in which this occurrent basic desire influences the processes that lead to choice and actions. There are two important aspect of the functional role that occurrent desires play: an attentional component, and an action-guiding component. These two elements of an occurrent desire combined, I believe, captures the phenomenology of an occurrent desire.

The attentional component of an occurrent basic desire involves a disposition to have one's attention exogenously drawn to salient features *F* of an outcome that one represents, or an object one perceives. For example, when one has an occurrent basic desire to eat strawberries one is disposed to have one's attention drawn to features such as the plump red flesh of the strawberries one perceives before one, or to memories of the flavour of the strawberries when one imagines eating strawberries.

Let us call the item possessing the features that captures one's attention, the *attractant*.¹¹⁸

The action-guiding component of an occurrent desire involves a non-alienating felt impulse to treat the attractant's having a property F as favouring a certain response; that is, as a *reason* for a certain response ϕ .¹¹⁹ To treat an attractant having property F as a reason to ϕ is for the representation of the attractant's being F to influence practical reasoning by assigning some weight to the option of ϕ -ing. For example, in virtue of one's occurrent desire to eat strawberries, one feels an impulse to treat the fact that the strawberries have the appearance they do as contributing some weight to ϕ -ing in the course of determining what it is one should do. This I think captures Jennifer Hawkins' suggestion that desires involve a non-conceptual "evaluative impression" glossed as "a very simple kind of experience, a primitive feeling as of certain responses making sense or feeling right" (Hawkins, 2008, p. 259).

At the end of Chapter 5, we outlined an outstanding issue concerning the way in which occurrent desires initiate the formation of intention or an action, specifically, how occurrent desires can give rise to intentions or actions in ways that do not disrupt rational agency. On the account of the way in which occurrent desires prompt action described here, occurrent desire initiates action or the formation of an intention not simply by causing it, but by drawing our attention to what are, ideally, reasons which can then inform a rational choice to act on a desire. For example, the subject with an occurrent desire to eat strawberries has her attention captured by features such as the plump delectable appearance of the strawberries in the greengrocer which she is inclined to treat as a reason to do what she believes will allow her to consume them. That the strawberries have these features operates as a reason, a reason to be weighed up in practical reasoning alongside other reasons. If there

¹¹⁸We have already considered an account of desire that emphasises the role of desire in directing one's attention. Tim Scanlon for instance locks on to the way in which an occurrent desire to buy a new model of computer directs one's attention to features of the new model and "taking [these] features to count in favor of buying them" (Scanlon, 1998, p. 43). Scanlon, recall, accounts for this in a way that I disagree with as involving "[a] tendency to judge that I have reason to buy a new computer" (Scanlon, 1998, p. 43). I will compare my view with Scanlon's shortly.

¹¹⁹Compare Schroeder (2007, p. 159) where Schroeder similarly considers the way which desires involves a disposition for certain conditions to "strike you in a certain phenomenologically familiar way", as non-alienating prompts to action.

appears to her to be a sufficient reason for her to act on this apparent reason, then a choice to act on this apparent reason would be rational. In this way, occurrent basic desires do not disrupt rational agency by compelling certain action, but operates through the standard processes of practical reasoning.

Occurrent Desires as Manifestations of a Reasons-Sensitive Competence

We can now connect these two elements: the characteristic activity of an occurrent basic desire and the conative system responsible for the production and regulation of such desires. These two elements stand in the kind of relation we observe in reasons-sensitive competences in general. On the account of apparent reasons we considered in the previous section, a consideration R is an apparent reason for an agent A to ϕ if it appears to A that R and A is competently attracted to treat R as a reason to ϕ . A *competent* attraction to treat R as a reason to ϕ is one that is a manifestation of a reasons-sensitive competence; that is, an underlying competence to act in ways that *are* favoured by the obtaining of certain facts.

The present suggestion is that occurrent desires are manifestations of a reasons-sensitive competence. To have an occurrent desire to ϕ is for one to be in a state which grounds an attraction to treat certain features of an outcome we represent in perception or thought as a reason to ϕ . For example, to have an occurrent desire to eat strawberries is for one to be attracted to treat some salient feature of the strawberries one perceives as favouring a certain kind of response. This attraction is a *competence* manifesting attraction because the desires that ground such attractions are the products of mechanisms that have the function to produce and regulate our desires in good-conducive ways. This competence thus provide the connection between the attractions that our occurrent desires ground and our normative reasons for action: they are manifestations of a fallible dispositions to respond to a certain class of normative reasons for action. It is this etiological connection between an agent's occurrent desires and her broader conative system that makes it *rational* for a subject in the grip of an occurrent desire to treat some

feature of a represented outcome as a reason for action.

Revisiting the Two Constraints

It is worth summarising how this account satisfies both the Rational Support and Intelligibility Constraint. The rational support constraint is satisfied by appealing to occurrent desires as having a nature and content which allows them to guide one toward what is good by drawing one's attention to considerations that, ideally, are a certain kind of reason, one pertaining to biological goods. In virtue of our system of basic desires having an indirect connection to what is good for the subject, we can provide an explanation of why rationality (defeasibly) favours one's doing what one believes will lead to the satisfaction of one's occurrent desires. The intelligibility constraint is satisfied by appealing to the way in which such desires attune you to an apparent reason for action. When a subject performs an intentional action, her action is intelligible to her in virtue of there being an apparent reason for which she acts. On this account, when one acts on a desire, one acts for the reason that one's occurrent desire attunes one to the presence of.

Now that we have an account of how these two constraints are normally satisfied, it would be instructive to revisit Quinn's Radioman case and try to explain some of the intuitions that were canvassed.

The first element concerns the nature of his strange desire. Plausibly, folk psychology contains a rough understanding of how desires are formed and, given a shared environment, an expectation of the normal set of basic desires subjects can be expected to form. What strikes me as, in part, responsible for our intuitions about Radioman is that it is not clear how, over the course of a lifetime, one could form intrinsic desires for such things. So we lack an explanation of *how* these strange motivational states are formed and indeed, whether they are formed in the course of Normal functioning of his DPMs. The oddness of the desideratum constitutes some reason to doubt that this motivational state is the product of a *Normally functioning* DPM. Thus, whether this motivational state can rationally support action comes into question. Indeed, it is unclear if the state is the product of a DPM at all. Given that desire are individuated etiologically, we can explain why highly bizarre motivations ground

the proprietary questions about whether they qualify as ‘desire’.¹²⁰ This second element focusses on Radioman’s perspective. Radioman is compelled by a brute urge or compulsion, a motivational state that fails to provide him with a perspective from which his acts make sense. What Radioman lacks is awareness of some apparent good-making feature of turning on radios that renders his actions intelligible. There is no feature for which he feels a felt inclination to treat as a reason and so lacks some of the phenomenological elements of a normal desire. For this reason, though Radioman finds himself motivated to turn on radios, he lacks the kind of desirability characterisation that normal desires typically avails one of.

6.3.2 The Reasons-Location Problem

In this section, I will consider the Reasons-Location problem raised at the start of this chapter and explain how this account of occurrent desires as manifestations of a reasons-sensitive competence allows us provide a neat story about the nature of the agent’s reasons for action when she acts on the basis of a desire. The Reasons-Location problem is generated by two simple claims. First, when a subject forms an occurrent basic desire, this desire is capable of rationally supporting choice and action. Second, rational subjects are those that act for apparent reasons. The Reasons-Location problem is concerned with tying these two elements together: when a subject acts on a basic desire that rationally supports her choice to pursue a course of action, what is the apparent reason for which a subject acts?

The solution to this problem can be found by attending to the notion of occurrent desire. The characteristic activity of an occurrent desires I have claimed is to ground an attraction to treat certain features of a represented outcome as a reason for action. On my view, we should identify the reason for which she acts as the features of the objects our hypothetical outcomes she is inclined to respond to in a certain way, as a result of the occurrent desire. For example, in the case of the occurrent desire for strawberries, the agent’s reason for selecting the strawberries

¹²⁰Here there is an echo of Anscombe’s point that intelligibility of the objects of desire coincide with the limits of desire ascription.

is the plump red appearance of the strawberries, or the anticipated taste of eating them.

It is worth comparing this position to the two standard approaches to the problem I identified in order to articulate the respects in which this account is in a position to avoid the problems faced by these other accounts.

One approach was to identify the agent's apparent reason for action when she acts on an occurrent desire with the fact that she has the occurrent desire. Here the thought was that we can seek to explain the rational significance of such desires by features that attend to having such an occurrent desire. The problem with such an approach however is that if my description of the activity of an occurrent desire is right, then most cases of acting on a desire involve responding to the worldly features of the desideratum that we are inclined to treat as reasons, *not* the fact that one has a certain desire. We can sharpen this point by drawing on a distinction due to Philip Pettit and Michael Smith concerning two ways in which desire can play a role in rationalising choice. In acting on a desire to ϕ , a desire to ϕ "figures in the *background* if and only if it explains the agent's choice of option [to ϕ]". In contrast, a desire to ϕ "figures in the *foreground* if and only if the agent reaches the choice via the recognition that he has that desire and that the option has the desirable property— the property justifying its choice— of promising to satisfy the desire" (Pettit and Smith, 1990, 567–8, emphasis added).

I want to suggest a diagnosis of the temptation to identify the agent's reason for action when she acts on an occurrent desire with the desire itself. The temptation I suggest comes from an elision of the thought that when one acts on an *occurrent* desire to ϕ that the occurrent desire thereby figures in the *foreground* of her choice to act. It does not follow from the fact that one acts on an occurrent desire to ϕ that one's choice to ψ proceeds via the recognition that one has a desire to ϕ which would be satisfied by ψ -ing. It is of course possible for an occurrent desire to lie at the foreground of a choice, however the point to be stressed here is that it need not. Many, if not most, cases of acting on the basis of an occurrent desire lack such a self-reflective component. In these cases, the activity of an occurrent desire comprises part of the causal explanation of that

action and so figures in the *background*.¹²¹ What lies at the foreground are the worldly features of the desideratum that we are inclined to treat as reasons.

Another approach identified the agent's reason for action when she acts on an occurrent desire with some objective feature of the desired course of action. The account I defend here is a version of this approach since I claim that acting on an occurrent desire involves responding to the features of a represented outcome that one is inclined to treat as a reason. However, I want to emphasise some differences between this version and the other versions of this approach we considered in Chapter 4.

According to the desire as belief view (DAB) defended by Gregory, a desire to ϕ is a belief that one has a normative reason to ϕ . According to the directed attention account of desire defended by Scanlon, a desire to ϕ involves an insistent tendency to see various considerations, or features of represented outcomes, as reasons for action. Despite their differences, both views maintain that desires involve a certain kind of representation with a normative content. In the case of Gregory, it involves a belief *about* reasons; in the case of Scanlon, it involves a feature or consideration quasi-perceptually seeming to be a reason. In this respect, both adhere to the De Dicto view of the apparent reasons that desires attune us to. I have already pressed the worry that this view of desire leads to an over-intellectualised account of desire as involving a kind of self-location in the space of reasons. I have suggested that the De Dicto view of apparent reasons in general is not a necessary condition on something's being an apparent reason.

The account developed here avoids this issue of over-intellectualising desire altogether. When one responds to an occurrent desire to ϕ , one responds to features one represents as possessed by the desideratum. One need not represent that these features constitute reasons for action and so it is not a requirement that the subject conceptualise her reason for

¹²¹This point is established by the possibility of occurrent desires that are not available to the subject's conscious awareness. Take an example due to Mele (2003) of a subject who is motivated by a desire to hurt the feelings of another. In so far as the desire is active in shaping the subject's actions, it is occurrent. However being an unconscious desire, the decision to say something hurtful is not done in recognition of the desire; that is, the relevant desire does not figure in the foreground of choice.

action *as* a reason for action. For example, when one responds to an occurrent desire to eat strawberries, one's reason for choosing strawberries are facts like the strawberries's plump red flesh and *not* the fact that these properties of the strawberries constitute reasons for one to choose them.

6.3.3 Putative Tie-Breaking Cases

There is one final issue to tackle concerning the putative tie-breaking cases involving desire we considered in Chapter 2 to motivate the thought that desires can rationally support choice. We also considered them in Chapter 3 where Chang appeals to a version of such cases to motivate her view that affective desires constitute reasons for action. This is especially pressing given that the position developed here is on the side of accounts that identify the agent's reason for action when she acts on an occurrent desire with the objective features of the course of action she desires to pursue. I want to start by briefly outlining these tie-breaking cases; in particular, I think it is important to distinguish between two types of putative tie-breaking cases. Failure to do this might lead one to illicitly assume that what holds in one variant holds in the other.

Type-1: At t_1 , options X and Y are both rationally eligible, and the subject takes there to be sufficient reasons to choose X , or to choose Y . At t_2 , the subject has a desire for X .

Type-2: At t_1 , options X and Y are both rationally eligible, both are identical in all good-making respects. The subject takes there to be no more reason for X over Y , vice versa. At t_2 , the subject forms a desire for X .

We considered a type-1 case in Chapter 2 involving a subject Uma who is faced with the choice of two options, for instance, eating a sandwich or some soup at t_1 . At t_1 , the apparent balance of reasons does not decisively favour sandwich over soup because she is correctly aware that there are sufficient reasons for either option. Both options are rationally eligible. At this point, it is perfectly intelligible to employ a random decision making procedure such a coin flip to break the tie, or simply to pick one (Morgenbesser and Ullmann-Margalit, 1977). We are then

to imagine that Uma forms an occurrent desire for a sandwich at t_2 . A compelling intuition we noted was that the onset of this desire seems to change her practical situation. It is plausible it would now be less than (ideally) rational for one to choose the soup over sandwich. We bolstered this by teasing out an auxiliary intuition that it would be *irrational* for Uma to disregard the desire and employ some random decision making procedure, e.g. a coin flip. The central difference between type-1 and type-2 cases is that type-2 cases involve the subject's taking *there to be no more reason to pick one option over the other*. The easiest way to secure this is by having two type-identical options as in the case of Buridan's Ass. For this reason, whilst type-1 cases are commonplace, type-2 cases are comparatively rare. We discussed these cases in Chapter 3 as they were employed by Chang to motivate her proposal.

Chang on Desires as Tie-Breakers

In order to clarify my position on such cases, it would be instructive to begin by comparing it with Chang's account of such cases in order to see where our positions diverge. Chang's view is that in both cases what it is rational for you to do has shifted because the onset of the occurrent affective desire for X at t_2 shifts the apparent balance of reasons in favour of the choice of X : the fact that you now have an affective desire for X constitutes a tie-breaking reason that you possess to choose X . The account I develop proposes a different explanation. The central difference is that whilst I accept that in both types of case an occurrent desire for X rationalises choosing X over Y , I deny that this is always to be explained in the way Chang proposes. Whilst an occurrent desire for X rationalises choosing X over Y , it does not necessarily do so by altering the balance of apparent reasons.

In order to justify this divergence, I want to consider the argument that Chang uses to defend the claim that desires constitute tie-breaking reasons for action in type-1 (Chang, 2004, p. 81) and type-2 cases (Chang, 2004, p. 83). Chang's argument begins with two intuitions that need explaining:

- A. If the agent acts *with* the recognised affective desire for X by choosing X over Y at t_2 , then her choice of X is rationalised.

- B. If the agent were to act against her desire by choosing Y over X at time t_2 , then she acts in a puzzlingly way that deviates from what seems to be the rational thing to do.

On this basis, she moves to the claim that:

- C. If one has an occurrent desire to ϕ , then one has an apparent reason that favours one's ϕ -ing over ψ -ing, viz. that one has a desire to ϕ .

As I understand Chang, she seems to be proposing (C) as the best explanation of (A) and (B). Interestingly, the movement from claims (A) and (B) to claim (C) is one that even philosophers who would be opposed to Chang's conclusion have been tempted to make.¹²² In her discussion, Chang flits between discussions of type-1 and type-2 cases and deploys the same line of reasoning to motivate claim (C) in both cases, thereby providing a uniform explanation of (A) and (B) across type-1 and type-2 cases.

An Alternative Explanation

Here I want to suggest a problem for this line of reasoning. The problem is that where (A) and (B) are plausible, there are alternative explanations of (A) and (B) and so (C) remains unmotivated across the board. Specifically, I will argue that:

- There is a plausible alternative explanation of (A) in both type-1 and type-2 cases.
- There is a plausible alternative explanation of (B) in type-1 cases, and that (B) fails to hold in type-2 cases.

¹²²Here I have Raz (1998) in mind. Chang appeals to this line of argument to argue against a pure value-based account of practical reasons which Chang attributes to Raz in defence of a hybrid account of practical reasons (for discussion, refer back to §1.1.3). Chang points out that Raz and others concede this point. The sandwich-soup case sketched is due to Raz. Raz considers a type-1 case involving the options of a pear or a banana, writing “[i]f when offered a pear or a banana, I have reason to take one and it does not matter which one, then if I want the banana but take the pear, I have acted *irrationally*” (Raz, 1998, p. 62). From this he concludes, making a striking concession, “[i]n these circumstances, *wants are reasons*, though in being limited to this case, they are very peculiar reasons” (Raz, 1998, 62, emphasis added).

I will start by considering (A) in type-1 and type-2 cases. This chapter avails one of a plausible alternative explanation of why ‘going with’ one’s occurrent desire is rational: desires are manifestations of a capacity to respond to reasons that bear on what would be practically good for the subject. When one acts on an occurrent desire, one acts for the reason that the desire draws our attention to. This is an adequate explanation of (A), one that does not appeal to the desires themselves as tie-breaking reasons for action.

I add one caveat. I do not think that Chang is wrong to think that an occurrent desire *could*, specifically in type-1 cases, shift the apparent balance of reasons, but *contra* Chang, this is not because (C) holds. In the sandwich-soup case one is aware at t_1 of enough of the features of both options for either to be supported by sufficient reasons. Yet, it is possible that the operation of an occurrent desire can make it seem to one that there is more reason to pick sandwich over soup. In having an occurrent desire for the sandwich, one is inclined to treat the attractive properties of the sandwich as a reason. What happens as a result is that only a *subset* of the total set of reasons at play are at the foreground of the agent’s choice. In this way, it might *seem* to a subject with a desire for the sandwich that there is more reason to select the sandwich, but this is simply due to the contingency of selective attention paid to just one of her options of choice as her function of her occurrent desire. But notice that this psychological explanation falls short of (C); we need not suppose that the apparent balance of reasons is shifted because the *desire* itself constitutes an apparent tie-breaking reason.

Turning now to (B), the intuition that it would be irrational to act contrary to desire in such cases. In motivating the Problem of Conative Significance, I relied on intuition (B) in type-1 cases. I will argue here that this intuition can be explained on the account of desire developed in this chapter, but I will raise some doubts as to whether (B) holds in type-2 cases. Does an explanation of (B) require that desire themselves constitute apparent tie-breaking reasons? It does not. In this counterfactual, the subject fails to respond to desire for no countervailing reason. This constitutes some evidence of a resistance to respond to desire in the default rational way that is dysfunction-indicating. The situation is akin to one where a subject perceives that p yet fails to form the belief that

p without reasons to disbelieve her perception.¹²³ Evidence of dysfunction to respond to psychological states that provide rational support is evidence of irrationality. So it is not the case that the *only* explanation of the irrationality is one that cites acting against where one takes the balance of reason to point. Note however that if the psychological fact about selective attention is right, then it is clear that there is an additional count of irrationality of which the subject is guilty, for she would be acting against the balance of apparent reasons. (And to re-emphasise a point already made, in such cases, we need not suppose that it is the *desire itself* that shifts the apparent balance of reasons by constituting a tie-breaking reason.)

What about the status of (B) in type-2 cases? Chang writes:

If [the subject] goes left, his action can be rationalised by pointing out that he felt like having the bale on the left. If he goes right, *we would need some explanation for this puzzling act*; what reason does he have to go right given that he would enjoy each of the two identical bales equally and feels like having the one on the left? *What we would be looking for is a reason to go right that counteracts his reason to go left provided by the fact that he feels like it...* The fact that one is attracted to something can *per se* provide a reason to go for it when all other reasons are evenly matched. (Chang, 2004, p. 81)

Here Chang employs the same structure of reasoning from (A) and (B) to (C). The problem is that (B) is far less compelling in type-2 cases. Chang claims that there is a legitimate demand for an explanation of the subject's decision to pick the undesired option, however it is clear that such an explanation can be given. After all, our subject is aware that either option is just as good as the other. Compare the following case. Suppose a thirsty subject looking for water is at a crossroads armed with a water detector. A truthful farmer who passes her informs her that both roads lead to bodies of potable water that are identically sized. Imagine now that the water detector lights up, indicating that there is water on the left. It is not at all clear why a subject would be irrational

¹²³We are given no reason why the subject would resist his attraction; indeed, that only goes to bolster the sense that this unexplained resistance is dysfunction-indicating, a breakdown in the kind of reasons-sensitive competences that comprise our capacity for practical rationality.

in going right; after all, she now knows that both roads lead to water. This case is analogous with the subject who finds herself in a type-2 case. To the extent that the subject takes there to be no more reason for one option than the other, then it is not at all clear that the subject *would* be irrational in just arbitrarily picking one option over the other. Notice that the kind of explanation of (B) given in type-1 cases need not apply in type-2 cases. This is because simply picking *in the knowledge that neither option is better than the other* does not evidence an unexplained resistance to rely on one's reason-sensitive capacity, for the subject selects an option in the awareness that the good-making features of both options are evenly matched in all salient respects. Like the water detector case, however the subject acts, she does so rationally since she relies on her awareness of a second-order fact about her reasons, viz. that there are no reasons favouring one option over the other.

Conclusion

The central aim of this chapter has been to develop an account of the operation of our system of basic desires to address the Problem of Conative Significance. The core idea is that our system of basic desires constitute a competence to respond to our reasons for action. Occurrent basic desires constitute manifestations of this reasons-sensitive competence, drawing our attention to features of outcomes we represent in perception and thought and grounding an attraction to respond to these features as reasons for action. This account allows us to provide an explanation of the Problem of Conative Significance that satisfies the two constraints we outlined. By connecting the operation of our occurrent desires with a capacity to respond to reasons, we can provide an account of the reasons for which we act when we act on an occurrent desires in ways that avoid the problems that I have raised for extant accounts.

Conclusion

This thesis has been an investigation into the nature of our basic desires and the role they play in our capacity for practical rationality. I want to finish by summarising the central account of this thesis; in particular, I wish to extract several issues that have been threads throughout this thesis. Doing so will provide us with a better sense of the concerns that unite the various problems I have raised for extant accounts in the literature, as well as a better understanding of how dealing with these concerns ultimately has come to shape the account I have developed of the rational role of basic desire.

One of the central aims of this thesis has been to bring a problem into view. Chapter 2 began with the natural thought that, at least pre-theoretically, there are cases where desires have the power to rationally support action and intention-formation. These desires assail us unbidden, exerting an influence on us to pursue a certain course of action. Such cases are not anomalous, but a pervasive part of everyday life. In normal cases, our basic desires motivate us to act in ways that are intelligible from the agent's perspective. The Problem of Conative Significance is one of providing an account of the nature of our basic desires that will explain these intuitions.

I have developed an account according to which our capacity to form basic desires comprise a part of our rational capacity to respond to our reasons for action. The foundation of this account consists of a teleo-functional elaboration of desire as a mental state individuated partly by its function within the agent's psychology. In Chapter 5, I suggested that our standing basic desires are the products of mechanisms that have the direct proper function to reliably produce standing desires that satisfy core biological functions (§5.2.2). These mechanisms produce new stand-

ing desires operating in a way that reliably tend to lead to the formation of desires for outcomes that are good-conducive (§5.3.2). Our standing desires too have a proper function: to initiate the formation of intentions and actions by interfacing with the subject's cognitive and perceptual capacities to cause the formation of occurrent desires. When a desire is occurrent, it drives psychological processes that engage the subject by drawing her attention to features of outcomes that she represents in perception or thought. Awareness of these features is then claimed to ground a felt inclination to accord greater weight to various courses of action in practical reasoning, specifically courses of action believed to be means to the desired end (§6.3.1).

In Chapter 6, I extended this account to motivate the view that the system of conation to which our basic desires belong constitutes a competence we have to respond to a class of our normative reasons for actions. Reason-sensitive competences in general are fallible dispositions to succeed in recognising and responding appropriately to our normative reasons for action. I motivated the view that our apparent reasons are a function of the exercise of reasons-sensitive competences (§6.2.3). Occurrent basic desires, I claim, are manifestations of a reason-sensitive competence in virtue of which certain considerations strike us as reasons for action (§6.3.1).

I turn now to the ways in which the foregoing has been shaped by the negative part of this thesis which was the central preoccupation of Chapters 3 and 4. Here I want to focus on isolating a few problems that have shaped my own approach to the Problem of Conative Significance.

In Chapter 3, I considered an approach that aims to explain the significance of desire for practical rationality by appeal to desire's connection with conscious occurrences like pleasure, discomfort and affective experiences. I argued that these views all fail to provide an adequate explanation of the rational significance of desire (§3.1.2; §3.1.4; §3.2.3). What these views get right in looking to sensations and affective experiences is the idea that the relation we stand in to our occurrent basic desires is a passive one. This intuition of passivity is accounted for on my view. Occurrent desires consist in psychological processes that depend on the activation of antecedently possessed standing basic desires. Whether a basic desire is occurrent depends on the presence of certain

conditions outside of the agent's control (e.g. detected physiological conditions) and/or the perceptual or cognitive representations that cue its activation. Moreover, the kind of practical engagement that characterises the agent in the grip of a desire includes features such as the exogenous capture of attention that is out of one's control.

An issue I emphasised in Chapter 6 concerned the way in which defenders of affective views construed the the reason for which one acts when one acts on a desire. As we discussed under the ambit of the 'Reasons-Location Problem', such views identify the reason for which one acts with the fact that one has the desire in question. Whilst this might seem a natural place to locate the agent's reason on the hedonic account, I have argued that, at least in the default case, acting on a desire does not involve the self-directed apprehension of one's own mental states as one's reason for action. The account of desire I have developed aims to remedy this. On my view, occurrent desires shape our practical perspective not by constituting elements in the 'foreground' that appear to justify certain courses of action, but by structuring it: drawing our attention to salient features of the world, putting us in touch with reasons that can then guide action.

In Chapter 4, I considered views that construe desires as being under the guise of the good. The first two views provided a cognitivist account of desires: the first identifying desires with a certain kind of normative belief (Gregory) and the second with the basis of certain attention-directing disposition (Scanlon). Here, I want to focus on the second of these views with which the present account shares some features in order to clarify the central differences between our accounts.

The present account incorporates Scanlon's observation that occurrent desires engage one's attention. An attractive feature of Scanlon's account is that it provides an account of the way in which desires rationalise action without construing desires themselves as the objects of reflection when one acts on a desire. The problem however is that Scanlon frames the relevant transition from desire to action as one in which it seems to the subject that certain states of affairs, or features of states of affairs, constitute reasons for action. This was the kind of normative content that characterised De Dicto views of apparent reasons. Thus, though Scanlon avoids making desires themselves our reasons for action

when one acts on desire, he makes appearances about one's reasons for action into the factor that rationalises action (§§4.3.1). The account I sought to develop avoids this by drawing on the Competence View of apparent reasons, allowing us to avoid the over-intellectualised account of apparent reasons that the De Dicto view would otherwise commit us to (§6.2.3).

The second disagreement, one that applies to both Gregory's and Scanlon's positions is a deeper one concerning the source or etiology of our basic desires. According to Scanlon, we should not think of desires as a motivational source distinct from the capacity we have to form judgements about one's reasons (Scanlon, 1998, p. 40). There is a respect in which I agree with Scanlon: our system of basic desire is a capacity we have to facilitate recognising and responding to our normative reasons for action. However, the central difference is that I want to hold that our system of basic desire is distinct from and isolated from our capacity to form judgements about what reasons obtain. As I argued in §4.3.1, there are pressing worries about whether Scanlon's explanation of recalcitrant desires in terms of a single capacity being exercised in two ways, seeing reasons and judging reasons, is well motivated.

To sum, the project of this thesis has been to provide an account of desire that allows us to appreciate the distinctive way in which they operate in rational animals, with an important role to play in our overall capacity for practical rationality. It has been a central concern of mine in grappling with its central problem to articulate an account of desire that takes seriously both the phenomenological and functional aspects of basic desires and the psychological systems to which they belong.

Bibliography

- Adams, Christopher D. (May 1982). “Variations in the Sensitivity of Instrumental Responding to Reinforcer Devaluation”. In: *The Quarterly Journal of Experimental Psychology Section B* 34.2, pp. 77–98. URL: <https://doi.org/10.1080/14640748208400878>.
- Aguilar, Jesús H. and Andrei A. Buckareff (2010). *Causing Human Actions: New Perspectives on the Causal Theory of Action*. Cambridge, Massachusetts: MIT Press.
- Alvarez, Maria (2008). “Reasons and the Ambiguity of Belief”. In: *Philosophical Explorations* 11.1, pp. 53–65.
- (2010). *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press.
- (2016). “Reasons for Action, Acting for Reasons, and Rationality”. In: *Synthese*, pp. 1–18.
- Anscombe, G. E. M. (1957). *Intention*. Cambridge, Massachusetts: Harvard University Press.
- (1967). “On the Grammar of ‘Enjoy’”. In: *The Journal of Philosophy* 64.19, pp. 607–614.
- Anselme, Patrick and Mike JF Robinson (2016). ““Wanting,” “Liking,” and Their Relation to Consciousness.” In: *Journal of Experimental Psychology: Animal Learning and Cognition* 42.2, p. 123.
- Archer, Avery (2013). “The Rational Significance of Desire”. PhD Thesis. Department of Philosophy: Columbia University.
- Arpaly, Nomy and Timothy Schroeder (2013). *In Praise of Desire*. Oxford: Oxford University Press.
- Bain, David and Michael Brady (Mar. 2014). “Pain, Pleasure, and Unpleasure”. In: *Review of Philosophy and Psychology* 5.1, pp. 1–14.

- Baker, Derek (2014). “The Abductive Case for Humeanism over Quasi-Perceptual Theories of Desire”. In: *Journal of Ethics and Social Philosophy* 8.2, pp. 1–29.
- Balleine, Bernard and Anthony Dickinson (Aug. 1991). “Instrumental Performance Following Reinforcer Devaluation Depends upon Incentive Learning”. In: *The Quarterly Journal of Experimental Psychology Section B* 43.3, pp. 279–296.
- Ben Yami, Hanoch (1997). “Against Characterizing Mental States as Propositional Attitudes”. In: *Philosophical Quarterly* 47, pp. 84–89.
- Berridge, K. C. and E. S. Valenstein (Feb. 1991). “What Psychological Process Mediates Feeding Evoked by Electrical Stimulation of the Lateral Hypothalamus?” In: *Behavioral Neuroscience* 105.1, pp. 3–14.
- Berridge, Kent C. (Apr. 2007). “The Debate over Dopamine’s Role in Reward: The Case for Incentive Salience”. In: *Psychopharmacology* 191.3, pp. 391–431.
- (2009). “Wanting and Liking: Observations from the Neuroscience and Psychology Laboratory”. In: *Inquiry* 52.4, pp. 378–398.
- Berridge, Kent C. and Morten L. Kringelbach (May 2015). “Pleasure Systems in the Brain”. In: *Neuron* 86.3, pp. 646–664.
- Berridge, Kent C. and Terry E. Robinson (1998). “What Is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Salience?” In: *Brain Research Reviews* 28.3, pp. 309–369.
- Berridge, Kent C. and Terry E. Robinson (Nov. 2016). “Liking, Wanting, and the Incentive-Sensitization Theory of Addiction”. In: *The American Psychologist* 71.8, pp. 670–679.
- Berridge, Kent C, Terry E Robinson, and J Wayne Aldridge (2009). “Dissecting Components of Reward: ‘Liking’, ‘Wanting’, and Learning”. In: *Current Opinion in Pharmacology* 9.1, pp. 65–73.
- Björnsson, Gunnar et al. (2015). “Motivational Internalism: Contemporary Debates”. In: *Motivational Internalism*. New York: Oxford University Press.
- Block, Ned (1978). “Troubles with Functionalism”. In: *Minnesota Studies in the Philosophy of Science* 9.
- Boswell, Paul (2016). “Affect, Representation, and the Standards of Practical Reason.” PhD Thesis. Department of Philosophy: University of

- Michigan. URL: <https://deepblue.lib.umich.edu/handle/2027.42/135756>.
- Boswell, Paul (2018). “Intelligibility and the Guise of the Good”. In: *Journal of Ethics and Social Philosophy* 13.1, pp. 1–31.
- Boyle, Matthew (2016). “Additive Theories of Rationality: A Critique”. In: *European Journal of Philosophy*.
- Boyle, Matthew and Douglas Lavin (2010). “Goodness and Desire”. In: *Desire, Practical Reason, and the Good*. Ed. by Sergio Tenenbaum. Oxford: Oxford University Press.
- Braddon-Mitchell, David and Frank Jackson (2006). *Philosophy of Mind and Cognition: An Introduction*. Oxford: Wiley-Blackwell.
- Bratman, Michael (1987). *Intention, Plans, and Practical Reason*. Stanford: Center for the Study of Language and Information.
- Bratman, Michael E (2009). “Intention , Practical Rationality ,” in: *Ethics* 119.April, pp. 411–443.
- Brentano, Franz (1874). *Psychology From an Empirical Standpoint*. London: Routledge.
- Broome, John (2007). “Does Rationality Consist in Responding Correctly to Reasons?” In: *Journal of Moral Philosophy* 4.3, pp. 349–374.
- (2013). *Rationality Through Reasoning*. Oxford: Wiley-Blackwell.
- Burge, Tyler (2003). “Perceptual Entitlement”. In: *Philosophy and Phenomenological Research* 67.3, pp. 503–548.
- (Sept. 2009). “Primitive Agency and Natural Norms*”. In: *Philosophy and Phenomenological Research* 79.2, pp. 251–278.
- (2013). *Cognition Through Understanding: Self-Knowledge, Interlocution, Reasoning, Reflection: Philosophical Essays, Volume 3*. Oxford: Oxford University Press.
- (forth.). “Entitlement: The Basis for Empirical Warrant”. In: *New Essays on Entitlement*. Oxford: Oxford University Press.
- Butlin, Patrick (2016). “The Direction of Fit of Desire”. PhD thesis. London: Kings College London.
- (2017). “Why Hunger Is Not a Desire”. In: *Review of Philosophy and Psychology* 8.3, pp. 617–635.
- Butlin, Patrick and David Papineau (2016). “Normal and Addictive Desires”. In: *Addiction and Choice*. Oxford: Oxford University Press.

- Byrne, Alex (2011). “I—Alex Byrne: Transparency, Belief, Intention”.
In: *Aristotelian Society Supplementary Volume* 85.1, pp. 201–221.
- Campbell, Douglas I. (2018). “Doxastic Desire and Attitudinal Monism”.
In: *Synthese* 195.3, pp. 1139–1161.
- Castro, Daniel C. and Kent C. Berridge (Mar. 2014). “Opioid Hedonic Hotspot in Nucleus Accumbens Shell: Mu, Delta, and Kappa Maps for Enhancement of Sweetness ”Liking” and ”Wanting”.” In: *The Journal of Neuroscience* 34.12, pp. 4239–4250.
- Casullo, Albert (2007). “What Is Entitlement?” In: *Acta Analytica* 22.4, pp. 267–279.
- Chang, Ruth (2004). “Can Desires Provide Reasons for Action?” In: *Reason and Value: Themes From the Moral Philosophy of Joseph Raz*. Ed. by R Jay Wallace et al. Oxford: Oxford University Press.
- (2013). “Grounding Practical Normativity: Going Hybrid”. In: *Philosophical Studies* 164.1, pp. 163–187.
- Chierchia, Gennaro (1989). “Anaphora and Attitudes de Se”. In: *Semantics and Contextual Expression*. Ed. by Renate Bartsch, J. F. A. K. van Benthem, and P. van Emde Boas. Foris Publications, pp. 11–1.
- Comesaña, Juan and Matthew McGrath (2014). “Having False Reasons”. In: *Epistemic Norms*. Ed. by Clayton Littlejohn and John Turri. Oxford: Oxford University Press, pp. 59–80.
- Copp, David and David Sobel (2002). “Desires, Motives, and Reasons: Scanlon’s Rationalistic Moral Psychology”. In: *Social Theory and Practice* 28.2.
- Cutter, Brian (2017). “Pain and Representation”. In: *The Routledge Handbook of Philosophy of Pain*. London: Routledge.
- Dancy, Jonathan (2000). *Practical Reality*. Oxford: Oxford University Press.
- Daveney, T. F. (1961). “Wanting”. In: *Philosophical Quarterly* 11.April, pp. 135–144.
- Davidson, Donald (1963). “Actions, Reasons, and Causes”. In: *The Journal of Philosophy* 60.23, pp. 685–700.
- Davis, Wayne A. (1984). “The Two Senses of Desire”. In: *Philosophical Studies* 45, pp. 181–195.

- Dickinson, A. (Aug. 1998). “Omission Learning after Instrumental Pre-training”. In: *The Quarterly Journal of Experimental Psychology Section B* 51.3, pp. 271–286.
- Dickinson, Anthony (2008). “Why a Rat Is Not a Beast Machine”. In: *Frontiers of Consciousness*. Oxford: Oxford University Press.
- (2011). “Goal-Directed Behavior and Future Planning in Animals”. In: *Animal Thinking - Contemporary Issues in Comparative Cognition*, pp. 79–92.
- Dill, Brendan and Richard Holton (2014). “The Addict in Us All”. In: *Frontiers in Psychiatry* 5.Oct, pp. 1–20.
- Dretske, Fred (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, Massachusetts: MIT Press.
- Ellsworth, Phoebe C. and Klaus R. Scherer (2003). “Appraisal Processes in Emotion.” In: *Handbook of affective sciences*. Series in Affective Science. Pp. 572–595.
- Feldman, Fred (1988). “Two Questions about Pleasure”. In: *Philosophical Analysis*. Ed. by D. F. Austin. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 59–81.
- Field, Hartry (1978). “Mental Representation”. In: *Erkenntnis* 13.1, pp. 9–61.
- Fischman, M. W. and R. W. Foltin (1992). “Self-Administration of Cocaine by Humans: A Laboratory Perspective.” eng. In: *Ciba Foundation symposium* 166, 165–73; discussion 173–180.
- Fodor, Jerry (1975). *The Language of Thought*. Language and Thought Series. Cambridge, Massachusetts: Harvard University Press.
- (1983). *The Modularity of Mind*. Cambridge, Massachusetts: MIT Press.
- Frost, K. (2014). “On the Very Idea of Direction of Fit”. In: *Philosophical Review* 123.4, pp. 429–484.
- Geach, Peter (1957). *Mental Acts*. Routledge and Kegan Paul.
- Gilbert, Daniel T, Romina W Tatarodi, and Patrick S Malone (1993). “You Can’t Not Believe Everything You Read.” In: *Journal of personality and social psychology* 65.2, p. 221.
- Ginsborg, Hannah (2006). “Reasons for Belief”. In: *Philosophy and Phenomenological Research* 72.2, pp. 286–318.

- Godfrey-Smith, Peter (1994). “A Modern History Theory of Functions”. In: *Noûs* 28.3, pp. 344–362.
- Goldman, Alan H. (2017). “What Desires Are, and Are Not”. In: *Philosophical Studies* 174.2, pp. 333–352.
- Gottfried, J.A. (2011). *Neurobiology of Sensation and Reward*. Frontiers in Neuroscience. CRC Press.
- Graham, Peter J. (2012). “Epistemic Entitlement”. In: *Noûs* 46.3, pp. 449–482.
- (2014). “The Function of Perception”. In: *Virtue Scientia: Bridges between Virtue Epistemology and Philosophy of Science*. Ed. by Abrol Fairweather. Synthese Library, pp. 13–31.
- Greenberg, Mark and Gilbert Harman (2009). “Conceptual Role Semantics”. English. In: *The Oxford Handbook of Philosophy of Language*. Ed. by Ernest Lepore and Barry C. Smith. Oxford: Oxford University Press.
- Gregory, Alex (2013). “The Guise of Reasons”. In: *American Philosophical Quarterly* 50.1, pp. 63–72.
- (2016). “Might Desires Be Beliefs About Normative Reasons for Action?” In: *The Nature of Desire*. Ed. by Julien Deonna and Federico Lauria. Oxford: Oxford University Press.
- (2018). “Why Do Desires Rationalize Actions?” In: *Ergo: An Open Access Journal of Philosophy* 5.
- Grice, H. P. (1974). “Method in Philosophical Psychology (From the Banal to the Bizarre)”. In: *Proceedings and Addresses of the American Philosophical Association* 48.n/a.
- Grzankowski, Alex (2015). “Not All Attitudes Are Propositional”. In: *European Journal of Philosophy* 23.3, pp. 374–391.
- Harris, Alison and Jane Wardle (1987). “The Feeling of Hunger”. In: *British Journal of Clinical Psychology* 26.2, pp. 153–154.
- Hawkins, Jennifer (2008). “Desiring the Bad under the Guise of the Good”. In: *Philosophical Quarterly* 58.231, pp. 244–264.
- Hawthorne, John and Ofra Magidor (2018). “Reflections on the Ideology of Reasons”. In: *The Oxford Handbook of Reasons and Normativity*. Ed. by Daniel Star. Oxford: Oxford University Press.
- Heuer, Ulrike (Oct. 2004). “Reasons for Actions and Desires”. In: *Philosophical Studies* 121.1, pp. 43–63.

- Holton, Richard and Kent Berridge (2016). “Addiction Between Compulsion and Choice”. In: *Addiction and Self-Control*. Ed. by Neil Levy. Oxford: Oxford University Press.
- Hornsby, Jennifer (2008). “A Disjunctivist Conception of Acting for Reasons”. In: *Disjunctivism: Perception, Action, Knowledge*. Ed. by Adrian Haddock and Fiona Macpherson. Oxford: Oxford University Press.
- Huemer, Michael (2006). “Phenomenal Conservatism and the Internalist Intuition”. In: *American Philosophical Quarterly* 43.2, pp. 147–158.
- Humberstone, Lloyd (1992). “Direction of Fit”. In: *Mind* 101. January.
- Hyman, John (1999). “How Knowledge Works”. In: *Philosophical Quarterly* 49.197, pp. 433–451.
- Kauppinen, Antti (forth). “Rationality as the Rule of Reason”. In: *Noûs*. URL: <https://philpapers.org/archive/KAURAT-3.pdf>.
- Kiesewetter, Benjamin (2017). *The Normativity of Rationality*. Oxford: Oxford University Press.
- Kolodny, Niko (2005). “Why Be Rational?” In: *Mind* 114.455, pp. 509–563.
- Korsgaard, Christine M. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Kringelbach, Morten L (2015). “The Pleasure of Food: Underlying Brain Mechanisms of Eating and Other Pleasures”. In: *Flavour* 4.1, pp. 20–20.
- Lauria, Federico (2017). “The Guise of the Ought-to-Be. A Deontic View of the Intentionality of Desire”. In: *The Nature of Desiree*. Ed. by Federico Lauria and Julien Deonna. Oxford: Oxford University Press, pp. 139–64.
- Lavin, Douglas (2011). “Problems of Intellectualism: Raz on Reason and Its Objects”. In: *Jurisprudence* 2, p. 367.
- Lehrer, Keith and Stewart Cohen (May 1983). “Justification, Truth, and Coherence”. In: *Synthese* 55.2, pp. 191–207.
- Lewis, David (1980). “Mad Pain and Martian Pain”. In: *Readings in the Philosophy of Psychology*. Ed. by Ned Block. Vol. 1. Cambridge, Massachusetts: Harvard University Press, pp. 216–222.
- (1994). “Reduction of Mind”. In: *Companion to the Philosophy of Mind*. Ed. by Samuel Guttenplan. Blackwell, pp. 412–431.

- Loar, Brian (1981). *Mind and Meaning*. Cambridge: Cambridge University Press.
- Lord, Errol (Jan. 2014). “The Coherent and the Rational”. In: *Analytic Philosophy* 55.2, pp. 151–175.
- (2018). *The Importance of Being Rational*. Oxford: Oxford University Press.
- Lord, Errol and Barry Maguire, eds. (2016). *Weighing Reasons*. New York: Oxford University Press.
- Lycan, William G (2012). “Desire Considered as a Propositional Attitude”. In: *Philosophical Perspectives* 26.Philosophy of Mind, pp. 201–15.
- MacCormack, Jennifer K. and Kristen A. Lindquist (Mar. 2019). “Feeling Hangry? When Hunger Is Conceptualized as Emotion.” eng. In: *Emotion (Washington, D.C.)* 19.2, pp. 301–319.
- Markie, Peter J. (2006). “Epistemically Appropriate Perceptual Belief”. In: *Nôus* 40.1, pp. 118–142.
- Martin, Michael G. F. (1993). “The Rational Role of Experience”. In: *Proceedings of the Aristotelean Society* 93.1993, pp. 71–88.
- (Sept. 2002). “The Transparency of Experience”. In: *Mind and Language* 17.4, pp. 376–425.
- McDowell, John (1995). “Knowledge and the Internal”. In: *Philosophy and Phenomenological Research* 55.4, pp. 877–893.
- McHugh, Conor (2012). “Reasons and Self-Knowledge”. In: *The Self and Self-Knowledge*. Ed. by Annalisa Coliva. Oxford: Oxford University Press, pp. 139–163.
- Mele, Alfred R. (2003). *Motivation and Agency*. Oxford: Oxford University Press.
- Millikan, Ruth Garrett (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, Massachusetts: MIT Press.
- (1986). “Thoughts without Laws: Cognitive Science with Content”. In: *Philosophical Review* 95.January, pp. 47–80.
- Monello, L. F. and J. Mayer (Mar. 1967). “Hunger and Satiety Sensations in Men, Women, Boys, and Girls.” In: *The American Journal of Clinical Nutrition* 20.3, pp. 253–261.

- Montague, Michelle (2007). “Against Propositionalism”. In: *Nous* 41.3, pp. 503–518.
- Morgenbesser, Sidney and Edna Ullmann-Margalit (1977). “Picking and Choosing”. In: *Social Research* 44.4, pp. 757–785.
- Mulligan, Kevin and Klaus R. Scherer (Sept. 2012). “Toward a Working Definition of Emotion”. In: *Emotion Review* 4.4, pp. 345–357.
- Nagel, Thomas (1970). *The Possibility of Altruism*. Oxford: Oxford Clarendon Press.
- Ninan, Dilip (2010). “De Se Attitudes: Ascription and Communication.” In: *Philosophy Compass* 5.7, pp. 551–567.
- O’Brien, Lucy (2005). “Self-Knowledge, Agency and Force”. In: *Philosophy and Phenomenological Research* 71.3, pp. 580–601.
- Parfit, Derek (2001). “Rationality and Reasons”. In: *Exploring Practical Philosophy: From Action to Values*. Ed. by Dan Egonsson et al. Surrey, UK: Ashgate, pp. 17–39.
- (2011). *On What Matters: Two-Volume Set*. Oxford: Oxford University Press.
- Peacocke, Christopher (1999). *Being Known*. Oxford: Oxford University Press.
- (2002). “Three Principles of Rationalism”. In: *European Journal of Philosophy* 10.3, pp. 375–397.
- (2003). *The Realm of Reason*. Oxford: Oxford University Press.
- Pearce, J.M. (2013). *Animal Learning and Cognition: An Introduction*. East Sussex: Taylor & Francis.
- Pettit, Philip and Michael Smith (1990). “Backgrounding Desire”. In: *Philosophical Review* 99.4, pp. 565–592.
- Piccinini, Gualtiero (2004). “Functionalism, Computationalism, and Mental Contents”. In: *Canadian Journal of Philosophy* 34.3, pp. 375–410.
- Price, Huw (1989). “Defending Desire-as-Belief”. In: *Mind* 98.389, pp. 119–127.
- Putnam, Hilary (1960). “Minds and Machines”. In: *Journal of Symbolic Logic*. Ed. by Sidney Hook. New York University Press, pp. 57–80.
- Quinn, Warren (1994a). *Morality and Action*. Cambridge: Cambridge University Press.

- (1994b). “Putting Rationality in Its Place”. In: *Morality and Action*. Ed. by Philippa Foot. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press, pp. 228–255.
- Railton, Peter (2010). “Rational Desire and Rationality in Desire: An Unapologetic Defense”. URL: <http://tiny.cc/mcuafz> (visited on 04/05/2016).
- (2012). “That Obscure Object, Desire”. In: *Proceedings and Addresses of the American Philosophical Association* 86.2, pp. 22–46. URL: <http://www.jstor.org/stable/43661297>.
- (2017). “Learning as an Inherent Dynamic of Belief and Desire”. English. In: *The Nature of Desire*. New York: Oxford University Press.
- Raz, Joseph (1998). “Incommensurability and Agency”. In: *Engaging Reason*. Ed. by Joseph Raz. Oxford University Press.
- (1999). *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press.
- (2010). “On the Guise of the Good”. In: *Desire, Practical Reason, and the Good*. Ed. by Sergio Tenenbaum. Oxford: Oxford University Press.
- (2011). *From Normativity to Responsibility*. Oxford: Oxford University Press.
- Robbins, Philip (2017). “Modularity of Mind”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2017. Metaphysics Research Lab, Stanford University. URL: <https://plato.stanford.edu/archives/win2017/entries/modularity-mind/>.
- Rummel, Jan and Laura Nied (Oct. 2017). “Do Drives Drive the Train of Thought?—Effects of Hunger and Sexual Arousal on Mind-Wandering Behavior”. In: *Consciousness and Cognition* 55, pp. 179–187.
- Saemi, Amir (2014). “The Guise of the Good and the Problem of Over-Intellectualism”. In: *Journal of Value Inquiry* 48.3, pp. 489–501.
- Scanlon, T. M. (2014). *Being Realistic About Reasons*. Oxford: Oxford University Press.
- Scanlon, Thomas (1998). *What We Owe to Each Other*. Cambridge, Massachusetts: Belknap Press of Harvard University Press.
- Schafer, Karl (2013). “Perception and the Rational Force of Desire”. In: *Journal of Philosophy* 110.5, pp. 258–281.

- Schapiro, Tamar (2009). “The Nature of Inclination”. In: *Ethics* 119. January, pp. 229–256.
- (2014). “What Are Theories of Desire Theories Of?” In: *Analytic Philosophy* 55.2, pp. 131–150.
- Scherer, K.R. et al. (2001). *Appraisal Processes in Emotion: Theory, Methods, Research*. Series in Affective Science. Oxford: Oxford University Press.
- Schiffer, Stephen (1976). “A Paradox of Desire”. In: *American Philosophical Quarterly* 13.3, pp. 195–203.
- Schnee, Ian (2016). “Basic Factive Perceptual Reasons”. In: *Philosophical Studies* 173.4, pp. 1103–1118.
- Schroeder, Mark (2007). *Slaves of the Passions*. Oxford: Oxford University Press.
- (2008a). “Having Reasons”. In: *Philosophical Studies* 139.1.
- (2008b). “How Does the Good Appear To Us?” In: *Social Theory and Practice* 34.1, pp. 119–130.
- Schroeder, Tim (2004). *Three Faces of Desire*. Oxford: Oxford University Press.
- Schueler, G. F. (1995). *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge, Massachusetts: MIT Press.
- Schultz, Wolfram (Mar. 2016). “Dopamine Reward Prediction Error Coding”. In: *Dialogues in clinical neuroscience* 18.1, pp. 23–32. URL: <https://www.ncbi.nlm.nih.gov/pubmed/27069377>.
- Setiya, Kieran (2010). “Sympathy for the Devil”. In: *Desire, Practical Reason, and the Good*. Ed. by Sergio Tenenbaum. Oxford: Oxford University Press, pp. 82–110.
- Siegel, Susanna (2013). “The Epistemic Impact of the Etiology of Experience”. In: *Philosophical Studies* 162.3, pp. 697–722.
- Silins, Nicholas (2012). “Judgment as a Guide to Belief”. In: *Introspection and Consciousness*. Ed. by Declan Smithies and Daniel Stoljar. Oxford University Press.
- Sinhababu, Neil (2015). “Advantages of Propositionalism”. In: *Pacific Philosophical Quarterly* 96.2, pp. 165–180.
- Smith, Michael (1987). “The Humean Theory of Motivation”. In: *Mind* 96.381.
- (1994). *The Moral Problem*. Oxford: Blackwell.

- (Oct. 2012a). “Four Objections to the Standard Story of Action (And For Replies)”. In: *Philosophical Issues* 22.1, pp. 387–401.
- (2012b). “Scanlon on Desire and the Explanation of Action”. In: *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. Ed. by R. Jay Wallace, Rahul Kumar, and Samuel Freeman. Oxford: Oxford University Press.
- Smithies, Declan and Jeremy Weiss (Mar. 2019). “Affective Experience, Desire, and Reasons for Action”. In: *Analytic Philosophy* 60.1, pp. 27–54.
- Sobel, David and David Copp (2001). “Against Direction of Fit Accounts of Belief and Desire”. In: January, pp. 44–53.
- Sober, Elliott (1985). “Panglossian Functionalism and the Philosophy of Mind”. In: *Synthese* 64. August.
- Sosa, Ernest (2001). “Human Knowledge, Animal and Reflective”. In: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 106.3, pp. 193–196.
- (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford: Oxford University Press.
- (2010). “How Competence Matters in Epistemology”. In: *Philosophical Perspectives* 24.1, pp. 465–475.
- (2015). *Judgment and Agency*. Oxford: Oxford University Press.
- (2017). *Epistemology*. Princeton University Press.
- Stalnaker, Robert (1984). *Inquiry*. London: MIT Press.
- Stampe, Dennis W. (1977). “Toward a Causal Theory of Linguistic Representation”. In: *Midwest Studies In Philosophy* 2.1, pp. 42–63.
- (1986). “Defining Desire”. In: *The Ways of Desire*. Ed. by J Marks. New Jersey: Precedent.
- (1987). “The Authority of Desire”. In: *Philosophical review* 96.3, pp. 335–381.
- Stanley, Jason (2010). ““Assertion” and Intentionality”. In: *Philosophical Studies* 151.1, pp. 87–113.
- Stocker, Michael (1979). “Desiring the Bad: An Essay in Moral Psychology”. In: *Journal of Philosophy* 76.12, pp. 738–753.
- Strawson, Galen (1994). *Mental Reality*. Cambridge, Massachusetts: MIT Press.

- Street, Sharon (2008). “Constructivism about Reasons”. In: *Oxford Studies in Metaethics*. Ed. by Russ Shafer-Landau. Vol. 3. Oxford: Oxford University Press, pp. 207–45.
- Sutton, R.S., A.G. Barto, and F. Bach (2018). *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning Series. Cambridge, Massachusetts: MIT Press.
- Sylvan, Kurt (2014). “What Apparent Reasons Appear to Be”. In: *Philosophical Studies* 172.3, pp. 587–606.
- Tanyi, Attila (2011). “Desires as Additional Reasons? The Case of Tie-Breaking”. In: *Philosophical Studies* 152.2, pp. 209–227.
- Tenenbaum, Sergio (2007). *Appearances of the Good: An Essay on the Nature of Practical Reason*. Cambridge University Press.
- (2008). “Appearing Good: A Reply to Schroeder”. In: *Social Theory and Practice* 34.1, pp. 131–138.
- (2013). “Guise of the Good”. In: *The International Encyclopedia of Ethics*. Ed. by Hugh LaFollette. Wiley-Blackwell.
- Teng, Lu (Feb. 2018). “Is Phenomenal Force Sufficient for Immediate Perceptual Justification?” In: *Synthese* 195.2, pp. 637–656.
- Thagard, Paul R. (2006). “Desires Are Not Propositional Attitudes”. In: *Dialogue* 45.1, pp. 151–156.
- van Reekum, Carien M. and Klaus R. Scherer (Jan. 1997). “Levels of Processing in Emotion-Antecedent Appraisal”. In: *Advances in Psychology*. Ed. by Gerald Matthews. Vol. 124. North-Holland, pp. 259–300. URL: <http://www.sciencedirect.com/science/article/pii/S0166411597801239>.
- Velleman, J. David (1992). “The Guise of the Good”. In: *Nous* 26.1, pp. 3–26.
- Wallace, R. Jay (1999). “Addiction as Defect of the Will: Some Philosophical Reflections”. In: *Law and Philosophy* 18.6, pp. 621–654.
- Way, Jonathan (2018). “Reasons and Rationality”. In: *Oxford Handbook of Reasons and Normativity*. Ed. by Daniel Star. Oxford: Oxford University Press, pp. 485–502.
- Wedgwood, Ralph (2017). *The Value of Rationality*. Oxford: Oxford University Press.

- Williams, Bernard (1979). "Internal and External Reasons". In: *Rational Action*. Ed. by Ross Harrison. Cambridge: Cambridge University Press, pp. 101–113.
- Wilson, Timothy D. and Daniel T. Gilbert (June 2005). "Affective Forecasting: Knowing What to Want". In: *Current Directions in Psychological Science* 14.3, pp. 131–134.
- Winkielman, Piotr and Kent C. Berridge (June 2004). "Unconscious Emotion". In: *Current Directions in Psychological Science* 13.3, pp. 120–123.
- Wright, Larry (1973). "Functions". In: *The Philosophical Review* 82.2, pp. 139–168.