An Affective Theory of Desire

A thesis submitted for the degree of Doctor of Philosophy of
The Australian National University

by
Daniel Gideon Friedrich

May 2008
I hereby declare that the work presented in this thesis is, to the best of my knowledge, original, except as acknowledged in the text.

Daniel Gideon Friedrich
2 May 2008
Acknowledgments

They say it takes a village to raise a child. Over the course of writing my thesis I have been helped by so many people in so many different ways that I feel the same might well be said about finishing a dissertation. It is with pleasure that I turn to acknowledge their support.

I have learned an incredible amount from numerous discussions with friends while studying for my PhD. Many have also read substantial parts of the thesis and provided insightful suggestions. I am very thankful to all, especially Tim Bayne, Peter Berggren, Ben Blumson, Michael Brady, Philippe Chuard, John O’Dea, Christoph Fehige, Sven Peetz, Anu Rafi, Sven Rosenkranz, Jorim Schraven, Irvin Studin, Dave Wall, and Ursula Wolf.

Supervisors play a major part in the process of writing a dissertation. I profited much from discussions with Michael Smith before he left for Princeton. Richard Joyce also provided valuable feedback throughout. I am particularly indebted to Martin Davies and Nic Southwood who both went far beyond reasonable expectations in terms of the time they invested in my supervision, the charity with which they approached my work, and their consistent striving to improve my thoughts without imposing their own views. Another special debt is owed to Bob Goodin, who took over as the chair of my supervisory panel in the final year and offered both insightful comments and constructive coaxing to ensure that the thesis was finished within a reasonable time frame.

It is hard to imagine that I would have taken up philosophy if my parents, Tiqvah and Axel, had not provided an environment that encouraged autonomous learning and critical thinking. They have also always backed my decision to study philosophy, thereby making the road much easier to travel. I am very grateful to them.

Last, but by no means least, I want to express my deepest thankfulness for all the love, patience and support that my partner Annabelle has given me over the years. Sharing my life with her has been wonderful and provided a healthy balance to philosophical reflection. Without it, I doubt that I would have kept up the motivation to come back to my dissertation day in and day out.
Desires play a vital role in our lives. They also figure centrally in philosophical thought about practical rationality, reasons for, and the explanation of action. But what exactly are desires? In this thesis I argue that the best philosophical account of desire characterises it as a state individuated by its relation to experiences of felt need. More precisely, I argue that a subject S (non-instrumentally) desires that $p$ if and only if from time to time $S$ has a felt need for $p$ to become reality.

According to a wide-spread view, desire necessarily involves a tendency to act. In chapter one, I argue that the truth behind this thesis depends upon a prior truth about desire, namely, that desire is a pro-attitude. Most contemporary philosophers aim to capture this dimension of desire in terms of evaluative cognition. In chapter two, I argue that no such analysis of the pro-attitudinal character of desire can succeed.

How should we characterise the pro-attitudinal character of desire? In chapter three, I develop the conceptual framework for a new answer. A key element of this framework is the thesis that there are two kinds of evaluation: cognitive evaluation and affective evaluation. To draw that distinction we must, in turn, distinguish between a cognitive and an affective mode of intentionality. I clarify these notions and argue that, contrary to much contemporary thought, modes of intentionality can’t be reduced to functional properties but must be understood in terms of their relation to distinctive conscious episodes.

In chapter four, I apply this apparatus to the case of desire. Desire is a state constituted by an affective mode of intentionality. It is a pro-attitude in as much as it involves affective evaluation. Moreover, I argue that the best account of the affective nature of desire is in terms of episodes of felt need. In particular, I argue that a subject S (non-instrumentally) desires that $p$ if and only if from time to time $S$ has a felt need for $p$ to become reality.

Much of chapter four is devoted to defending this felt need theory of desire against objections. However, discussion of a central objection is deferred to chapter five. This is the objection that any theory that individuates desire in terms of distinctive conscious episodes must misconstrue the role of desire in the explanation of action. This objection is built upon the Humean thesis that desire is necessary for motivation and action. In chapter five, I examine the Humean view in detail and argue that it is false.

Finally, chapter six considers whether desire can provide us with (normative) reasons for action. I argue that it can and that the felt need theory of desire is well-equipped to explain why this should be so.
Contents

Acknowledgments
Abstract
Contents

Introduction ........................................................................................................ 1

Chapter 1: Desires as Causal Engines ............................................................... 4
  1.1 The Causal Engine Thesis ...................................................................... 4
  1.2 A Counterexample to the Causal Engine Thesis ................................. 9
  1.2.1 The Weather Watchers .................................................................. 10
  1.2.2 Non-Actable Desires ..................................................................... 13
  1.3 Causal Engines and the Rational Significance of Desire ..................... 19
  1.4 Rationality and Evaluation .................................................................. 20

Chapter 2: Cognitivism about Desires ............................................................. 24
  2.1 Desire-as-Belief .................................................................................. 25
  2.2 Desire-Implies-Belief ......................................................................... 26
  2.2.1 The Argument from Limited Cognitive Sophistication .................. 26
  2.2.2 The Argument from Recalcitrant Desires ....................................... 29
  2.3 Non-Doxastic Cognitive Theories of Desire ......................................... 33
  2.4 Desire and the Norm of Truth ............................................................ 34

Chapter 3: Modes of Intentionality ................................................................. 39
  3.1 Modes of Intentionality .................................................................... 40
  3.2 Affective Intentionality .................................................................... 43
  3.3 Modes of Intentionality and Functionalism ....................................... 48
  3.4 Modes of Intentionality and Consciousness ...................................... 52
  3.5 Cognitive and Affective Evaluation .................................................... 58

Chapter 4: An Affective Theory of Desire ....................................................... 60
  4.1 Hedonic Theories of Desire ................................................................. 60
  4.2 Desire and Felt Need .......................................................................... 65
  4.3 Benefits of the Felt Need Theory of Desire ........................................ 72
  4.4 Objections to the Felt Need Theory of Desire .................................... 74
    4.4.1 Is There a Phenomenology of Desire? ........................................ 74
    4.4.2 The Argument from Content ..................................................... 76
    4.4.3 The Argument from Epistemology ............................................ 77
    4.4.4 The Argument from Motivation ............................................... 79
There were some moments, after long periods of suffering, when what Ivan Ilyich wanted more than anything else – however embarrassed he would have been to admit it – what he wanted was for someone to take pity on him as if he were a sick child. He wanted to be kissed and cuddled and have a few tears shed over him in the way that children are cuddled and comforted. He knew he was a big man and something of a greybeard, which made this impossible, but nevertheless this is what he wanted.

(Tolstoy, *The Death of Ivan Illyich*)
Introduction

To have no desire”, Hobbes once said, “is to be dead” (Hobbes [1994]: 1.8.16). His point was not, of course, that a lack of desire would literally put our survival at stake but rather that without desire our lives would lose much of their vitality and the world, much of its splendour. Given the enormous importance that desires play in our lives it should, therefore, come as no surprise to learn that desires also figure strongly in the philosophical endeavour to shed light on our condition, especially in the philosophy of mind and in moral philosophy. But what exactly are desires? The aim of this thesis is to provide an answer to this question – in short, to give a philosophically illuminating account of the nature of desire.

A project of this kind raises important methodological questions concerning the scope of the inquiry and the criteria of its success. With respect to the former question, we must note that people frequently associate desire with our animal natures. Thus, desire is often understood to be a stand-in for sexual desire and, to a lesser extent, desire for food and drink or excretion. But people also desire many other things: to be held tight, to win a competition, to be promoted, to have children, to paint a portrait, and so on and so forth. Note also that some desires are very strong, while others are weak. Some are short-lived; others last for years or even decades. Put another way, some of our desires are full-blown yearnings or longings, whereas others are mere niggling or fleeting wants. The aim of this thesis is to shed light upon our desires no matter what object they take, or strength or duration they have.

However, giving an account of the nature of desire-like states such as wishing, hoping, or regretting falls outside the scope of this thesis. This is because although it is natural to think that desires for different things, or with differing strength or duration have a unified nature, it is much less clear whether the same is true with respect to desires and desire-like states. Clearly, desire-like states are importantly related to desire, and the results of this thesis may in an indirect way shed light upon these states as well. Moreover, it will be necessary at times to touch upon desire-like states. Nonetheless, the aim of this thesis is to give an account of the nature of desire; not to give an account of desire and desire-like states.

Moving on to the second methodological question, what are the criteria of success for a theory of desire? One thing a successful theory of desire must do is identify in virtue of what we judge that an agent has or lacks a desire in particular circumstances. Of course, our intuitions are neither completely precise or conflict free and we should, therefore, be prepared to revise some of our pre-reflective judgments about particular
cases. Yet, while some tweaking of these intuitions is to be expected and acceptable, a theory of desire that failed to capture our most central intuitions about particular cases would only notionally be a theory of desire.

This isn't all a successful theory of desire must do. Our pre-reflective conception of desire also involves various views about the kind of state desire is. For example, we intuitively think that desire is a state that places relatively few demands upon the cognitive sophistication of agents, that we are in a privileged epistemic position with respect to our own desires but also that judgments about our own desires aren’t infallible, that desire is a state that can be unconscious but also that there are conscious episodes that reveal desire, that desire is subject to certain kinds of criticisms but not to others, that desire is a state with which the world must 'fit' not vice versa, and so forth. A successful theory of desire must strive to respect these intuitions.

But a successful theory of desire must do even more than identify in virtue of what we judge that an agent has or lacks a desire in particular circumstances and respect our intuitions about the kind of state desire is. Desire, it seems, plays an important role in practical rationality. The rational agent, we think, looks after satisfying her desires as well as she can. Desire also seems to be vital for our (normative) reasons for action. If a person desires to spend more time gardening, then we think this provides her with a reason to do so. Finally, desire also seems to be crucial to the explanation of action. We often explain someone's actions by reference to their desires, and according to a venerable tradition in philosophy any fully explicit explanation of action must appeal to desire. Desire, then, appears to play important roles in rationality, (normative) reasons for, and the explanation of, action. If this is so, then a successful theory must also help us understand how desire can play these theoretical roles.

This means that we can’t divorce the task of assessing a theory of desire from getting clearer about what exact role desire actually plays in practical rationality, (normative) reasons for, and the explanation of action. Unfortunately, claims about the role of desire in these fields are highly contentious – with claims about the role of desire vis-à-vis our (normative) reasons for action and the role of desire in the explanation of action being especially controversial. Nonetheless, we must take up the task of clarifying the role of desire in practical rationality, (normative) reasons for, and the explanation of action if we are serious about giving the best philosophical theory of desire, and I shall, therefore, devote substantial parts of this thesis to addressing these matters.

\[1\] For stylistic reasons, I shall frequently use feminine pronouns to signify the generic, gender-neutral 'she or he,' 'her or him,' or 'hers or his.' I shall also assume that 'agent,' 'person,' 'someone' and so forth are terms that can refer to persons of either gender. However, I shall assume that 'man' is not gender neutral and refers exclusively to a male person.
The three criteria of success against which I shall measure a theory of desire are, thus, (i) its ability to unify our intuitions about specific cases of desire, (ii) its capacity to respect our intuitions about the kind of state desire is, and (iii) its power to help us understand how desire can play the role it does in practical rationality, (normative) reasons for, and the explanation of action. To be sure, we should not assume in advance that there exists any theory of desire that satisfies all three criteria. Reflection might lead us to make deep revisions to our pre-reflective intuitions about desire, or it might lead us to disentangle different notions that are conflated in our ordinary thought and talk about ‘desire’. However, while we can’t rule out that our inquiry will force us towards such conclusions, we should not take them as our starting point.  

Very briefly, the plan for this thesis is as follows: In the first chapter, I shall begin by scrutinising the appealing idea that desire necessarily involves a tendency to act. The main conclusion will be that desires that involve action-directed tendencies do so in virtue of the rational connection between the desire for the end and the motivation to act so as to promote it, and that such a rational connection presupposes that desire is a pro-attitude. But how can desire be a pro-attitude? In chapter two, I shall discuss the influential view that desire is a pro-attitude in as much as it implies a cognition of the desired end as good. I shall argue that this view fails. In chapter three, I shall put the conceptual framework in place to make way for a more satisfactory account of desire as a pro-attitude. In chapter four, I shall then apply this framework to desire, arguing for what I call the felt need theory of desire. This theory, I argue, can explain the pro-attitudinal character of desire and, consequently, also elucidate the role of desire in rationality. In the fifth chapter, I consider the role of desire in the explanation of action and argue that the felt need theory can do justice to it. In the sixth and final chapter, I discuss the role of desire with respect to our (normative) reasons for action and argue that, once again, the felt need theory can explain the role of desire.

---

2 If there really are distinct kinds of desire, I would claim that the one that I try to explicate in this thesis is particularly worthy of our attention, because it is connected with what we find most important about desire: its role in rationality, reasons for, and the explanation of action.
Chapter 1
Desires as Causal Engines

There is no shortage of things that we can come to desire. As Anthony Kenny once remarked, we can desire “haircuts, fine weather, health, revolutions, laws and repeals of laws, more free time, less petty restrictions, room to live in, brighter colours and gayer fashions, and so on until we exhaust Aristotle’s categories and Roget’s thesaurus” (Kenny [1963]: 118). But what holds this multitude together? In virtue of what is a desire for a haircut a desire and thus, a state of the same kind as a desire for a revolution? A powerful thought is that any answer to this question must begin with the potential of desire to move us towards action, with desire being an engine for action. In this chapter, we shall scrutinise this popular and commanding idea. In the first section, I shall explore the thought that desires are causal engines for action in much greater detail. In the second section, I shall argue that not all desires are causal engines. And in the third and final section, I shall argue that there are nonetheless some important lessons to be learned from the idea that desires are engines for action.

1.1 The Causal Engine Thesis
Remember the last time you really wanted something? Perhaps you had worked way too hard during the past year and what you wanted, really wanted, was a holiday. A break from work, just a few days relaxing on the beach, doing nothing. Or perhaps a relationship didn’t work out the way you wanted, and you had a strong desire to mend it. Or maybe it was something altogether different. Be that as it may, the chances are that your desire motivated you to take some action. Naturally, there is nothing inevitable about this. Even strong desires sometimes wane without us having had the opportunity to satisfy them. But if your desire did not move you all the way to action, it would still have involved a tendency to act. Indeed, a tendency to act seems to be a central part of what it is to desire. To be sure, this action-directed tendency will be more obvious in some cases. Strong desires, presumably, impel us more strongly towards action than weak desires. But who could doubt that even the weakest desires include a minimal tendency towards action? Indeed, upon reflection the claim that an action-directed tendency is a necessary condition of desire looks exceedingly plausible.

---

3 As Kenny himself notes, strictly speaking, objects of desire must have a propositional structure (Kenny [1963] Ch. 5; De Sousa [1974]). Thus, strictly speaking, we can’t desire ‘health’ but only that we are healthy, not ‘revolutions’ but only that more revolutions occur, and so forth.
Call this the causal engine thesis. The causal engine thesis has been almost universally accepted. Indeed, some have wanted to go even further: they have claimed that an action-directed tendency is also sufficient for desire. Here, though, we shall focus only on the weaker claim that an action-directed tendency is necessary for desire.

The causal engine thesis is intended to capture the thought that desire, by its very nature, inclines agents to act so as to promote the desired end. Of course, this is not to say that desire, by its very nature, mysteriously directs agents towards such action that will in fact bring about the desired end. Rather, it is to say that desire inclines the agent to act in ways that the agent believes will promote the desired end. As such, the causal engine thesis is logically stronger than the claim that a tendency to act simpliciter is necessary for desire. The latter claim would, for example, be vindicated if it turned out to be a brute metaphysical truth that desire necessarily involved a tendency to lift one's arm. Needless to say, such a discovery, however fascinating, would be of the wrong kind to corroborate the causal engine thesis. What needs to be the case for the causal engine thesis to be true is that a specific kind of tendency to act is necessary for desire, namely, a tendency to act in ways that the agent believes will promote the desired end.

There are two main interpretations of the causal engine thesis. On the motivation interpretation, to have an action-directed tendency is to have some motivation to perform an action. Consequently, on this interpretation, desire implies some motivation to act in ways that the agent believes will promote the desired end. On the disposition to be motivated interpretation, desire only implies a disposition to be so motivated. As we shall see, this interpretation is to be preferred as it avoids some difficulties that plague the motivation interpretation.

A number of philosophers have endorsed the motivation interpretation as they think that desiring just is being motivated. Scott Meikle, for example, has claimed that the “distinction between being motivated and having a desire ... is a distinction without a difference” (Meikle [1974]: 55). Likewise, Jonathan Dancy has said that he is “left thinking that what is common to all desires (desirings) is that they are states of being motivated” (Dancy [2000]: 88), indeed, that he sees no persuasive obstacles to a flat-out “identification of desiring with being motivated” (Ibid.).

This doesn’t seem right. One way to see this is to consider the respective ‘success’ conditions for motivation and desire. Motivation can be successful only if it issues in action. Desire, in contrast, is successful if the desired state of affairs comes to obtain.

---

4 I shall assume a causal interpretation of the role of desire in motivation and action. Some philosophers are skeptical this. (e.g. Schon [1994]). However, I do not think that any of the key arguments essentially depend upon this assumption.


6 This has been influentially argued for by Armstrong [1968]; Smith [1994]; Stalnaker [1987].
irrespective of whether or not this is the result of the agent's own doings. To illustrate, suppose I desire to be rich and am motivated to bring it about that I am rich. Can we identify the two states? No, because whereas my desire is satisfied by a generous benefactor who transfers a large sum of money into my bank account, my motivation is rendered pointless in these circumstances. It hasn't been satisfied or carried towards success by the enormously generous action of my benefactor.

However, none of this rules out that being motivated is necessary for desire. A second argument shows that this can't be quite correct either. Remember, desire is said to involve a tendency towards such actions that the agent believes will promote the desired end. If the action-directed tendency consisted in being motivated, then having a desire would have to involve being motivated to act in ways that the agent believes will promote the desired state of affairs. Since one can't be so motivated if one doesn't believe there is something one can do to further the desired state of affairs, it would follow that desire would have to involve some such belief. Yet, this doesn't seem to be the case. You can desire something even if you think it is outside your agential reach. A fan might strongly want her team to win. An old lady might desire to die peacefully in her bed. Both believe, correctly or incorrectly, there is nothing that they can do to make these things happen. This does not mean that they think the desired state of affairs is impossible or even that it is enormously unlikely, which, arguably, would make it impossible for them to desire the state of affairs. Yet, if they think the desired state may well come to obtain, if only through the actions of other people or through the course of nature, they can desire it and also believe that there is nothing they themselves can do to promote it.

It might be said that one must at least be motivated to think about how one could advance the desired state of affairs. Such deliberation can be of two kinds, though. It can be based upon the belief that it will, if only in an indirect way, promote the desired end. Or it can occur in the absence of such belief. In the latter case, deliberation is of the wrong kind to support the motivation interpretation of the causal engine thesis. Even if it were a brute metaphysical truth that desire necessarily induced such deliberation, it would fall short of demonstrating what is at issue – namely, that a motivation to act in ways that the agent believes will promote the desired end is necessary for desire. Focusing on the former kind of deliberation, on the other hand, avoids these difficulties. Unfortunately, it doesn't help, since an agent can desire something without being remotely motivated to think about how to advance it precisely because it isn't in her power to advance it, even through a process that begins with deliberation. It might be said that one must always believe that deliberation carries some significant chance of furthering the desired end, but I don't see why this should be so. Suppose, for example, that an astronomer desires that there will be intelligent life in a distant

---

7 In some cases, it will be logically impossible that the desired state of affairs comes about in any other way, as when one desires that one is performing some action. Still, this does not affect the present point that a desire is successful just in case the desired state comes to obtain.
universe even after all life on earth has vanished. Clearly, this astronomer can desire this whilst also knowing that she can’t do anything at all to promote the desired end, not even through a process that begins with deliberation. In that case, all deliberation would be pointless and a waste of time, and the kind of deliberation required to support the motivation interpretation of the causal engine thesis, namely, deliberation based upon the belief that it will further the desired state, would be impossible.8

While these considerations refute the claim that desire implies being motivated to act in ways that are believed to promote the desired state, they don’t undermine the thesis that desire entails a disposition to be so motivated. This is because they don’t undermine the thesis that the desiring agent would have to be motivated if certain conditions were to obtain. What are those conditions? First, the agent needs to believe there is an action within her agential register that would promote the desired state; she needs an appropriate bridging-belief. Second, the agent needs to be fully rational.

The relevance of the first condition should be obvious, but what purpose does the second serve? Before answering that question, let me clarify what I mean by saying that an agent is fully rational. An agent is fully rational just in case she complies with all requirements of rationality. Importantly, rational requirements, as John Broome has put it, “supervene upon the mind” (Broome [2007]: 352). Therefore, mental duplicates will be subject to exactly the same rational requirements no matter how different their non-mental properties are. Clearly, not all normative requirements supervene upon the mind. What moral or prudential requirements a person is subject to, for example, depends in part upon non-mental facts. But rational requirements are different. They are determined entirely by properties of the person’s mind.9

Now it might be objected that we sometimes say things such as “the rational thing for her to do is to look for a new job” or “to stay put in her job given the recent change of management is just plain irrational”. The truth of these statements depends in part upon non-mental facts. Does this show that some rational requirements do not supervene upon the mind? It does not. Instead, what it shows is that we sometimes use the language of rationality to make claims about, say, prudential requirements. Put another way, in everyday speech, the language of rationality is ambiguous. It can be used to express claims about rational requirements as well as about other normative requirements. Throughout this thesis I shall use the language of rationality only for the purpose of referring to rationality and rational requirements (in contrast to what a person ought to do or has reason to do).

8 Moreover, mutatis mutandis the arguments of section two also apply to the motivation version of the causal engine thesis.
9 For similar conceptions of rationality, see Broome [2007]; Kolodny [2005]; Scanlon [2007].
Rationality requires different things of us. For example, rationality requires that one believe that \( q \) if one believes that \( p \) and believes that if \( p \) then \( q \).\(^{10}\) Rationality also requires that one intend to \( \psi \) if one intends to \( \varphi \) and believes that one must \( \psi \) if one is to \( \varphi \). However, what is distinctive of rational requirements is that they govern the *internal coherence* among our mental states.\(^{11}\) To comply with rational requirements is, of necessity, to exhibit internal coherence. A lack of internal coherence, of necessity, makes one subject to rational criticism. (This does not entail, however, that rational requirements are merely requirements for internal coherence, that all rationality requirements have ‘wide-scope’).\(^{12}\)

With these clarifications in place, let us return to the question of what purpose the requirement of full rationality serves. Suppose Brian wants to have a drink and believes he could do so by getting a beer out of the fridge. Despite this, Brian might not be motivated to go to the fridge because he can fail to rationally combine his desire and belief. Consider an analogy. Suppose Brian believes that \( p \) and believes that if \( p \) then \( q \). Does it follow that he will believe \( q \)? Clearly not. For while \( q \) must be true if \( p \) and if \( p \) then \( q \) are true, Brian may simply fail to put two and two together. Sometimes such mistakes are due to inattention, but this need not be so. Brian may attend to both of his beliefs yet nonetheless fail to draw the required conclusion. As most of us have plenty of experience to attest, a solution can stare us in the face without us recognising that it is there. Moreover, as Lewis Carroll forcefully argued, the failure to draw a conclusion is not just a matter of a missing piece of knowledge – such

---

10  This way of stating what rationality requires is somewhat cumbersome. I have chosen it because it retains impartiality between a narrow- and a wide-scope interpretation of rational requirements. See also footnote 12 below.

11  Many accept this (e.g. Broome [1999]; Kolodny [2005]; Scanlon [1998]). However, there are two possible challenges to this. First, it might be argued that some requirements of rationality govern the internal coherence among mental states and *action*. I believe this is a mistake. True, we sometimes say that it is rationally required of a person that she perform some action if she has certain mental states. But I believe what we really mean in these cases is that a person is rationally required to intend to act (or perhaps, to try to act) if she has certain mental states. One argument in favor of this view is that ought implies can. But we do not retract rational requirements if we discover that the agent can’t act, that is, if she lacks the ability to perform the relevant action. Second, it might also be argued that the rationality of belief and desire depends on certain non-mental facts. Derek Parfit, for example, argues that “the rationality of most of our beliefs depends on whether, in having these beliefs, we are responding to apparent reasons for having them. We have such apparent reasons if the evidence available to us makes it sufficiently likely that these beliefs are true” (Parfit [2001]: 29). This may be understood as claiming that the rationality of belief depends upon evidence for which, as we might put it, there already has been some cognitive uptake. Alternatively, the claim might be that the rationality of belief depends upon evidence that is available to the agent but that may be wholly external to the agent’s mind. In the first case, there is no objection. And in the second case, I think we should deny that the rationality of belief depends upon evidence so understood. An agent who doesn’t appropriately respond to evidence so understood fails to comply with her epistemic requirements, but her beliefs can’t be faulted as irrational. Parfit also claims that “the rationality of our desires depends … on whether, in having these desires, we are responding to apparent reasons for having these desires. We have such apparent reasons if we have beliefs whose truth would make what we want worth achieving” (Ibid.). Whether this is correct or not, it is consistent with the claim that all rational requirements govern the internal coherence among our mental states.

12  I shall remain neutral here as to whether all requirements have wide-scope (see Broome [1999]; Kolodny [2005]).
as not knowing that if $p$ is true and that if it is true that if $p$ then $q$, then $q$ must be true. Add as many premises as you like, the possibility always remains that one fails to draw the implied conclusion.\textsuperscript{13} What Brian needs is not more attention or knowledge, but more rationality. The same applies in the case of desire. In the absence of sufficient rationality, an agent can always fail to rationally combine her desire and bridging-belief to come to be motivated. Per contra, an ideally rational agent who desires that $p$ and believes that she can promote that by $\varphi$-ing must be motivated to $\varphi$ – any failure to be motivated would \textit{ipso facto} be a failure of rationality.\textsuperscript{14} Thus, only a set of conditions that requires full rationality will be sufficiently strong for a credible version of the thesis that desire implies being motivated under certain conditions.

The best interpretation of the causal engine thesis, therefore, amounts to the claim that desire implies a disposition to be motivated to act in as much as an agent who desires that $p$ would be motivated to $\varphi$, were she to believe that $\varphi$-ing would promote that $p$ and were the agent to be fully rational. So understood, the causal engine thesis has a great deal of intuitive force. Nonetheless, in the next section I shall argue that desire does not imply this disposition.

\textbf{1.2 A Counterexample to the Causal Engine Thesis}

According to the causal engine thesis it is impossible to have a desire and not have the relevant disposition to be motivated to act. In this section, I shall argue that this is not impossible. There are possible cases of desire – indeed actual cases of desire – that are unaccompanied by such a disposition. Consequently, the causal engine thesis is mistaken. I shall approach the task in two subsections. I shall begin by considering a well-known attempt by Galen Strawson to develop a counterexample to the causal engine thesis (section II.A). I shall argue that, despite its considerable ingenuity, Strawson’s case is ultimately less convincing than might first be thought. I shall then try to develop a more persuasive kind of counterexample (section II.B).

\textsuperscript{13} Carroll [1895].

\textsuperscript{14} Perhaps this is not quite right. Suppose that Brian also believes that, all things considered, he ought not to go and get a beer out of the fridge. In that case it seems clear that Brian could be fully rational yet lack the intention to get a beer out of the fridge. Indeed, it might be said that Brian intending to get a beer out of the fridge would, under these circumstances, constitute a breach of rationality. But then it might be argued that lacking any motivation to get a beer out of the fridge can’t constitute a failure of rationality. For it might be said that it is odd to suppose it is ever a condition on full rationality that one is both motivated to perform an act and lack the intention to do so. Even if we should accept this – and I’m not sure we should – it doesn’t impact upon the argument in any substantive way. For we could deal with this and similar complications by restricting the claim to cases in which the agent has a desire and an appropriate bridging-belief and does not have any additional mental states that would conflict with the rationality of coming to be motivated to perform the action the agent believes will promote the desired end.
1.2.1 The Weather Watchers

In his book *Mental Reality*, Strawson presents a putative counterexample to the thesis that desiring necessarily involves being disposed (to be motivated) to act. This is the fictional example of the *Weather Watchers*. The Weather Watchers are in many respects just like us. They have sensations, thoughts, emotions, beliefs, desires. But there is one fundamental difference: the Weather Watchers are constitutionally unable to act. They are just not the kind of creatures that can do anything. Their lives are completely passive. They watch and observe (particularly the weather). They have feeling and beliefs, desires and hopes. But all of this occurs in a passive mode. The possibility of the Weather Watchers, Strawson claims, should convince us that the causal engine thesis is false. For while they do not actually exist, they could exist. Yet, as they are ex hypothesi constitutionally unable to act, desirous Weather Watchers would not even be disposed to act or disposed to be motivated to act.

The example is cute, but it isn't obvious how it's supposed to support the conclusion. We can't change the weather. Still, it is plausible that if we desire a rainy day, we thereby have a disposition to be motivated to act in ways we believe to promote rainy days. To be sure, most of us know we can't change the weather, and our desire will, therefore, not motivate us to do anything. Nonetheless, having that desire does seem to dispose us towards action: if we had the relevant bridging-belief and were fully rational, then our desire would motivate us to take action. Why should matters be any different in the case of the Weather Watchers?

One suggestion would be that since the Weather Watchers can't act, it would be impossible for them to form an appropriate bridging-belief. It is not clear, however, that this follows. Even though we can't change the weather, we might still falsely believe we can do so (think of rain-dancers). Why should the Weather Watchers be any more immune from the possibility of false belief than we are? Another suggestion would be that just as a super-vase that can't break isn't even disposed to break, so a Weather Watcher that can't act isn't even disposed to act or disposed to be motivated to act. However, once again we must query this. We can't act to change the weather, yet this does not undercut us being disposed to be motivated to act. True, we can perform other actions, whereas the Weather Watchers can't perform any. Thus, it might be said, we can act in ways we falsely believe to change the weather, but the...

---

15 Strawson [1994] ch.9. Strawson's presentation is complex, and my presentation of his argument involves substantial reconstructions. I hope no injustice is done to Strawson's thoughts. I should also point out that Strawson's general target in that chapter is the claim that "it is impossible for there to be a being that has mental properties but is not even disposed to behave in any way" (Strawson [1994]: 251), though his particular focus is upon desire. Finally, I should note that as I read him, Strawson has at least two major arguments against the causal engine thesis. First, there is an argument that tries to show that it is possible to have desires without being motivated to act or even disposed to be motivated to act. This is the argument considered in the text. Second, there is the argument that tries to show that, even if desire were to include an action-oriented disposition, this would not be of the essence of desire. The latter argument can be naturally seen as akin in spirit to the argument I shall develop in the next section.
Weather Watchers can't even act on the basis of false belief. However, suppose that unbeknownst to us we have been paralysed. In that case, we can't change the weather nor perform a bodily action that we falsely believe will change the weather. Desiring a rainy day, we could nonetheless be motivated to do a rain dance if we believed that this would make it rain. Again, why should matters be different in the case of the Weather Watchers?

Strawson’s defence of the claim that the Weather Watcher could not even be disposed to be motivated boils down to this: something is disposed in a certain way, Strawson claims, just in case it would manifest a certain response under certain conditions. More formally speaking, the claim is that an object O is disposed to R if and only if were C to obtain then O would R. But Strawson claims that the example of the Weather Watchers shows that having a desire does not entail the truth of some such conditional – or more precisely, he claims that the example shows that desire does not entail the truth of a conditional of the right kind.\(^\text{16}\)

There is a worry that this argument gets off on the wrong foot, since it is highly contentious whether dispositions can be analysed in terms of conditionals. Consider the simple conditional analysis that states that an object O is disposed to R in conditions C I and only if it would R were it the case that C. This analysis fails as it can be true that O would R if C obtained without it being true in virtue of the way O is constituted. Examples include: The conditional being trivially true because C is impossible; the response being the result of external interference that obtains in C (cases of ‘mimicking’); and the object acquiring a new disposition when C obtains (‘finkish’ dispositions). The flipside is that the conditional can be false, but not in virtue of the way O is constituted. Examples include: The manifestation of the response being prevented by external interference in C (cases of ‘masking’ and of ‘antidotes’); and the object losing the disposition when C obtains (‘finkish’ dispositions again).\(^\text{17}\) However, I think none of this fatally impacts upon Strawson’s argument. Indeed, it is precisely these sorts of complications that Strawson wants to bypass by flagging the need for conditionals of the right kind. More fully elaborated, his idea is that a true conditional is necessary and sufficient for having a disposition if these complications are excluded.

In other words, the claim is that an object O is disposed to R if and only if it would R were C to obtain, C is possible, and the disposition is not ‘finkish’, mimicked, masked or subject to an antidote. (Building these requirements into the bi-conditional may make the bi-conditional unusable as an analysis of dispositions, but this is irrelevant to Strawson’s argument. His aim is not to give an analysis of dispositions but to establish that the Weather Watchers lack a certain disposition.)

\(^{16}\) Strawson [1994]: 269-276.

\(^{17}\) For the standard semantics of counterpossibles (counterfactuals with necessarily false antecedents) see Lewis [1973]. For ‘finkish’ dispositions see Martin [1994]. For ‘mimicking’ and ‘masking’ of dispositions, see Johnston [1992]. For ‘antidotes’ see Bird [1998]. For an influential attempt to circumvent these objects to conditional analyses of dispositions, see Lewis [1997]. See also Choi [2006]; Gundersen [2002] and Malzkorn [2000] for arguments that the criticisms to the simple conditional analysis are misplaced.
Now let us turn to the Weather Watcher who desires that \( p \). If Strawson is right, this Weather Watcher is disposed to act (or, \textit{mutatis mutandis}, disposed to have a motivation to act) if and only if she would act if certain conditions \( C \) were to obtain, and the truth of this is not due to some extraneous factor. But Strawson thinks there is no true counterfactual of this kind. Why? Here is one possible argument: the Weather Watchers are unable to act. For them to act would require that they first acquire the ability to act. But this is metaphysically impossible because, \textit{ex hypothesi}, the Weather Watchers are constitutionally unable to act. It follows that the required condition \( C \) is metaphysically impossible. But then it would only be trivially true that a Weather Watcher would perform some action were \( C \) to obtain – guaranteed by the semantics of counterpossibles and not by anything in the nature of the Weather Watcher. Here is a second argument: though not strictly impossible, the only conditions under which the Weather Watcher would act are ‘finkish’, that is, conditions wherein the Weather Watcher would \textit{come to acquire} the disposition to act.

Both arguments are suggested by things that Strawson says.\(^{18}\) Is either convincing? The trouble with the first argument is that it deploys a very strong interpretation of the idea that the Weather Watchers are constitutionally unable to act. A weak interpretation is that the Weather Watcher would only be able to act under circumstances that are \textit{remote}. The strong interpretation goes further than this: it holds that it would be strictly \textit{metaphysically impossible} for them to act. But once this is made explicit, we lose our grip on the example. It is exceedingly hard to say what creatures that are metaphysically unable to act would be like, and it would be even harder to say whether such creatures could have desires. Arguing on such contentious grounds against a well-established thesis is unlikely to be successful and will invite the response that one person’s \textit{modus ponens} is another’s \textit{modus tollens}.

The second argument is even less persuasive. Objects with ‘finkish’ dispositions are objects that would \textit{acquire} (or lose) dispositions in conditions under which an object with the disposition would give the relevant response. To illustrate, consider a vase

---

18 For the first argument, see his emphasis that the Weather Watchers are \textit{constitutionally} unable to act and the use of examples where an object can only come to have a disposition once its \textit{nature} has been changed (e.g. the lump of \textit{plastic} that would have to be changed into a lump of \textit{gold} before being disposed to conduct electricity (Strawson [1994]: 275)), so that it is impossible that \textit{it} could have the disposition. For the second argument see his emphasis that the Weather Watchers would not be disposed ‘as they \textit{now} are’ even if they could \textit{become} disposed to act and the examples of objects that would only give certain responses in conditions under which they acquire a new disposition (e.g. the table that is disposed to act because it would act if it acquired beliefs, desires, and the ability to act (Strawson [1994]: 273)). Note also that certain passages suggest a third possible argument, an argument that would seek to disqualify the required \( C \)-antecedent in the case of the Weather Watcher simply on the grounds of it being \textit{remote}. Now there seem to be some dispositions that an object can have only if the relevant \( C \)-antecedents involve relatively nearby situations (e.g. fragility). However, this does not seem to be true of dispositions in general. Some dispositions are only instantiated in extremely remote circumstances. To see this, however, we must take care not to confuse the case in which a disposition is only manifested in extremely remote circumstances with the case in which it is impossible for the disposition to be manifested, the case of ‘finkish’ dispositions, or the case in which a disposition is mimicked.
that is not fragile but that would turn into a fragile vase just in case it is dropped (say, because of a magic spell). This vase would break if dropped, not because it has been fragile all along but because being dropped would transform it into a fragile vase. Next, suppose I want it to rain. This desire disposes me to \( \phi \), and I would \( \phi \) if I were able to \( \phi \), believed that \( \phi \)-ing would promote rain, took it to be possible for me to \( \phi \), had no motivation not to \( \phi \) and were fully rational. The same is true of a Weather Watcher who wants it to rain. She too would \( \phi \) if she were able to \( \phi \), believed that \( \phi \)-ing would promote rain, took it to be possible for her to \( \phi \), had no motivation not to \( \phi \) and were fully rational. Moreover, there is no reason to say that either the Weather Watcher or I would \( \phi \) only because we would acquire a disposition to \( \phi \) under these conditions. Per contra, what makes is true that both the Weather Watcher and I would \( \phi \) under these conditions is simply that we both have a desire that \( \phi \), and that, given this desire, we are already disposed to act under these conditions.\(^{19}\)

### 1.2.2 Non-Actable Desires

According to the causal engine thesis, desire implies a disposition to be motivated to act in ways that are believed to be conducive to achieving the desired end were the agent to have an appropriate bridging-belief and be fully rational. An appropriate bridging-belief is part of the manifestation conditions for this disposition, because without one it isn't possible to be motivated to act in ways that are believed to be conducive to achieving the desired. Being fully rational is also part of the manifestation conditions for this disposition, because a less-than-fully rational agent can fail to be motivated in accordance with her desires and beliefs. Therefore, if desire is to imply a disposition to be motivated, it must be a disposition to do so given the bridging-belief and full rationality.

But what if being motivated on the basis of a desire would be irrational? In that case, the condition of full rationality would work to prevent the agent from coming to be motivated. Consequently, the agent could have the desire, bridging-belief and be fully rational yet not be motivated. Moreover, this lack of motivation would not be due to an extraneous interference, but be based upon the same structural conditions that are typically responsible for ensuring that the agent is motivated. Thus, we would have to conclude that, at least in these cases, desire does not involve a disposition to be motivated. Consequently, desire could not imply such a disposition and the causal engine thesis would be mistaken.

---

\(^{19}\) Contrast Strawson's case of a table that would act in some way if it came to acquire agential abilities, bridging-beliefs, and desires (Strawson [1994]: 273). This table is indeed merely 'finkishly' disposed to act because the conditions under which it would act involve acquiring a disposition to act. And they do so because they involve acquiring a desire. Evidently, the same does not apply in the case of the Weather Watcher who, \textit{ex hypothesi}, already has desires.
Yet, is it ever irrational to be motivated in accordance with one’s desire? One interesting case concerns irrational desires. Which desires, if any, are to count as irrational is controversial. But suppose a desire D is irrational. Suppose further, that the agent believes she could satisfy D by \( \varphi \)-ing. Would it be irrational under these circumstances to be motivated to \( \varphi \)? This is a difficult question, but on balance I think it wouldn’t. This, in any case, is suggested by reflection upon an analogy with theoretical rationality. Suppose a person irrationally believes that \( p \). It still seems that, if the person also believes that \( p \) implies \( q \), there is something rational in the agent coming to believe that \( q \). We are inclined to say that her starting point is flawed, but that the way she moves forward can’t be faulted. Likewise, one could argue that even though D provides a flawed starting point for practical reasoning, the way in which the agent comes to be motivated to \( \varphi \) on that basis is flawless.

However, a second type of desire provides a stronger challenge. Suppose Harry fancies Paula. Suppose, further, that there are a number of things Harry could do that would make Paula fall for him, and that he is well aware of this. But Harry abhors the thought of doing anything in order to make Paula fall for him. Acting with such an end in mind, Harry feels, is contrary to the ideal of romantic love, and in any case, simply not what he wants. Rather, what he wants is for Paula to come to love him without his engaging in any intentional effort designed to make her fall in love with him. In short, Harry has what I call a non-actable desire (or N-desire). What is distinctive about N-desires is their content. They are desires for a state of affairs to obtain without the intervention of the agent. The fully specified content of N-desires is “\( p \) without me doing anything to promote that \( p \)”. With N-desires there is of necessity nothing that the agent can do to promote the desired state of affairs. Any attempt designed to further the obtaining of the desired state of affairs is self-defeating. With N-desires the agent is best off to sit back and relax. 20

N-desires are not irrational. But being motivated to act so as to bring the N-desired state about is. Or so I shall argue. First, however, let me address a complication that arises in this context. Remember, desires are not said to dispose us to be motivated simpliciter but rather to dispose us to be motivated in a way designed to promote the desired end. That is, the aim is to account for our intuition that desire, by its very nature, moves agents towards actions they think will bring about the desired end. But

---

20 Strawson might hint at this kind of argument when he says that

When I want Wimple rather than Ivanov to win the World Chess Championship, I do not wish – let alone necessarily wish – that I could affect the outcome. I want something to happen. I do not wish that I could do anything about it. Desire does not necessarily involve the will.

(Strawson [1994]: 287)

However, he does not elaborate upon this idea, and at another place even seems to indirectly deny it: “Any desire has the following property: it is necessarily true that there are beliefs with which the desire can combine in such a way as to give rise to, or constitute, a disposition to act or behave in some way” (Strawson [1994]: 276). Schroeder [2004]: 17-18 also comes close to giving the argument, though his discussion is very brief and the argument somewhat cryptic.
Desires as Causal Engines

being motivated to act so as to further some aim implies the belief that acting thus promotes the desired end. In the case of N-desires, such motivation would require the belief that one can φ so as to promote that p without promoting that p by φ-ing. This is obviously absurd.

Indeed, it isn't even clear whether the belief is so much as possible. Appendix I provides an argument for thinking that it isn't. If this argument goes through, then N-desires refute the causal engine thesis simply because the conditions under which the said disposition is to be manifested are impossible. In that case, the corresponding counterfactual may be true, but this truth would be too weak to support a disposition ascription, since by the same reasoning it could be shown that N-desires involve every possible disposition, including a disposition not to be motivated.

Even if the required bridging-belief in the case of N-desires is possible, it would obviously be a highly irrational belief. And this, it might be thought, is why I claim that motivation on the basis of the N-desire and the bridging-belief would be irrational. Put another way, I could be understood as claiming that it is simply because the motivation would be based on an irrational state that it can't be rational. So understood, however, the case of N-desires would only give rise to the same kinds of difficulties present in the case of irrational desires, and an argument based on the latter was already found to be inconclusive. Needless to say, this would not quite capture the argument I want to make. I should also, briefly, caution against a second misunderstanding. A fully rational agent, it would be correct to point out, could not even have the required bridging-belief in the case of N-desires. This, it might be said, shows that the antecedent of the relevant conditional in the case of N-desires must be impossible: no-one could have the required bridging-belief and be fully rational. And in that case, a disposition ascription would be unjustified. But again, this is not the basis upon which I object to the causal engine thesis. For it seems that this objection refutes the letter of the thesis but not the spirit. More precisely, it seems only to show that we have been somewhat imprecise in formulating the exact conditions under which the said disposition manifests itself. Instead of requiring full rationality, we ought to have required merely that the agent is fully rational in the way she combines her desires and beliefs.

The real reason why motivation by N-desires is irrational is best illustrated by a comparison with the case of theoretical rationality. Typically, it is rational to believe the believed logical consequences of one's beliefs. But not necessarily. Suppose you believe that p and also believe not p. What rationality requires of you in that case is to resolve that inconsistency and not to believe the conjunction. This is because rationality is an a priori guide to true belief, and it is an a priori truth that p and not p can't both be true. So, rationality can't dictate that you believe that p and not p. Mutatis mutandis for desire. It is typically rational to be motivated to do what one regards as promoting one's desires. But not necessarily. Not if you believe you can satisfy your
N-desire by \( \varphi \)-ing. It isn’t rational for you to be motivated to \( \varphi \) because rationality is an \textit{a priori} guide to the satisfaction of your desires, but it is an \textit{a priori} truth that you can’t possibly satisfy your N-desire by \( \varphi \)-ing. So, rationality can’t dictate that you are motivated to \( \varphi \) if you N-desire that \( p \) and believe you can satisfy your N-desire by \( \varphi \)-ing, any more than rationality can dictate that you believe a contradiction just because you hold contradictory beliefs.

We might sum up the argument as follows: Suppose an agent is \textit{not} fully rational in the way she combines her desires and beliefs in her motivations. Then this agent could desire that \( p \), hold an appropriate bridging-belief and not be motivated. To hold onto the claim that this agent is nonetheless disposed to be motivated in virtue of her desire, we must make the further claim that she would be so motivated if she held the desire and appropriate bridging-belief and were \textit{also} fully rational in her motivations. But while this would go to show that desires other than N-desires imply a disposition to be motivated, it would not show the same about N-desires. For in the case of N-desires it would simply not be rational to be motivated to act on the basis of the desire and bridging-belief. It would not be rational because rationality is an \textit{a priori} guide to the satisfaction of desire, but in the case of N-desires it is \textit{a priori} that every action designed to satisfy the desire would have to fail.

I know of three ways of trying to resist the impending conclusion that the causal engine thesis is mistaken. The first two purport to show that appearances notwithstanding, ‘N-desires’ are not really possible. The third purports to show that we can safely ignore N-desires.

First, it might be argued that cases of ‘N-desires’ are, strictly speaking, cases of \textit{wishing} or \textit{hoping}. They are, therefore, \textit{desire-like} states but not desires. Clearly, this objection will go through only if desire-like states do not imply desires. Let us, however, grant that they don’t. Yet, why should we assume that cases of ‘N-desires’ \textit{must} be cases of wishing or hoping? Desire differs from wishing, for example, in so far as desire aims at the \textit{attainable}, whereas wishing aims at what is taken to be \textit{unattainable}.\footnote{Velleman [2000]: 116-117. Note, Velleman is quite clear that attainable does not mean “attainable by the agent” but rather it “being a possible future outcome” (Ibid.).} Therefore, it might be said, since the outcome is unattainable by the agent in the case of ‘N-desires’, we are really dealing with wishes and not desires. In reply to this objection we should distinguish between a state of affairs being attainable \textit{tout court} and a state of affairs being attainable through the agent’s own actions. It’s plausible that desire presupposes believing the desired end to be attainable \textit{tout court}; that is, believing that it is a live possibility. But desire does not also presuppose that one believe that it is attainable by one’s own deeds. Once that distinction is in place, the force of the objection evaporates because an N-desirer can believe that the desired state of affairs is attainable, even though it is not attainable through her own actions.
It’s even less clear on what basis one could argue that cases of N-desires must be cases of hoping, if only because it’s unclear in virtue of what a state is one of hoping rather than desiring. One possibility is that hoping that \( p \) is set apart from desiring that \( p \) by being a state that involves being less confident about the prospects of \( p \). But this would not do to support the present challenge, because someone who N-desires that \( p \) can be quite confident that \( p \) will come to obtain; it’s just that she also thinks this won’t be so in virtue of her own doing. To illustrate, it is compatible with Harry knowing that he can’t get what he wants by acting so as to make Paula fall in love with him, that he should also be brimming with confidence that she will sooner or later have to recognise what a fantastic catch he is and fall in love with him. Another possibility is that hoping that \( p \) is distinguished from desiring that \( p \) because it involves being anxious about the possibility of not \( p \). But even if this should be true, there is no reason to think people like Harry need be troubled by anxiety. Yet another suggestion is that hoping that \( p \) is characterised by a peculiar experience of one’s own powerlessness. But while the N-desirer knows that the desired state of affairs lies beyond her powers of agency, she need not have the kind of experience of her own powerlessness so characteristic of many instances of hoping, especially if she is as cocky as Harry. Finally, it might be said that hope is desire minus any disposition to act. But surely, the challenger who hopes to win the race is still disposed to run as fast as she can. In short, I can see little to commend the position that apparent cases of N-desires can’t be cases of desire, but only involve wishing or hoping.

Second, it might be argued that I have mischaracterised the nature of N-desires. To illustrate, consider another example. Abe wants his son to win the chess championship on his own. He realises that he can’t further the desired state of affairs by bribing or drugging the opponent. Abe, it might be claimed, has a N-desire. But close attention reveals that there are nonetheless some actions that Abe can take to try and satisfy his N-desire. His desire may motivate him to practise with his son or drive him to the tournament if he believes this will promote his son winning on his own. So, it seems that, contrary to what I claimed, the required belief in the case of N-desires would not have to be incoherent. If this is correct, then my objection to the causal engine thesis falters.

---

22 The nature of hoping is little discussed. Downie [1963] says that hoping that \( p \) is desiring that \( p \) and believing that it is possible for \( p \) to be attained. However, this is not sufficient for hoping that \( p \), as it is possible that one simply desires that \( p \) under these conditions. Schueler [1991]; Searle [1983] and Velleman [2000] note that hoping that \( p \) excludes the belief that \( p \) is true. But again, this does not seem to set hoping apart from desiring (nor do these authors claim to provide an analysis of hope). Another possibility is that hoping is essentially an attitude directed towards the past whereas desire is essentially an attitude directed towards the future. However, this also doesn’t work since one can hope for things in the future. My own view is that hoping involves distinctive negative experiences. However, I doubt that an N-desirer would have to undergo these experiences.

23 Wall [2007].
No doubt there are people like Abe. If the objection is to be successful, however, it must be shown that there can be no people like Harry – people who desire the obtaining of a state of affairs without them taking any action whatsoever. In short, it must be shown that genuine N-desires are impossible and that any appearance of a N-desire only demonstrates a lack of imagination as to what action would, after all, be compatible with the desired state of affairs. I do not see how this claim could be established.

Suppose Bert also wants his son to win the chess tournament on his own. But when Bert says ‘on his own’, he really means it. Bert will shun any action that could improve the chances of his son winning the tournament. He will not practise with his son, he will not drive him to the tournament, and he will not even offer encouraging words. Bert may be a tough guy and a terrible father, but his desire seems possible.  

Third, our pre-theoretical conception of desire, some people will say, does indeed include the possibility of N-desires, but this proves little as our pre-theoretical conception of desire is an unilluminating mess. Philosophical progress is made by imposing some structure, a structure that shows desire to play a unified role in practical reason. This is achieved by assigning desire the role of being an engine for action. If this should violate some of our pre-theoretical intuitions, so be it. This cost is more than made up for by the explanatory power gained. However, this response is flawed on a number of counts. It assumes too quickly that there is no interesting underlying unity between N-desires and other desires. As I shall argue from the next section onwards, this assumption is mistaken. But the response is also tarnished by another flaw. Suppose we assign desire the role of moving us towards action. That is, suppose we say that our conception of desire is that of a state that is a source of motivation. This is in tension with the idea that other states can also be sources of motivation, an idea that, in chapter five, I shall argue we have every reason to think is true. Consequently, the

24 I should note that Smith [1998] has offered another route of defense for the advocate of the causal engine thesis against the troubles of N-desires. He says that a subject desires that Wimple beats Ivanov fairly and squarely only if, in the closest possible world in which she has a desire to gamble, and is offered a choice between a gamble in which the pay-off is that Wimple beats Ivanov fairly and squarely and a gamble in which the pay-off is something that she wants less but assigns only a somewhat higher probability, she chooses the gamble in which the pay-off is that Wimple beats Ivanov fairly and squarely. (Smith [1998]: 450)

However, introducing gambles will not let the causal engine thesis off the hook. Rather, the argument trades on an ambiguity in how to understand the scenario. Let us distinguish two cases. First, imagine a subject who loves to gamble and is being offered two bets. Bet A offers $10 to her if Poldie loses the Horse Race. Bet B offers $5 to her if Wimple beats Ivanov. Suppose furthermore, that the subject judges the likelihood of Poldie losing and of Wimple winning to be the same. In that case, the rational subject will accept bet A. She would do so because that bet has the higher expected pay-off. Note, that this is all that matters. A desire for Wimple to beat Ivanov is completely irrelevant to the subject’s choice of gambles. Second, imagine a subject who is being offered a bet with a rather special pay-off: If she wins the bet, she won’t get money, but her pay-off will be that Wimple beats Ivanov. If the subject desires that to happen, this bet will hold some attraction to her, and she might accept it. In this (rather bizarre) case, her desire for Wimple beating Ivanov will be relevant for which bet she takes, because it affects the expected utility of the bet. However, for an agent with a genuine N-desire for Wimple beating Ivanov, there could be no bet which offered the desired state of affairs as a pay-off. For wanting Wimple to beat Ivanov without any intervention of the agent rules out, amongst other things, Wimple beating Ivanov as a result of the subject having made and won a bet with this pay-off.
attempt to set aside an inconvenient kind of counterexample by appeal to theoretical progress is rather too simplistic to deal with the inherent complexity of the nature of desire and action.

This concludes my argument against the causal engine thesis. Not all desires dispose us towards action. Some desires, at least, have a content that makes it impossible for them to play that role.

1.3 Causal Engines and the Rational Significance of Desire
We began the discussion of the causal engine thesis in the hope of learning something about the nature of desire. Yet, the causal engine thesis is mistaken. Does this mean that there is nothing to be learned from it? I don’t think so. Despite its limitations, there is also an important grain of truth therein.

Not all desires involve a disposition to be motivated to act; N-desires do not. But most desires really do involve a disposition to be motivated to act. One could not have these desires without thereby being disposed to be motivated to act. If one had some such desire and appropriate bridging-belief and were fully rational, then one would be motivated to act in ways that one believes to bring about the desired end.25 Patently, however, this is true only because of a prior truth – namely, that it is a requirement of rationality that one be motivated to act if one desires something and holds the appropriate bridging-belief (and it is a priori possible for the belief to be true). This should give us pause. What kind of a state must desire be for it to be true that desiring something can make one subject to a rational requirement of this kind? A complete answer to this question will have to wait until chapters three and four. But here I shall defend the claim that any answer must revolve around the idea that desire is a pro-attitude.

This is crucial. If desire is a pro-attitude, then we can explain how desiring an end and believing an action to be conducive to that end can make it rational to be motivated to pursue this end. Moreover, understanding desire to be pro-attitude offers an account of a deeper unity between desires that do and desires that do not dispose us towards action. For it allows us to say that all desires are pro-attitudes, although not all pro-attitudes can make motivation rational, because some pro-attitudes are towards states that it is a priori impossible for the agent to promote.

---

25 As noted in footnote 14, this may not be quite correct. We might also have to rule out the presence of additional mental states that would conflict with the rationality of being motivated. Nothing of substance would be affected.
1.4 Rationality and Evaluation

As we shall see in chapter six, the role of desire with respect to (normative) reasons for action is highly controversial. Many people believe that desires can generate full-blown reasons to act, but many others staunchly deny that this is possible. The role of desire in practical rationality is much less controversial. It is widely accepted that there are rational requirements governing desire and bridging-belief on the one hand, and motivation or intention on the other. Moreover, as I have just pointed out, only if there is a rational requirement governing the relation between desire and bridging-belief on the one hand and motivation on the other can we capture the truth behind the causal engine thesis. So, it seems there are good reasons for believing that rationality really does require that one be motivated to act if one has a desire and the appropriate bridging-belief (and it is a priori possible for the belief to be true). This raises the following important question: what kind of a state must desire be such that this can be true and that desire can play this role in practical rationality?

To answer this question, let us consider what I shall call the rationality-evaluation thesis. This is the thesis that it is a requirement of practical rationality that one be motivated to \( \varphi \) if one is in state S only if S involves a (positive) evaluation. Now I shall understand evaluation as, in the widest sense of these terms, casting things in a positive or negative light. On this broad understanding of evaluation, evaluative states include states that involve conceiving of something as good or bad. But it also involves such things as conceiving something as worthwhile or despicable, admirable or reprehensible, obligatory or forbidden. So understood, the rationality-evaluation thesis has a great measure of plausibility. It isn’t rationally required that one be motivated to perform some action if one believes water is H2O. Nor is it a requirement of rationality that one be motivated to perform some action if one sees a tree or imagines horses flying through the air. And intuitively, the reason there are no such requirements of rationality is simply that these mental states do not involve any evaluation.

The rationality-evaluation thesis supports the view that desire involves an evaluation. This is because it is a requirement of rationality that one be motivated to \( \varphi \) if one desires that \( p \) and believes one can promote that \( p \) by \( \varphi \)-ing. Since it is distinctively implausible to suppose that the bridging-belief involves an evaluation, desire must do so. Only if, as we might put it, desire were a pro-attitude could it play the required role in practical rationality.

The claim that desire is a pro-attitude and embodies an evaluative dimension seems quite plausible anyway. Before and after desiring a state, a person may look with calm disinterest upon the possibility of that state coming to obtain. However, while in the grip of the desire, her conception of the end changes. Suddenly, the allure of the state is all too apparent. Suddenly, the object is presented to the agent in a positive light.

---

26 I shall say more about evaluation in chapter three.
This is not to say that the object can’t also be seen to have its flaws. But it appears to be impossible to desire something and to be wholly neutral or wholly negative towards it. Desiring, it seems, requires at least some kind of endorsement.27

Nonetheless, some people reject the claim that desire is a pro-attitude in this sense.28 The reasons for resistance are multifarious, but three central ideas stand out: (i) Thinking of desire as comprising an evaluative endorsement precludes a proper account of cases in which an agent desires what she knows to be bad and, perhaps, even desires it because it is bad; (ii) taking desire to be a pro-attitude forces a cognitive theory of desire upon us, but no such theory can be correct; (iii) the conception of desire and motivation that is being offered is unappealingly hyper-rationalist.

In the next chapter, I shall argue that the first point does indeed put considerable pressure upon a theory of desire that aims to explicate the pro-attitudinal nature of desire in terms of belief. Moreover, I agree that no cognitive theory of desire can be correct and shall also argue this point extensively in the next chapter. However, the task of the third and fourth chapters will be to show that we can hang on to the insight that desire is a pro-attitude without being forced to embrace a cognitive analysis of desire. Thus, I think, the first two objections fail because they mistakenly presuppose that if desire were a pro-attitude then it would have to involve an evaluative belief, or more generally, an evaluative cognition. I shall end this section by arguing that taking desire to be a pro-attitude need not be hyper-rationalist.

One line of thought that can fuel the charge of hyper-rationalism is the conviction that the present account confuses evaluation with motivation. Velleman, for instance, says, “a person who desires something can be said to find it attractive, but this description does not necessarily mean that he makes an attractiveness-judgment about it. It may mean simply that he is attracted to the thing” (Velleman [2000]: 109). In other words, desire is a pro-attitude only in the sense that it moves us towards the desired end. In this sense, and in this sense alone, it can correctly be said that desire essentially involves being attracted towards something. However, this can’t be right. For one, sometimes we desire something without being motivated because we lack an appropriate bridging-belief. Yet, intuitively, having such a desire involves the same kind of pro-attitude as desires that do motivate us. Further, I argued that an essential connection between desire and motivation presupposes a rational connection, which in turn presupposes that desire is a pro-attitude in the evaluative sense. The claim that desire is a pro-attitude because of its connection to motivation, therefore, gets things the wrong way around.

27 This claim has often been made in recent years. See for example, Anscombe [1976]: §37-41; Dancy [2000]: 42; Davidson [1980a]: 3-4, Davidson [1980c]: 87,97,102; De Sousa [1974]: 538; Goldman [1970]: 94; Quinn [1993]: 247; Raz [2001b]: 56; Stampe [1987]: 363. The most vocal critics of this claim are Stocker [1979] and Velleman [2000].
28 See, for example, Stocker [1979] and Velleman [2000].
Another worry that can stand behind the charge of hyper-rationalism is this: some philosophers who think that desire involves a positive appraisal also think that desire must be a response to this appraisal. That is, they think we desire because we antecedently have a positive conception of the desired end. This, I think, really is hyper-rationalist (in chapter six I shall also present an argument against this hyper-rationalist view). But no such commitment is essential to the idea of desire as a pro-attitude. What is essential is simply that desiring involves some kind of positive assessment of the desired end; some way of being attracted to, or seeing the appeal of, what one desires. It doesn’t require the further claims that these evaluations predate our desires and that we desire something because we take a positive view of it.

There is a related issue here about the connection between desire and motivation that can lead to confusion. I don’t think we should rule out the possibility of an agent finding herself motivated to do something without having any desire that could serve to rationalise her motivation. Consider, for example, Warren Quinn’s well-known illustration: a man finds himself set up to turn on radios whenever his eyes come to rest upon a turned-off radio. As Quinn notes, this seems possible even if the agent sees no good whatsoever in turning the radio on – even if, that is to say, the motivation eludes any rational explication and is purely and brutally causal. The man finds himself driven, but does not find himself in the driver’s seat. He is impelled to turn on a radio, but he can’t rationalise this influence from his own point of view. Being in that state would be a weird experience, maybe even frightening. For all that, it does seem possible.29 Quinn says he is not at all sure whether being so set up is to have a desire, but I think it would certainly be very strange to say of a man in that functional state that he desired to turn on the radio (or anything else). In fact, it is hard to see how one could avoid even more absurd consequences if one took this bizarre state to be a desire. Suppose that whenever a man sees a white wall, he is brutally compelled to stand on one leg. Should we say that this man desired to stand on his leg, even if he sincerely denies having any desire and confesses that he finds being in that functional state utterly bizarre? This would clearly stretch our notion of desire too far. But how could we justify thinking that Quinn’s imagined functional state would be a desire but that this man’s state wouldn’t? What would be the principled difference if both states compel the agent in practical ways, and both do so in ways that are brutally causal and rationally inexplicable?30 If we wish to avoid becoming embroiled in this quagmire, we had better make sure not to stretch our notion of desire too far.

29 Quinn [1993]. There may even be actual cases of this sort. See, for example, Della-Sala et al. [1991]; Lhermitte [1986].
30 It might be thought that in Quinn’s imagined case, but not in the other, the contents of the respective states do stand in logical relations. However, as I shall argue further in chapter three, that fact alone is not sufficient to guarantee a rational relation. But in the absence of a rational relation it is hard to see why this difference between the two cases should be of any deeper significance.
While being in a Quinn-type state is not sufficient for desire, it is plausible to think that *some* desires have, or at least could have, their causal origins in such states. That is, it is plausible to think that sometimes being brutally set up in a certain way can cause a desire which, in light of our beliefs, allows us to make rational sense of our practical dispositions. Indeed, from the first person point of view, it might well seem that the desire preceded the motivation, when in fact the motivation was first in the causal order of things. The important point to note is that all of this is entirely consistent with desire being a pro-attitude, and there being a rational connection between desire and motivation.

Let me sum up. I argued that the causal engine thesis is false. Not all desires imply an action-directed tendency; N-desires do not. However, I have also argued that the truth behind the causal engine thesis is that, save for N-desires, desires do indeed involve being disposed to be motivated. Further, I argued that a proper appreciation of this must emphasise the rational connection that exists between desire and belief on the one hand, and motivation on the other. As long as it is *a priori* possible for the agent to act so as to satisfy her desire, it is rational for her to be motivated to do so if she can. Finally, I argued that accounting for the role of desire in practical rationality requires taking note of its pro-attitudinal nature. The question we now need to address is how desire can be a pro-attitude. How can we make sense of the idea that desiring an end involves some kind of positive evaluation? The most obvious answer is to look towards cognitive accounts of desire, accounts that analyse desire as consisting, at least in part, in an *evaluative cognition*. In the next chapter, we shall consider the strengths and weakness of this approach.

---

31 See Nisbett and Wilson [1977] for some suggestive ideas in this context.
Chapter 2

Cognitivism about Desires

Many philosophers who have thought long and hard about the nature of desire have come to the conclusion that desire is, at least in part, a cognitive state.\textsuperscript{32} Specifically, they take desire to imply a belief that or, more generally, a cognition that the desired object is good.\textsuperscript{33}

The motivation for a cognitive account of desire is straightforward. As I argued in the previous chapter, we must take the rational significance of desire into account in order to do justice to the truth behind the causal engine thesis. That, in turn, I argued, requires us to recognise that desire is a pro-attitude – that is, a state that comprises a positive evaluation of its object. But what could that mean other than that desire involves a cognition or judgment that the desired object is good?

Despite this apparently compelling case in favour of cognitivism, I shall argue that no cognitive account of desire is acceptable. I shall begin by arguing that the most simple version of cognitivism – the theory that identifies desire with evaluative belief – is obviously too strong. In the next two sections, I shall then consider two successive ways of weakening the simple cognitive theory. In the second section, I shall discuss and reject the thesis that evaluative belief is necessary for desire. In the third section, I shall consider the claim that a non-doxastic evaluative cognition is necessary for desire, but in the fourth section I shall argue that even this claim is false.

\textsuperscript{32} Anscombe [1976]; Audi [2002]; Dancy [2000]; Davidson [1980c]; De Sousa [1974]; Goldman [1970]; Oddie [2005]; Quinn [1993]; Raz [2001b]; Stampe [1986]; Scanlon [1998]. All of these authors hold that desiring that $p$ entails cognising that $p$ to be good. Still, some of them might object to their being labeled as holding a cognitive theory of desire. These authors might press that there is a distinction between thinking that $X$ is an \textit{a priori} presupposition of $Y$ – and, consequently, that $Y$ entails $X$ – and thinking that $X$ is a part of $Y$ (e.g. Davidson [2001]). In addition, they might clarify that they merely wish to claim that desire implies a cognitive state with evaluative content. Since these subtleties are irrelevant to the main issue, though, everyone who thinks that desire implies or presupposes an evaluative cognitive state will here be lumped into the camp of cognitive theories.

\textsuperscript{33} Note, it is plausible that the desire that $p$ presupposes the cognition that not $p$. On that basis, it could be said that desire has a cognitive part. However, I am not concerned with cognitive theories of that kind. Rather, as indicated in the text, I shall understand cognitive theories of desire to be characterised by the more particular claim that the desire that $p$ entails the cognition that $p$ is good.
2.1 Desire-as-Belief

Let us begin by clarifying the notion of a cognitive state. A cognitive state is a representation of purported fact. Belief is a cognitive state. Suppose you believe Antananarivo is the capital of Madagascar. Then you are in a state that represents Antananarivo's being the capital of Madagascar as a purported fact. Perception is also a cognitive state. Suppose you see a red wall. Then you are in a state that represents it as a purported fact that there is a red wall. In contrast, imagining and wondering are not cognitive states. To imagine a flying elephant or to wonder if Moses led the Jews out of Egypt does not involve having anything represented as a purported fact.

With this clarification in place, let us turn to cognitive theories of desire. The most simple of all cognitive theories, and a good place to start our discussion, is the desire-as-belief account. This is the thesis that an evaluative belief is both necessary and sufficient for desire.

The desire-as-belief thesis is too strong because holding an evaluative belief is not sufficient for desire. Suppose I believe it would be good, because morally required, to visit an ailing friend in hospital. Yet, despite that I somehow lack any desire to visit him. Nor is this possible distance between evaluative belief and desire a feature peculiar to moral belief. The fact that I know that exercising is a healthy and most pleasant pastime is entirely compatible with me lazing about on the couch without any desire whatsoever to exercise. Indeed, the intuition that we can know of all the good to be won and yet lack the desire for it applies to any kind of evaluative belief. We surely do often desire in accordance with our evaluative beliefs. But the fickleness of desire prohibits the claim that evaluative belief is sufficient for desire.

---

34 In some contexts, the notion of a cognitive state is also used to refer to a state that figures in the processing of information or simply to a content-bearing state. However, I shall use the notion exclusively to refer to representations of purported fact because what is characteristic of those theories that I group under the label 'cognitive theories of desire' is the idea that having a desire involves a representation in which it seems that the desired end is good.

35 In contrast to the philosophy of emotions, where emotion-as-belief theories have had their past and present advocates, the desire-as-belief thesis is most noteworthy for the amount of vigorous criticism it has been subjected to (see especially Lewis [1988]; Smith [1994]; Stocker [1979]). Some have defended the desire-as-belief thesis against these objections while stopping well short of endorsing the thesis (e.g. Price [1989]). Indeed, I am aware of only one person who explicitly endorses the desire-as-belief thesis: Humberstone [1987]. Moreover, Humberstone's motivation for this is not to account for desire being a pro-attitude. Indeed, he says that his account is not based upon "any special insight in philosophical psychology. It is, rather, from philosophical logic (broadly understood) that the inspiration for this account comes" (Humberstone [1987] p.50).

36 Some people claim that evaluative beliefs, or at least, beliefs to the effect that oneself morally ought to ϕ imply some motivation. I will discuss this thesis in chapter five. Note, however, that, even if true, it would not follow that there are some beliefs that imply having a desire. This would only follow if being motivated was the same as having a desire (which I argued in the previous chapter that it is not), or at least implied having a desire (a thesis I shall argue to be mistaken in chapter five).
2.2 Desire-Implies-Belief

The cognitivist can avoid this problem by weakening the thesis that desire is belief. Most obviously, she can drop the claim that evaluative belief is sufficient for desire, while continuing to maintain that an evaluative belief is necessary for desire. Call this the desire-implies-belief thesis. But is evaluative belief really necessary for desire? I shall discuss two arguments that show that this is not the case.

2.2.1 The Argument from Limited Cognitive Sophistication

Belief and desire are paradigm propositional attitudes: they are states that essentially have content that is propositional in form. But this may also disguise a difference between them. Some states necessarily have their content in a conceptualised way. To be in such a state involves possessing the concepts that are required to specify the state's content. Belief is such a state. One can't believe more people live in Sydney than in Canberra, or indeed, believe anything whatsoever, without possessing the concepts required to specify the content of the belief. However, it is controversial if all propositional attitudes must have their content in a conceptualised way. Intuitively, at least, this does not seem to be the case. For example, it seems that one can have a perception as of things being thus-and-so – say, as of there being a red apple – without possessing the concepts that are required to specify what one's perception is a perception of.

Intuitively, desire resembles perception more closely in this respect than it resembles belief. It seems that there are at least some desires that do not require the agent possessing the concepts necessary to specify the desired state of affairs – an intuition that is especially forceful in the case of desires for nourishment, excretion, or physical stimulation. This would help to explain why there is a certain oddity in ascribing beliefs to animals and very young children but no comparable oddity in ascribing perceptions and desires to them. The possession of concepts, after all, is no trivial cognitive affair. It certainly requires more than a disposition to reliably track instances that fall under the concept, and it arguably requires the demanding ability to have some grasp of the kinds of inference patterns that are being warranted by the deployment of the concept. Animals and very young children, we think, don't have this degree of cognitive sophistication. At the same time, they don't seem to lack anything required for perceptions and desires. If this is correct, then desire is a cognitively less demanding state than belief because it does not necessarily have its content in a conceptualised way. It would follow that desire can't imply belief.

37 Authors apparently sympathetic to the desire-implies-belief approach include Anscombe [1976]; Dancy [2000]; QSuinn [1993]; Raz [2001b].
The argument has a good measure of force. But it also relies upon some pretty contentious claims about the relation between content bearing mental states, the possession of concepts, and the nature of concept possession. Fortunately, a modified version of the argument does not suffer from these flaws. Let us concede that both desire and belief presuppose the possession of the concepts required to specify their respective contents. Even so, we must recognise that some concepts are much easier for us to grasp than others. For example, it is one thing to have the concept of a round object and quite another to have the concept of a nuclear fusion. Evaluative concepts fall in between these two extremes. Children acquire evaluative concepts such as something being good, desirable or right long before they understand what a nuclear fusion is but quite some time after they have mastered some non-evaluative concepts, such as shape and colour concepts and concepts of objects that figure importantly in their environments. Now consider a child who has not yet acquired the cognitive sophistication required to possess evaluative concepts. This child could not have evaluative beliefs. Even so, it seems, that the child could, say, desire to be fed. As Velleman succinctly puts the point, “a young child can want things long before it has acquired the concept of their being worth wanting” (Velleman [2000]: 104).

Still, the present argument is not foolproof, because it can be said that it is one thing to note that certain abstract evaluative concepts are beyond the ken of young children. It is another thing to establish that this holds true of all evaluative concepts. The concept of being pleasurable, for example, seems to be a fairly simple one and it is much less obvious that children could have desires prior to being able to believe that the desired end will be pleasurable. But then the present argument fails to show that desire does not imply evaluative belief because it does seem that children who have desires must also have the cognitive sophistication required to have certain simple evaluative beliefs.

This response clearly assumes that the belief that something is pleasurable is an evaluative belief. One might hold this if one thinks that pleasure is good and overlooks the distinction between evaluative properties and evaluative concepts. If we keep that distinction in mind, however, we see that, even if the belief that something is pleasurable represents a valuable state of affairs, it does not follow that the state of affairs is being thought about in an evaluative way, that is, in terms of it possessing value or being desirable. Another reason why some might think that the belief that something is pleasurable is an evaluative belief is that this belief is accompanied by an evaluative belief. Note, though, that this need not be the case. We can, for example,

39 This connects with the debate whether evaluative concepts are reducible to non-evaluative (and non-normative) concepts. I accept the force of Moore’s open question argument to show that this is not the case (Moore [1988]. See also Ewing [1944] and Hare [1952] for illuminating comments on the nature of distinctively evaluative or normative thought). Note, that this leaves open whether evaluative properties are reducible to non-evaluative properties. I think Moore’s open question argument is illuminating in this context as well, but this claim is significantly more controversial (for some important arguments and positions see Blackburn [1985]; Frankena [1938]; Jackson [1998]; Scanlon [1998]; Stratton-Lake [2002]).
believe that something is pleasurable in a way that includes the judgment that it is bad and that it ought to be avoided (think of taking pleasure in a malicious act). Likewise, it seems possible to believe that something is pleasurable without this being linked up with an explicitly evaluative belief at all. Indeed, it seems that this is precisely what is going on in the case of young children. Their earliest thoughts of pleasure seem to be devoid of any genuine cognitive evaluative content. They think something is or will be pleasant, but they do not think this in a way that includes the belief that this state of affairs has value or that it is desirable or that there is a reason to promote it.

Suppose that the critic modifies her objection in response. Suppose she says that desire only requires a belief that the desired object has a property F, which we in turn consider as a belief in the possession of a good-making property. A prominent proponent of the desire-implies-belief thesis, Joseph Raz, for example, clarifies that this is how he wants to be understood:

Stocker is mistaken to think that I attribute a kind of high order reflectiveness to people: for example that they not only think of their actions as pleasurable or thrilling or beneficial to X or Y, etc. but also think of them as good in virtue of possessing those properties. All I ever claimed is that people act for considerations which we classify as a belief in the possession of a good making property. (Quoted in Stocker [2004a]: 305)

However, this also will not do as a defence against the argument from limited cognitive sophistication. On the one hand, Raz might mean that desire merely implies the belief that the desired object has a property F and that we think that F is a good-making property. This would be compatible with the agent herself lacking the belief that F is a good-making property. Suppose, now, that all of this were true. It would then not only be misleading to say desire implies evaluative belief, but it would also undermine the motivation for cognitive theories. If it is merely us and not the agent who thinks of F as a good-making property, then we are giving up on the idea that we can illuminate the status of desire as a pro-attitude by appealing to her beliefs; after all, we can’t explain in virtue of what desire casts its objects in a positive light by recourse to evaluative beliefs that are not held by the desiring agent herself. On the other hand, Raz might mean that desire implies the belief that the desired object has a property F and that we understand this as the agent believing that the desired object has a good-making property. The agent may not believe the object is good, but at least she believes it has a good-making property. Whereas this would come closer to claiming that desire implies evaluative belief, it would face the full brunt of the argument from limited cognitive sophistication. Not only are there desirous agents who are incapable of believing that something is good, but there are, a fortiori, also desirous agents who

40 In this quote Raz does not explicitly speak about desire. The rest of his work (e.g. Raz [2001b]), however, makes clear that he would have to say the same about desire as well.
are incapable of believing that something has a good-making property, that is, incapable of believing that it has a property in virtue of which something can be good. In order to avoid the argument from limited sophistication, therefore, it is not enough to point out that simple desirous agents can deploy concepts that we often deploy in an evaluative way. Nor is it enough to point out that they could believe an object to have a property and that we (perhaps correctly) think that this is a good-making property. Rather, it must be shown that even the most simple creatures capable of having desires are nonetheless sophisticated enough to actually possess evaluative concepts and to think in evaluative terms. The responses just considered do not establish this thesis nor does there appear to be any other good reason to think that it is true.

2.2.2 The Argument from Recalcitrant Desires

In the Republic Socrates recounts a curious episode:

There is a story of a certain Leontius [...]. He was coming up from the Piraeus outside the north wall, and he passed a spot where there were dead bodies lying by the executioner. He felt a longing desire to see them and also an abhorrence of them; at first he turned away and shut his eyes, then, suddenly tearing them open, he said, “Take your fill, ye wretches, of the fair sight.” (Plato [1986]: 439e-440a)

The story of Leontius provides us with a stark reminder that believing something to be bad does not preclude desiring it; some of our desires are recalcitrant. The phenomenon of recalcitrant desires poses a problem for the desire-implies-belief thesis because recalcitrant desires seem to be accompanied by the belief that the desired end is bad rather than good.41

However, why should the agent not harbour both beliefs? That the agent believes the desired end to be bad is, it may be said, perfectly consistent with the claim that she also believes it to be good. To assess this response we have to distinguish between a strong and a weak version of the desire-implies-belief thesis. According to the strong version, to desire that \( p \) entails believing that \( p \) is good all things considered. On the weak version, in contrast, to desire that \( p \) only entails believing that \( p \) is good in some respect. Further, let us clarify that in the case of recalcitrant desires, the agent believes the desired end is bad all things considered. The strong version of the desire-implies-belief thesis would then have to say that an agent must believe the desired end is good all things considered even in those cases in which her desire is accompanied by the belief that the desired end is bad all things considered; in short, it would have to ascribe flat-

---

out contradictory beliefs to the agent. And this is a bad thing to do, because although all of us hold beliefs that are jointly contradictory, this is only possible so long as we remain unaware of the tension between them. But in the case of recalcitrant desiring, the agent is typically fully aware both of her belief that the desired end is bad all things considered and of her continuing attraction to the desired end.

In its attempt to handle the phenomenon of recalcitrant desires, the strong version of the desire-implies-belief thesis must ascribe contradictory beliefs in situations where we think no contradictory beliefs could survive. The weak version of the desire-implies-belief thesis avoids this problem because there is no contradiction in believing something to be bad all things considered and also believing it to be good in some respect. But I want to maintain that even the weak version of the desire-implies-belief thesis is at odds with an adequate account of recalcitrant desires. In some instances, at least, ascribing the belief that the desired end is good in some respect to the recalcitrant desirer seems to distort the phenomena. Consider, for example, the case of the wife-beating husband. He knows that his brutality is condemnable and often feels genuinely remorseful and ashamed. But that does not stop the desire to beat his wife cropping up in him from time to time, and it does not stop him from acting on his desire. Like Leontius, he desires something that he is fully aware is bad all things considered. Must he also believe that the beating of his wife is at least good in some respect, say, because he thinks it will feel good, or because he thinks it will establish his power and authority? He might hold such a belief. But to think that he must strains credulity, precisely because the relevant beliefs are so obviously implausible. If establishing one’s power is a good thing at all, it surely is not good if it has to be achieved by violence against the innocent. Likewise, while the fact that an action feels good is often a redeeming feature, it surely isn’t in this case. Indeed, taking pleasure in beating his wife compounds the overall badness of the act. Nor need the violent husband be deluded about any of this. Indeed, it seems that he could want to beat his wife, while knowing all along that doing so is not only bad all things considered, but bad all out.42

A second complication must now be addressed though, because it might be argued that this reasoning derives its plausibility from an improperly narrow focus upon the moral good. Moral goodness is an important kind of value, but it is not the only kind. An action can be morally bad, even morally bad all out, and yet be good in some other way, say aesthetically good or prudentially good. This suggests that while the violent husband does not think of his actions as morally good he may still think of them as, say, good with respect to his own well-being. But again, while the violent husband might believe that beating his wife is good for his own well-being, it is hardly credible to suppose that he must believe that. Surely, it can be glaringly obvious that his actions are not only morally bad but also detrimental to his well-being, say, because

42 Stocker [2004a] makes a similar point.
he knows himself to be under police surveillance. Nor need this only be obvious from a third person point of view. The violent husband, it seems, could truthfully say he never thought of his actions advancing any good whatsoever, but that he nonetheless found the desire inexorably arise within him.43

However, at this point, a challenge seems to arise. Suppose that the violent husband really does not believe his actions to be good in any way. In that case, it might be said recalcitrant desires would be pathological urges, mere causal forces that propel the agent forward, but not desires proper. If, on the other hand, the violent husband is to really desire beating his wife and not merely be inflicted by a causal force that drives him towards beating her, then it might be said he must also believe them to be good in some respect or other. Either way, appealing to recalcitrant desires can't undermine the desire-implies-belief thesis.44

This challenge is an important one, and it highlights a dimension of the desire-implies-belief thesis that its critics have all too often failed to pay sufficient attention to. They point out that it is implausible to insist that the desired end must be believed to be good, but fail to fully engage with this challenge. After arguing with much sensitivity and insight that desiring does not imply evaluative belief, Michael Stocker, for example, wonders why we should need or want to ascribe evaluative beliefs to all desirers. The answer, he says,

can't be that we need to do this in order to see their acts and reasons as intelligible. As argued, we can do that in light of their non-evaluative concerns, in terms of what they find attractive and compelling, in terms of their character and desires. (Stocker [2004a]: 328)

But here Stocker moves too quickly. No doubt, we can make actions 'intelligible' in terms of what agents find 'attractive and compelling'. However, that is not the same as saying that we can do this in terms of their 'non-evaluative' concerns. Desires bereft of any evaluative dimension could figure in what John McDowell has called "a style of explanation in which one makes things intelligible by representing their coming into being as a particular instance of how things generally tend to happen" (McDowell

43 I agree with Stocker that there are people who do not think in evaluative terms. [...] I do not mean only that they do not explicitly think, for example, 'This is good'. Rather, I also have in mind the following sorts of things: they may well deny thinking this; they can't be (easily) brought to see that they do think this; and, most importantly, the best accounts of their thinking and planning, of their character and intentions, do not involve such views. (Stocker [2004a]: 331)

The present point, however, is not that evaluative belief has no firm role in the psychology of the agent with recalcitrant desires but only that they need not back the desire. In my view, the structure that sets recalcitrant desires apart from cases in which limited cognitive sophistication hinders the deployment of evaluative concepts is that the agent is capable of having the appropriate beliefs but that these beliefs do not determine her desires.

44 Raz [2001a]; Raz [2001b]; Dancy [2000].
Cognitivism about Desires 32

This, as McDowell points out, is to be contrasted with “explanations in which things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be” (Ibid.). But as I argued in the previous chapter, for desires to figure in explanations of this kind, they have to have an evaluative dimension. So, only if desires are seen to be internally connected to the agent’s evaluative point of view will they be able to render actions intelligibility in the relevant sense. And this, the proponents of the desire-implies-belief thesis say, shows that desire must indeed involve some belief in the goodness of the desired end.

The problem is that this implicitly makes an assumption that is also accepted by critics like Stocker: namely, that the evaluative point of view is exhausted by evaluative belief. Let us agree that recalcitrant desires do indeed contain an evaluative dimension. The violent husband does not just find himself with an inexplicable compulsion to beat his wife in the way in which a patient suffering from Tourette-Syndrome finds herself with the urge to swear or a person suffering from Alien Hand-Syndrome can be driven to pick up objects in front of her. In some sense, let us say, the violent husband does, while in the grip of his desire, see some good in beating his wife. But let us also ask if the best way to explicate this intuition really is in terms of evaluative belief. A more realistic conception of the psychology of recalcitrant desires suggests otherwise. The violent husband does not, after all, seem to be suffering from what is, at bottom, an intellectual infliction. His problem is not that he desires to do what is good but is confused about the value of beating his wife. His problem simply is that, at bottom, his desire is orientated towards the bad. He is afflicted with a perversity in his desire, a perversity that can persist despite the recognition of the real disvalue of the desired state of affairs. What recalcitrant desires, therefore, call our attention to is the difficulty in accounting for the evaluative dimension in terms of evaluative belief. They highlight that this analysis violates, and in the end misconstrues, our intuitions about the conflicts that are possible between what a person desires and what she believes to be good and bad.

---

45 There arguably exists a third notion of intelligibility over and above mere regularity and rationality. One may find an oversight intelligible if an agent is stressed or tired not because one knows that agents in these conditions do regularly make that kind of oversight, nor because it is rational, but because one can imagine making the oversight oneself if one would have been in those conditions. This notion of intelligibility raises interesting questions, which are, however, not essential to the primary concerns of this thesis.

46 See Della-Sala et al. [1991]; Jankovic [1997]; Lhermitte [1986]. In this context, see also Johnston [2001a] for subtle observations on the phenomenal intelligibility of desires.

47 Raz writes,

The classical approach allows that a gap may open between what is good and what attracts agents. But it accounts for the gap by the fact that agents can make mistakes. Certain actions appear to them attractive because they appear to them to possess good-making properties, but in fact they do not possess them. (Raz [2001a] p. 28).

This, I think, is a caricature of the phenomenon.
All of this, then, goes to reinforce two points. Desiring does involve some kind of positive evaluation of its object. But at the same time the evaluative element in desire is not a matter of belief. The problem is, how else shall we understand this if not in terms of belief?

### 2.3 Non-Doxastic Cognitive Theories of Desire

Suppose I look at a picture of the Mueller-Lyer illusion. It *seems* to me that the two arrows are of different lengths. At the same time, I do not *believe* they are of different lengths. I know of the nature of the illusion, and I know that the two arrows really have the same length. Despite that, however, it continues to *appear* or *seem* to me that one arrow is longer than the other. My mental state represents the arrows being of different length, and this representation is one of *purported fact*, but it is not an instance of belief. Instead, it is an instance of a non-doxastic cognition.

This suggests a way to develop the cognitive account of desire that avoids the above problems. According to the *non-doxastic cognition* thesis, a non-doxastic evaluative cognition is necessary for desire. This don-doxastic cognition thesis seems to be immune to the argument from limited cognitive sophistication, because there is no obvious reason to think a non-doxastic evaluative cognition would have to place higher demands upon the cognitive sophistication of an agent than desire does. It may also seem to escape the argument from recalcitrant desires because that argument specifically targeted our intuition that an agent need not always believe the desired end to be good in any way. In short, the non-doxastic cognition thesis promises to explain what the cognitive theories set out to explain – the pro-attitudinal character of desires – while avoiding the objections that burdened the most obvious way of pursuing that agenda.

What is this non-doxastic evaluative state? One suggestion is that it is an evaluative *perception*, since perception is the paradigmatic non-doxastic cognitive state. It would also make good sense of the pervasiveness of perceptual language in the theory of desire: that the desirer *sees* some good in the desired end, that the desired end *appears* good to her, and so on and so forth. But we need not understand the non-doxastic cognition thesis this way. On another interpretation, what is necessary for desire is a *non-doxastic and non-perceptual* evaluative cognition. Indeed, I think this is probably the best thing to say. Here is why. Perception stands in a conceptual connection to the idea of causation: under ideal conditions a perception has to be caused by its object. Suppose I am childless but desire that I have children. On a perceptual interpretation of the non-doxastic cognition thesis, this involves the perception that my having children is good. But note, since I am childless it can’t be the case that my having

---

48 See Davidson [1980a]; Davidson [1980c]; Oddie [2005]; Scanlon [1998]; Stampe [1986]; Stampe [1987]. See also the rich account in Johnston [2001a]. Note, a number of these authors also think that a non-doxastic evaluative cognition is sufficient for desire. This stronger claim is not crucial to our discussion.

children is good. Since non-obtaining state of affairs do not have causal powers, the perception that my having children is good could not have been caused by its object. But in ideal conditions, a perception has to be caused by its object. It would follow, that in ideal conditions one can’t be childless and desire to have children. That seems false.

The non-doxastic cognition thesis is false whether or not we give it a perceptual twist. One can’t claim that “the desire that p is p’s seeming good” (Oddie [2005]: 42) or that “what a desire is is a perceptual state, a state where in it seemed as if some state of affairs would be good” (Stampe [1987]: 363). One can’t claim this because there is a deep problem with the very idea of desire implying any evaluative cognition. Or so I shall argue.

2.4 Desire and the Norm of Truth

All cognitive theories maintain that an evaluative cognition of some sort is necessary for desire. That is, they maintain that desiring involves a representation of purported evaluative fact. The question is whether desire really can be understood as involving such a representation. I shall suggest a simple argument to the effect that this is not the case. The basic idea is this: cognition is subject to the norm of truth. But desire is not subject to the norm of truth; it can’t, as David Hume put it, “be oppos’d by, or be contradictory to truth” (Hume [2003]: 3.1.1.9). So, desiring can’t be or imply a cognitive state.

The premise that cognition is subject to the norm of truth should be uncontroversial. A cognition just is a state which represents things as being a certain way and as such is subject to an assessment in terms of whether things really are the way it represents them as being. In the ideal case, cognitive states will satisfy the norm of truth; they will represent things to be the way they actually are. If, on the other hand, they represent things as being otherwise than they actually are, then cognitive states

50 Perhaps my having children would be good, but it can’t be good. However, might it be said that the content of the perception is that the desired end would be good? No, because perception does not have counterfactual objects.

51 Stampe [1987] who endorses a perceptual theory of desire attempts to avoid these difficulties in two steps (Stampe [1987] p. 371-376). First, he claims that there is a bodily or mental state Σ which is the object of ’nepistemic desiderative perception’. Second, he claims that being in this state (or perhaps perceiving this state, Stampe is not clear on which option he has in mind) causes it to seem to the subject that the desired end is good. How exactly this is supposed to work is left unclear, but even so the proposal does not solve the difficulty. After all, if the evaluative cognition is a perception, it would have to be caused by its object, at least under ideal conditions. But Stampe accepts that it isn’t so caused – instead it is caused by a bodily or mental state Σ. (There are other worries here as well. For example, Stampe who had previously identified desire with an evaluative perception now says that ”desire is one’s nepistemic perception of this state of oneself” (p.372) and that this state is ‘distinct’ from the evaluative perception (Ibid.).

52 Note, Stampe speaks of ‘seems’ to distinguish the seeming involved in desiring from the phenomenal seeming characteristic of sense perception. Note further, that while Stampe says that desire is a perception he also says that desire is a ‘peculiar’ kind of perception (Stampe [1987] p. 371).
manifest some kind of failing. That is not to say one ought always to rid oneself of such cognitions, but it does mean that the cognition is false or unveridical and to that extent falls short of an inbuilt standard.

That a state is subject to the norm of truth must be distinguished from its having semantic truth-conditions. Any state with content determines semantic truth-conditions but not every contentful state is subject to the norm of truth. To imagine that \( p \), for example, is to be in a state that has \( p \) as its semantic truth-conditions. But being in that state does not make one subject to the norm of truth. There is no standard one falls short of by being in that state if \( p \) is not in fact the case. To mark this distinction, we might say states like imagining that \( p \) are states with content but not states with cognitive content.53

Are desires subject to the norm of truth? It does not seem that way. To see this, consider the criticisms we do level against desire. First, we criticize instrumental desires if they are based upon false belief. Thus we might object to Chris’ desire to holiday in Fiji if it is based upon the belief that Fiji is much like downtown New York, just a bit bigger. However, not all desires are based upon beliefs in this way – non-instrumental desires aren’t.54 Second, we do object to desires on broadly ethical grounds. Suppose, for example, that Ronald is a virtuous man who has worked hard to rise to his position and, because of this, I desire his downfall. Clearly, such a desire is deeply flawed. But this is not because it is epistemically problematic but because it is morally objectionable. The appropriate criticism to level is not that it is false or unveridical but that it is mean-spirited and base. Likewise, if you want to gamble all your savings on a highly unlikely business proposition, then there is something seriously amiss with your desire. But again, this is not because it is false but because it is imprudent. So, there are moral and prudential standards to which we hold desire. But we do not subject desire to the norm of truth.

The cognitivist might respond that this is true on the level of what we say but false on the level of substance. We call a desire base and a belief false, but the underlying criticism is the same. In both cases our criticism is effectively that the norm of truth has been violated.

53 See Martin [2002] and Schwitzgebel [1999] for discussions of the potential confusion between these two notions of content in the philosophy of mind.

54 Moreover, it is not sufficient for cognitivism to be true that desire is subject to the norm of truth. Suppose that desire implies the cognition that the desired state of affairs does not obtain. In that case, desire would be, or at least imply, being subject to the norm of truth with respect to existence of the desired state. What the cognitivist needs, however, is for desire to be subject to the norm of truth in a different way. The cognitivist needs it to be the case that by desiring one is subject to the norm of truth with respect to the desired state of affairs being good.
Upon reflection this response is quite implausible. The criticisms we bring against desire do not just seem to be different in terminology to epistemic criticism, but different in nature. When we call the belief that his downfall would be good ‘false’, and the desire for his downfall ‘base’, it does not appear as if we are levelling the same kind of criticism just clothed in different terms. Rather, we seem to identify different kinds of faults with the belief and the desire. The fact that the circumstances in which a desire is subject to ethical censure may overlap to a certain extent with the circumstances under which the respective evaluative cognition would be subject to censure under the norm of truth easily obscures this point. But note, even if we criticised desires in all and only those circumstances in which the desired object was not good, it would not follow that the point of our censure was epistemic and due to desire falsely purporting the desired object to be good. In fact, however, there are circumstances in which ethical and epistemic appraisal come apart. For example, if I dislike Ronald and want to see him lose his job, then this is an ugly and vicious desire. This remains the case, even if losing his job would be the best thing that ever happened to Ronald. In that case, the cognition that losing his job would be good could not be faulted with respect to the norm of truth, but our criticism of the desire would lose none of its force. In a similar vein, we can point out that in some cases the desired end is in fact bad, but that the desire for it still is praiseworthy. Next, the cognition that \( p \) is good fails to live up to its inbuilt standard whether or not \( p \) is a great evil or slightly bad. In other words, assessment under the norm of truth knows no degrees of failing. Assessment of desire, on the other hand, follows a different logic. It does allow for degrees. A desire for \( p \) is, prima facie, subject to less stringent criticism if the desired state is slightly less bad than if the desired state of affairs involves great evil.\(^{55}\) Finally, the cognition that \( p \) is good is also subject to criticism if \( p \) is evaluatively neutral. And though there are cases in which a desire for an evaluatively neutral state can be subject to criticism (just as there are cases in which a desire for a state of affairs which is in fact good can be subject to criticism), it is not credible to believe that just because the desired state of affairs is evaluatively neutral the desire manifests some failing.

All in all it seems clear, therefore, that the fact that desire is subject to ethical norms does not corroborate the claim that desire is subject to the norm of truth. However, the cognitivist may say that we have overlooked additional forms of criticism of desire and that these forms of criticism reveal our commitment to the idea that desire is, after all, subject to the norm of truth. What other sources of criticism might the cognitivist have in mind? The cognitivist might claim that we also criticise sets of desires if their objects are logically inconsistent or if they could not be jointly satisfied if the agent’s beliefs were true.\(^{56}\) Moreover, the cognitivist can claim that like the criticism of

---

\(^{55}\) Might the cognitivist reply that even though cognitions are either true or false, their truth can be more or less obvious and that we censure more if the evils are more obvious? No, for apart from making nonsense of the praise of desires, we don’t object more virulently to some desires than others because the desired end is more obviously bad, but because the desire is worse.

\(^{56}\) Oddie [2005] pp. 33-34 makes these claims, though he does not make them in the context of trying to show that desires are subject to the norm of truth. As far as I can see, he does not address that challenge.
inconsistent sets of beliefs, that criticism is an epistemic one: it hinges on the fact that jointly inconsistent propositions can’t be true. Therefore, the cognitivist concludes, desires are subject to demands of truth. This, however, is a weak argument for two reasons. First, it is not obvious that it is indeed appropriate to criticise sets of desires in the fashion the cognitivist claims. Suppose that Luciano desires to see the soccer game and also desires to go to the final night of the opera but knows that he can’t do both. Surely, there is nothing wrong with him having both these desires. Moreover, suppose that Luciano has some attachment to both soccer teams and finds himself desiring that one win and also desiring that the other win. Would Luciano in that case be subject to criticism? Perhaps, but it is hard to be certain. Moreover, even if it would be appropriate to censure inconsistent sets of desire (or sets of desires that could not be jointly satisfied if the agent’s beliefs were true), this does not support cognitivism. This is because it does not follow from the fact that the desire that \( p \) is inconsistent with the desire that not \( p, \) that \( p \) and not \( p \) could not both be good. A state may be good in respect \( R, \) and its negation good in respect \( R^* \). A state may be good all things considered, but its negation still be good in some respect. A state may be morally good all things considered, but its negation be good prudentially (in a respect, or all things considered). Consequently, there is no easy inference from a set of desires being objectionable, because inconsistent, to the thesis the cognitivist needs, namely, that it is objectionable because the implied cognitive evaluations could not all be true.

Desire is not subject to the norm of truth, but cognition is. It follows that desire can’t be a cognition, not even in part. Still, the cognitivist might maintain that this does not contradict her thesis, because her claim is that an evaluative cognition is necessary for desire, and this, she might say, is a weaker claim than saying that such a cognition is a constitutive part of desire. The premises of the argument, therefore, do not imply the needed conclusion. Does this move save the cognitivist? I don’t think so. The same kind of reasons that should convince us that desire itself is not subject to the norm of truth should also convince us that a person is not necessarily subject to the norm of truth when having a desire. In other words, when we carefully consider the kinds of norms and standards that are implied by desire, then we must conclude that they do not include the norm of truthfully representing the desired state of affairs as good. This contradicts the thesis that an evaluative cognition is necessary for desire. If cognitivism about desire is to be saved, it would have to be shown that desire really does make us subject to the norm of truth; that we necessarily fall short of an inbuilt truth-relative standard if our desires are aimed at ends that are not, in fact, good (not just if they are aimed at ends that are bad, but also if they are aimed at ends that are

58 The cognitivist could claim that desires necessarily represent all things considered evaluative properties of a certain kind. Desiring that \( p \) would then involve the cognitive representation that, say, \( p \) is prudentially good all things considered. Even that, however, is consistent with not \( p \) also being prudentially good all things considered. Moreover, once the cognitivist committed herself to such a stark thesis about the cognitive content of desire, she would face countless counterexamples in which a desire is unobjectionable but the desired end is not, say, prudentially good all things considered.
evaluatively neutral!). I do not see how the cognitivist could show this. Therefore, I conclude, that any cognitive theory of desire is flawed. Evaluative cognition is not necessary for desire.

There is a natural inclination to think this conclusion can't be right. After all, doesn't the evaluative dimension of desire and its rational significance require some cognitive theory to be true? To give up on cognitivism, it seems, comes at the cost of being unable to vindicate the thought that desire is a pro-attitude and all that is implied by it. Clearly, there is a serious challenge here. To refute cognitivism it does not suffice to show its weaknesses, but we must also put forward a credible alternative. In the next chapter, I shall take up the task. I want to conclude the current chapter with some more general reflections that underpin the need to go beyond cognitivism. We are familiar with a whole array of mental phenomena that appear to have evaluative and rational significance without being cognitive in nature. The assumption that drives cognitivism about desire and that threatens to introduce a fundamentally foreign element into our thought about it, would therefore also seem to steer us towards awkward analyses of such ubiquitous phenomena as liking or disliking something, taking pleasure in something or enjoying it. A case can be made that being in such states involves having an evaluatively coloured conception of the object of one's liking or disliking. But, just as in the case of desire, it seems intuitively wrong to explicate the evaluative character of these states in terms of cognition. Indeed, when reflecting upon both the phenomena of taste and desire we are drawn to ideas and metaphors wholly removed from the domain of cognition. We are drawn to speaking of the agent as a creator of value and not just as holding up a mirror to the evaluative structure of the world. We intuitively think these phenomena involve the agent's investing objects with importance and value. That through taste and desire agents impose their ineliminably subjective evaluative colouring onto the world without simply being beholden to an epistemic standard set by the world. Of course, merely to recall these age-old intuitions that form the central pillar of the non-cognitivist tradition in philosophy is not yet to answer the cognitivist's challenge. Nothing has been said to show how agents can invest the world with value without being in a state with cognitive evaluative content. Nothing, that is, has been said to show that non-cognitivism about taste and desire is a coherent alternative. But enough should have been said to highlight the wider significance of challenging the motivations that stand behind cognitivism about desire and to draw attention to the importance of considering afresh the possibility of a non-cognitivist conception of desire.
Chapter 3

Modes of Intentionality

The results of the first two chapters pose something of a dilemma. We know that desire must be a pro-attitude – this much we learned from the first chapter. And the second chapter showed that this can't be understood in terms of desire involving an evaluative cognition of the desired end as good. Yet, it seems that the only way for desire to be a pro-attitude is for it to involve such an evaluative cognition.

A natural reaction at this point would be to reject the idea of desire being a pro-attitude. But this comes with a heavy cost. It is intuitively plausible to think that desire comprises some kind of positive appraisal of its object. Also, it is exceedingly plausible to think that desire bears upon the rationality of motivation and action and it can do so only if it is a pro-attitude. And there does seem to be a grain of truth behind the causal engine thesis, and to capture this insight we must think of desire as a pro-attitude. At the same time, the arguments to the effect that desire does not involve an evaluative cognition remain forceful. Taking evaluative belief to be necessary for desire imposes implausibly strict requirements on the cognitive sophistication of desirous agents and also conflicts with our understanding of the rift that can open up between recalcitrant desire and evaluative belief. If we construe the relevant evaluative cognition in a non-doxastic way, we may avoid these difficulties, but we would still not avoid the implausible treatment of desire as being subject to the norm of truth. There is no avoiding the conclusion that desire appears to be a mental state of a quite different kind from cognitive mental states. Yet, it seems we can reject cognitivism about desire only on pain of giving up on desire's status as a pro-attitude. We have come full circle without having found a theory that can accommodate even the most basic ideas about the nature of desire.

The central difficulty is to see how desire could be a pro-attitude if not by incorporating an evaluative cognition. It seems we must choose between an evaluatively blind, mere causal, conception of desire on the one hand and a cognitive theory on the other. The task of this chapter is to put the conceptual framework in place to show that these options are not exhaustive. There is a third way. There is room for a viable form of non-cognitivism about desire.

To put the needed framework in place we shall need to think about the nature of intentionality and evaluation more generally. This will allow us to recognise that there are two dimensions of intentionality and two dimensions of evaluation. There are two dimensions of intentionality because intentionality is not only a matter of content but also a matter of the mode under which a content is being had. And there are two dimensions of evaluation because evaluation is not only a matter of evaluative content...
but also of evaluative modes of representation. With these key claims in place, an interesting possibility emerges: desire is a pro-attitude in as much as it is constituted by an evaluative mode of intentionality.  

In the first section, I shall defend the claim that there are two dimensions of intentionality. I shall explain what a mode of intentionality is and argue that it is distinct from content. In the second section, I shall defend the claim that affective intentionality is a fundamental mode of intentionality. In the third section, I shall consider whether modes of intentionality are reducible to functional properties. I shall argue that they are not. In the fourth section, I shall argue that modes of intentionality are instead anchored in distinctive conscious episodes. Finally, in the fifth section, I shall draw out the implications of all this for our understanding of evaluation. Having done so, the path towards a viable form of non-cognitivism about desire should be clearly visible.

3.1 Modes of Intentionality

Many, perhaps all, mental states have a peculiar property: they are about things. Consider, for example, thinking that Shakespeare wrote Hamlet. That thought is about Shakespeare’s writing Hamlet. Likewise for remembering that Shakespeare wrote Hamlet, doubting that Shakespeare wrote Hamlet, wishing that Shakespeare wrote Hamlet and so forth. All of these states are about Shakespeare’s writing Hamlet.

Call states that have this peculiar property of aboutness intentional states. Mental states are not the only things marked by intentionality. Linguistic entities provide another paradigm example. The English sentence “Shakespeare wrote Hamlet”, for instance, is also about Shakespeare writing Hamlet. Not all sentences are about something; some sentences are meaningless. But what a sentence is about, if it is about anything at all, is some actual or merely possible state of affairs. This state of affairs is the extension of the sentence. As Frege pointed out, however, the meaning of a sentence is not exclusively determined by its extension. The sentence “Shakespeare wrote Hamlet” and the sentence “The Bard wrote Hamlet” have the same extension, but differ in meaning. This is because although they have the same extension they have a different intension; both are about the same state of affairs, but they represent that state of affairs in different ways.

59 The basic tenets of this framework are not new. They can be found in Brentano [1902] and Brentano [1911]. However, while Brentano’s thesis that intentionality is the mark of the mental has generated wide interest, his discussion of different intentionality relations has almost entirely been ignored in the contemporary literature. The only contemporary writers who make mention of modes of intentionality are, as far as I am aware, Crane [2001]; Crane [2003]; Searle [1983]. However, neither Crane nor Searle seriously attempt to elucidate the notion of a mode of intentionality.

60 Frege [1994]. Idiosyncratically, Frege took the extension of a sentence to be its truth-value.
As has often been observed, something similar appears to apply in the case of mental intentionality. There is a difference in thinking that Shakespeare wrote Hamlet and thinking that the Bard wrote Hamlet, which consists in the different ways in which the same state of affairs is being thought about. We can mark this by saying that although the two states have the same extension, they involve distinct *modes of presentation* – they differ in the concepts that the agent deploys to think about the state of affairs. (Speaking of the *content* of an intentional state is ambiguous between these two ideas. It can refer to the represented state of affairs or to the particular mode of presentation. Often there will be no need to resolve this ambiguity.)

The notion of a mode of presentation is well entrenched in the debate about the nature of mental intentionality. It is, therefore, natural to assume that my emphasis on modes of intentionality is but an idiosyncratic way of drawing attention to the phenomenon of modes of presentation. Not so. The phenomenon of a mode of intentionality, as I understand it, is quite distinct from the phenomenon of a mode of presentation. To clarify what I mean by a mode of intentionality, we must return to the case of linguistic meaning.

The extension and the intension of a sentence do not fully characterise it – a point once again first noted by Frege. Importantly, sentences that have the same extension and the same intension can nonetheless differ in a significant way if they have been put forward with different *force*. Contrast the *assertion* ‘Sam smokes habitually’ with the *command* ‘Sam, smoke habitually’. The two sentences have the same extension and the same intension – they refer to the same person, predicate the same property and do so under the same mode of presentation. Nonetheless, different things are being conveyed by the two sentences. If you knew only the content of a sentence, but did not know with what force it had been put forward, you would be missing an essential element.

What I want to draw attention to is that there is a close analogue of sentence force in the case of mental intentionality. A full characterisation of, say, the belief that Shakespeare wrote Hamlet must not only take into account the extension and mode of presentation of this belief, but must also emphasise its peculiar ‘force’; that in believing this one regards something as true. It is this phenomenon of mental ‘force’ that I have in mind when I speak of modes of intentionality.

---

61 There may also be non-conceptual ways for representing an object in a particular way (Peacocke [1986]). See Stalnaker [1987] for an important discussion of how to do away with the theoretical appeal to modes of presentation.

62 Frege [1986]. The example is Searle’s (Searle [1969]: 22). Further important reflections on the phenomenon include Austin [1962]; Davidson [1984] and Dummett [1993].
It is crucial to recognize that modes of intentionality are not a dimension of the content of mental states. Two states can have the same content yet differ in ‘force’. Believing that Shakespeare wrote Hamlet and entertaining the thought that Shakespeare wrote Hamlet have the same content, but the belief is characterized by a ‘force’ that is absent in the case of a mere thought. Of course, this might be called into question. It might be said that a belief that \( p \) is simply a thought that \( p \) is true. The supposed difference in ‘force’, on this account, reduces to a difference in content. Alas, there are numerous objections to this proposal.

(i) The proposal makes it mysterious that we should have voluntary control over our thoughts but not over our beliefs. In Hume’s words

the mind has the command over all its ideas, and can separate, unite, mix, and vary them, as it pleases; so that if belief consisted merely in a new ideas, annex’d to the conception, it wou’d be in a man’s power to believe what he pleas’d. (Hume [2003]: Appendix)

(ii) While entertaining a proposition in thought and believing it to be true are both states with extensions and – as we may put it, semantic truth-conditions – only belief is subject to the norm of truth. In the case of belief it makes sense to assess the state against an ideal of truth. In the case of mere thought or imagination it does not.\(^6^3\)

(iii) If one person believes there to be a red ball and another does not, then there is genuine disagreement between them. But how could this be if belief was just a matter of having a certain kind of thought? After all, if one person thinks of a red ball and another of a blue ball, then the two people have different thoughts but there is no disagreement between them.\(^6^4\)

(iv) Believing that \( p \) can’t consist in a thought with the content \( p^* \), because a person can entertain any content in thought without believing it to be true. After all, any content can appear in ‘unasserted contexts’, such as the antecedent of a conditional or as part of a disjunction. Entertaining a content in such a context does not imply believing it to be true. This shows that we can’t capture the ‘force’ of belief by appeal to what is being represented.\(^6^5\)

If this is so, however, why should we speak of modes of intentionality? The answer is that, although modes of intentionality are distinct from mental content, they are inextricably tied to it. A mode of intentionality is a particular manner in which a

\(^{63}\) Brentano [1911] Ch. 3, §9.

\(^{64}\) Ibid.

\(^{65}\) Frege reached the same conclusion with respect to sentence force: “In an assertion we express the recognition of truth. We do not need the word “true” for that. And even if we do make use of it, the assertoric force does not consist in it ... where the sentence has lost its assertoric force, the word “true” can’t re-establish it” (Frege [1986]: 35-36. My translation).
given content is apprehended, for example, as true. As such, it is essentially related to mental content. We can't make sense of modes of intentionality without taking into account the ways in which they are linked to mental content. And this is why we would be missing something in the nature of the subject if we took mental 'force' to be wholly non-intentional.

How many modes of intentionality are there, and are all intentional mental states marked by a mode of intentionality? There are two equally legitimate ways of answering these questions. The first answer was given by Franz Brentano, whose close attention to the phenomenon remains unsurpassed. Brentano thought that there are three basic modes of intentionality – which I shall dub cognitive intentionality, affective intentionality, and pure intentionality – and that all intentional mental states are marked by at least one mode of intentionality. Cognitive intentionality involves a state’s having a content in a manner that involves regarding it as true or false. Affective intentionality involves a state’s having a content in a manner that involves being for or against it. And pure intentionality simply involves a state’s having a content. The other answer is to restrict modes of intentionality to two basic categories: cognitive and affective intentionality. The guiding idea here is that modes of intentionality ought to positively add something to the mental reality of a state and that the notion of pure intentionality fails to do that. On this account, some intentional states, like entertaining the thought, or imagining that, Shakespeare wrote Hamlet will not be marked by any mode of intentionality. Whichever answer we prefer will not make a substantive difference for our purposes.

Let me sum up. I argued that in addition to being characterised by mental content, intentional mental states can also be characterised by a peculiar mental ‘force’, analogous to sentence force. This is the phenomenon I refer to by speaking of modes of intentionality. Modes of intentionality are essentially linked to mental content as they are manners in which the content is being apprehended. Cognitive and affective intentionality are the two basic categories of mental ‘force’. Some intentional mental states aren't characterised by either. Of these we can say that they involve a pure mode of intentionality or that they lack a mode of intentionality.

### 3.2 Affective Intentionality

This brief account of modes of intentionality is sure to raise many questions. One of these is whether there really is an affective mode of intentionality. This question will be taken up in the present section. Other questions, such as the relation of modes of intentionality to functionalism and consciousness will be deferred until later sections.
I have said that affective intentionality is one fundamental mode of intentionality. Put another way, the claim is this: some intentional mental states are characterised by a peculiar evaluative ‘force’; they involve a pro-ness or con-ness towards the content. Moreover, this dimension of these intentional mental states can’t be reduced to a pure or cognitive mode of intentionality.

These claims are perhaps easiest to explicate and defend in the case of hedonic mental states. Hedonic mental states include simple sensory pleasures, such as enjoying the taste of chocolate or taking pleasure in the soothing touch of a massage. But as I understand them, they are not limited to sensory pleasures but also include states such as being happy about the blooming of the flowers, being glad to learn that a brooding conflict had been resolved, or taking pleasure in one’s child uttering its first word. (Of course, there are also negative hedonic states such as disliking the taste of oysters or being saddened by the news that Aunty Linda’s health is deteriorating. For simplicity’s sake, however, I shall focus upon positive hedonic states.)

It is very natural to say that hedonic states are characterised by a distinctive evaluative dimension. Consider taking pleasure in the blooming of the flowers. In that case, one does not merely represent that the flowers are blooming, but one represents it in a certain positive light, with a certain kind of positive ‘force’. One’s representation is not dispassionate, as we might say, but infused with a pro-ness towards the blooming of the flowers. On the view I want to endorse, taking pleasure in the blooming of the flowers involves such a positive ‘force’ because it involves apprehending the blooming of the flowers under an affective mode of intentionality.

Hedonic phenomena provide a good test case for the claim that affective intentionality is a fundamental mode of intentionality. There are two primary objections that can be made against the thesis that hedonic states are characterised by a distinctive affective mode of intentionality. The first is that appealing to an affective mode of intentionality is redundant as we can explain hedonic phenomena in terms of pure or cognitive intentionality alone. The second is that hedonic phenomena aren’t characterised by any mode of intentionality as they are essentially non-intentional phenomena. I shall consider these objections in turn.

Suppose it is said that being happy about the blooming flowers is the same as being in a mental state with a certain content, such as that the blooming of the flowers is good. On this view, affective intentionality reduces to pure intentionality. This is not credible. It is obvious that any thought could be had without the subject’s being happy. Now suppose it is said that being happy about the blooming flowers is the same as apprehending that the blooming of the flowers is good under a cognitive mode of intentionality. This too is not credible. Again it seems that one could regard as true
any proposition without being happy. It is one thing to believe, judge, or cognise more generally and quite another thing to be happy or sad. Being happy, as we might put it, is a matter of sentiment and not of intellect, an exercise of Gefühl and not Erkenntnis.

This suggests a more sophisticated way to account for the affective quality of happiness. It is easy, at least to begin with, to be sympathetic to the thought that what is special about sentiment is that it involves being affected or touched by something; that it involves certain kinds of feelings, and that these feelings are nothing less, but also nothing more, than bodily feelings. Powerfully expounded by William James a century ago and still influential today, this contention has a measure of plausibility. The way we feel our muscle tone relaxing, our heartbeat easing, or tears welling in our eyes certainly captures part of the idea of happiness being a sentiment. To be sure, being happy about the blooming of the flowers can’t be the same as experiencing a bunch of bodily sensations as it is evident that one could experience any set of bodily sensations without being happy about the blooming of the flowers. But it might be said that we can do justice to hedonic phenomena by combining the claim that hedonic states comprise being in a state with a certain content (or with a certain content under the cognitive mode of intentionality) with the claim that hedonic states involve experiencing distinctive bodily sensations.

This proposal strikes me as inadequate on phenomenological grounds. Suppose you experience a range of bodily sensations characteristic of being happy, like tears welling in your eyes, and the easing of your heartbeat. Suppose, further, that you also believe the blooming of the flowers is good. It does not seem that this must add up to anything like experiencing happiness about the blooming of the flowers. You may simply believe one thing and also experience some bodily sensations at the same time. In being happy about the blooming of the flowers, the evaluative seems fused with, rather than just accompanied by, the affective. It is hard to see how the present proposal could do justice to this.

There are also other problems with this proposal. In being happy about the blooming flowers it seems that one does evaluatively endorse the blooming of the flowers in some way. This evaluative element can hardly consist in the experience of some bodily sensation nor in merely entertaining a proposition in thought. The thesis that the evaluative element in hedonic states consists in apprehending an evaluative proposition under a cognitive mode of intentionality stands at least a fighting chance of accounting for this evaluative dimension, but must, ultimately, fall prey to the same difficulties that plagued the cognitive theory of desire. First, taking hedonic states to involve evaluative beliefs would impose demands that are too onerous upon the cognitive sophistication of creatures capable of being in hedonic states. After all, it seems

---

67 For more extensive criticism of the attempt to reduce affect to bodily sensation see especially Goldie [2000]; Green [1992]; Roberts [1988]. See also Lambie and Marcel [2002].
creatures that lack evaluative concepts can nonetheless be happy about things and take pleasure in the contents of their perception. Second, intuitively we are familiar with the possibility that someone may be happy that $p$ but fails to believe that $p$ is good – not because she lacks the ability to deploy evaluative concepts, but because she simply does not regard it as true that $p$ is good; her happiness can be recalcitrant. We might, at this point, consider whether happiness involves a non-doxastic evaluative cognition. However, even that seems implausible, because being happy about something is not a state that is evaluable under the norm of truth. To be sure, like desiring, being happy is subject to certain kinds of criticisms and assessments. One can exhibit a serious moral failing by being happy about the wrong kinds of thing, such as the wanton infliction of pain upon the innocent. But being happy is simply not a state that is subject to the norm of truth.

The first objection fails: we can’t explain hedonic phenomena in terms of cognitive or pure intentionality. But what about the second objection? A mode of intentionality is inextricably tied to the content of an intentional state since it is a manner in which this content is being apprehended. Yet it might also be objected that hedonic phenomena lack any intentional nature; they can accompany intentional states, but in and of themselves they are wholly non-intentional.

This view gets some support from hedonic states that appear to lack any object, such as when one apparently is merely happy and full of joy or merely displeased and depressed. One may, for example, wake up one morning to find oneself in a positive or negative mood – that is, in a state that appears to be one of happiness or sadness without there being anything that one is happy or sad about. The point is not merely that one does not know what caused the mood. Rather, there does not appear to be anything upon which the positive or negative mood is directed.

In this respect, hedonic states are similar to other emotions. While, in paradigmatic cases, one is afraid of something, there are instances in which one feels afraid without there seeming to be anything that one is afraid of. Likewise, there is a phenomenon of being angry without there seeming to be anyone whom one is angry with. Apparently objectless emotions have a tendency to transform themselves into object directed emotions. Crossing my path, you may become the object of my anger that hitherto did not seem to be about you nor about anyone else. The same goes for hedonic moods. Objectless positive and negative moods have a propensity to turn into states of taking pleasure or displeasure in something. Despite this, the fact remains that, to begin with, both experiences seem to be of a hedonic or emotional state lacking an object; an entirely free-floating mood or emotion.

Are there really such entirely free-floating hedonic states? Plausible alternative explanations of the phenomena are available. It may be, for example, that apparently free-floating moods do not lack content altogether, but that their content is very
general, vague, or highly complex and hard to be put into words. Consequently, the difficulty may simply lie in articulating what it is that we affectively feel good or bad about. Another possibility is that an epistemic shortcoming stands behind the impression that one's mood is without an object. That is to say, while the hedonic state in fact has content, the agent for one reason or another, lacks access to it. A great strength of these explanations is that they can make sense of incidents in which a person becomes aware that her apparently free-floating happiness was in fact all along her being happy about something. A further virtue is that they offer us a unified picture of hedonic phenomena. All things considered, therefore, I think we should conclude that “no experience could be merely pleasant or painful as that nothing could be black or white without also having some shape and some size” (Broad [1962]: 235).  

Suppose, however, we grant the existence of entirely free-floating moods. It would follow that these hedonic states could not be characterised by an affective mode of intentionality. But it wouldn't follow that no hedonic states can. Consider the pleasure of playing a game of chess. It seems absurd to try to abstract away the pleasure from the playing of the game. As Duncker puts the point,

a product like gold emancipates itself from, and exists independently of, its source. Is the pleasure separable from the flavor in the same sense? Clearly not. The experience of the pleasure remains dependent upon the experience of the flavor (or whatever other source it may have). (Duncker [1941]: 399)

In this case, therefore, pleasure can't be seen as a non-intentional state that merely happens to accompany an intentional state as we can't conceive of it independently of that which it is a pleasure in. More generally, if we do take pleasure in something, the pleasure can't be separated from its (intentional) object. Since these pleasures are essentially tied to their object, they can't be wholly non-intentional in nature. So, even if we thought that some hedonic states are non-intentional, we should not generalise this to all hedonic states. At most, therefore, the apparent existence of free-floating moods would call for a distinction between intentional and non-intentional hedonic states. Such a distinction alone would suffice for the current objection to affective intentionality to fail.

I have argued that hedonic states, such as being happy about the blooming of the flowers, provide a test case for the thesis that affective intentionality is a fundamental mode of intentionality. These states are characterised by a distinctive evaluative ‘force’. To be happy about the blooming flowers is not just to apprehend that the flowers are blooming, but to represent this favourably or in such a way as to be for it. This, I argued can’t be reduced to pure or cognitive intentionality, not even if these are

combined with the experience of bodily sensations. Finally, I argued that even apparently free-floating hedonic states might be intentional. And even if there should be some non-intentional hedonic states, I also argued that there are at least some hedonic states that definitely can’t be regarded as a composite of an intentional state and a non-intentional state. The upshot then is this: if we wish to capture the richness of our mental lives, we must recognise both cognitive and affective intentionality. Put another way, we must recognise that just as there are ways of having a content before the mind that include an appraisal of it as being true or false, so too there are ways of having a content before the mind that include a positive or negative appraisal.

3.3 Modes of Intentionality and Functionalism

There is a complaint against the conceptual framework I am trying to develop in this chapter that I have so far chosen to set aside. The complaint is that a characterisation of mental phenomena in terms of modes of intentionality fails to get to the heart of the matter — or at least, fails to express the crucial points in the most fundamental terms. This is because the present conceptual framework encourages us to think of the nature of belief or the nature of pleasure as being determined by a mode of intentionality. To believe that \( p \) just is to apprehend \( p \) under a (distinctive) cognitive mode of intentionality; to take pleasure in \( p \) just is to apprehend \( p \) under a (distinctive) affective mode of intentionality. But according to an influential doctrine, holding a mental attitude towards a content reduces to a functional property. On this functionalist view, to believe that \( p \) just is to be in a state with the content that \( p \) and for that state to play one kind of functional role in the mental economy. To take pleasure in \( p \), on the other hand, just is for that state to play a different kind of functional role. Consequently, functionalists will say while it isn’t exactly false to say that belief is marked by a cognitive mode of intentionality or that hedonic phenomena are marked by an affective mode of intentionality, this is an unnecessarily cumbersome and potentially misleading way of getting to the key point: that these states have a distinctive functional profile. 69,70

The case for functionalism may look especially strong in the case of belief. Beliefs are states “by which we steer” (Ramsey [1978b]: 134);\(^\text{71}\) they are states that “occupy an executive office in the system’s functional organization” (Dretske [1983]: 198). This, many feel, is what belief essentially consists in, and it is natural to think that this also gives us an account of what the mental ‘force’ of belief consists in. States that lack this steering role, lack the mental ‘force’ of belief. “The difference between the belief that

69 For classic statements of functionalism, see Armstrong [1968]; Lewis [1980]; Putnam [1975].
70 A related objection could be made based on Hume’s famous claim that the distinguishing mark of belief is its ‘force and vivacity’. See Appendix II for a discussion of Hume’s attempt to capture the difference between belief and thought.
71 Ramsey is often interpreted as having held a causal analysis of belief. There is a case to be made, however, that while he thought the causal properties of belief to be paramount to the importance of belief (Ramsey [1978a] p. 50), he nonetheless accepted a phenomenal analysis of belief (e.g. Ramsey [1978a] p. 46).
modes of intentionality are determined by their rational role. on this view, belief is a state with
cognitive intentionality in virtue of the rational role belief plays in our mental lives. according to causal role functionalism,
mental attitudes and modes of intentionality are determined by their causal role. on this view, belief is a state with
cognitive intentionality in virtue of the causal role belief plays in our mental lives. i think
neither rational role functionalism nor causal role functionalism is true and that the
widespread appeal of functionalism owes much to a confusion of the two.

belief is a state that plays a distinctive rational role. for example, suppose you don’t
believe that not q nor anything that implies not q, but that you do believe that p and
also that q follows from p. suppose, too, if you like that the inference from p to q is
“immediate and obvious” (broome [2002]: 93). then it is rational for you to believe
that q. or suppose you desire that p and believe that you can bring p about by ϕ-ing.
in that case it is rational for you to have a motivation to ϕ. in contrast, entertaining a
proposition in thought lacks this rational role. suppose, for example, that you don’t
believe that not q nor anything that implies not q, but that you do believe that p and
that you entertain the proposition in thought that q follows from p. that doesn’t mean
it is rational for you to believe that q. so, the functionalist is right in saying that belief
and entertaining differ in the ‘roles they play in our mental lives and behaviour’. but
can we reduce modes of intentionality to rational role properties. surely not. doing
so would put the cart before the horse because the rational role properties of a mental
state type are determined by its mode of intentionality. belief has the rational role
it does precisely because it involves regarding a proposition as true. it is because
in believing a proposition one is committed to its truth that one is subject to the norms
of rationality. and it is because this commitment is absent when one merely entertains
a proposition in thought that this does not make one subject to the same norms of
rationality. this is crucial. there is no way we can capture the rational significance of
belief without appeal to the fact that belief involves regarding a proposition as true.
and there is no way we can explain why entertaining does not play the same role in
rationality, without recourse to the fact that entertaining a proposition in thought

72 see also, e.g. bennett [1990] p. 97; chalmers [1996] p. 19; fodor [1991]; schiffer [1981] p. 331; shoe-
(braithwaite thought that he is following the lead of alexander bain. what exactly bain thought about belief
is controversial, though. see, for example, wernham [1986]). note, i shall focus solely on the causal powers of
belief, not its causal genesis. it would be consistent with a functionalist theory of belief to regard some causal
antecedents as necessary. the problem is just that this is not particularly plausible. some beliefs are caused by
perception, others by prior beliefs. yet others, it seems, could be caused by reading a novel, by taking a pill or
by having a good night’s sleep. some beliefs may even be innate.
does not involve regarding it as true. Since the rational difference between belief and mere thought presupposes a difference in modes of intentionality we can’t reduce the latter to the former.

Next, consider causal role functionalism. According to that doctrine, there is some causal power in virtue of which an intentional state is a state of belief. Note how strong this thesis is. What is at stake is not the obvious truth that beliefs typically have certain causal powers nor that beliefs must play a certain causal role in a fully rational agent (since a fully rational agent must comply with all rational requirements that she is subject to in virtue of her beliefs). Rather, what is at stake is the thesis that there is some casual power which makes it the case that an intentional state is a state of belief.

What causal power could this be? One obviously inadequate answer would be that it is the causal power to effect a belief in what is (obviously and immediately) implied by the content of the intentional state. This is blatantly circular. To break out of this circularity, it might be said that the causal power in virtue of which an intentional state is a state of belief is the causal power to guide one’s behaviour relative to our desires. Is that, or something like it, true? We might think so if we confuse the fact that belief is the rational guide to behaviour with the claim that a state is a belief in virtue of causing behaviour. The latter claim is doubly implausible. First, it is possible for a belief to lack this causal power. Suppose, for example, that a person wants some ice-cream and believes she can get it by going to the super-market. Nonetheless, these states fail to guide her behaviour because the agent fails to rationally put them together. Yet, if an intentional state were a state of belief in virtue of the causal power to guide behaviour in concert with desire, this would have to be impossible. Second, it is possible for states other than belief to have this causal power. Suppose, for example, an ice-cream starved person is stuck in a boring meeting with no means of satisfying her desire. Her thoughts start to drift, and she daydreams about a Genie who offers to supply her with ice-cream if she orders him to do so. To the surprise of her colleagues and of herself, the person says out loud, ‘I order you to get me ice-cream’. Again, if an intentional state were a state of belief in virtue of the causal power to guide behaviour in concert with desire, this would have to be impossible.

One might try to defend causal role functionalism against these counterexamples. For instance, it might be thought that a belief that fails to display this causal power on one occasion must still exhibit it on most other occasions. Consequently, it might be said that the first example fails to refute the claim that an intentional state is a state of belief in virtue of the causal power to typically guide behaviour in concert with desire. However, the problem with this response is that it assumes that people can’t
have stable rational blindspots. It assumes that it is impossible for a person to have a belief that fails to guide her behaviour in concert with desire on most occasions. I do not see why we should think that this is impossible.\textsuperscript{73}

A related response to the first counterexample is to claim that we can avoid its force if we define the relevant causal power in more fine-grained ways. Thus, it might be said that the causal power in virtue of which an intentional state is a state of belief is the power to guide behaviour relative to desire, as well as given a background of attention, non-intoxication, and so forth. However, as I pointed out in the first chapter, unless we explicitly build in the requirement of rationality, an agent can satisfy any set of conditions yet fail to make the rationally required transition. For instance, someone can desire ice-cream, believe she can get it by going to the super-market, be consciously aware of her desire and her belief, not be under the influence of alcohol yet still not come to be motivated to go to the super-market simply because she rationally fails to put two and two together. The only way to guarantee the causal transition is by building in the requirement of rationality. Yet, requiring the agent to be rational can secure the causal transition only if there is a rational requirement to make the transition in the first place. Therefore, to the extent that a causal power is necessary for belief—namely, the causal power to guide behaviour relative to desire and rationality—it is a causal power that itself presupposes the rational properties of belief, which in turn presuppose belief’s cognitive mode of intentionality.

What about the second counterexample? It is striking that the behaviour caused by the entertaining of the Genie is somehow deficient, falling short of full-blooded intentional action. This suggests that the counterexample can be avoided by specifying that the relevant causal power is the power of a state to guide intentional action in the light of our desires and not mere behaviour. But, let us ask, what accounts for the intuition that there is something deficient in the behaviour? It is not that the behaviour lacks goal-directedness as, for example, behaviour displayed in an epileptic fit does. Rather, the behaviour appears deficient because it isn’t expressive of the person’s agency. Now, let us also ask the further question, why the behaviour isn’t expressive of the person’s

\textsuperscript{73} It is plausible to claim that an agent can have beliefs only if they typically have certain causal powers relative to her behavior. However, this will help the causal role functionalist only if she were to endorse the thesis that an intentional state is a state of belief in virtue of it being a token of a state type, other tokens of which, typically play a certain causal role in the psychology of the agent. This raises deep questions about the appropriate way of individuate the state type. Evidently, it would be viciously circular to say an intentional state is a state of belief in virtue of it being a token belief, other tokens of which, typically play a certain causal role in the psychology of the agent. A more promising proposal would be the claim that there is a certain type of brain state, tokens of which typically play a certain causal role. It might be said that an intentional state is a state of belief in virtue of it being a token of this brain state type. Yet, unless we are antecedently convinced that mental state types map neatly on to brain state types, I think we would object that something could be both a token of the appropriate brain state type and not be a belief. (On the other hand, the causal role functionalist might challenge us to provide an alternative explanation of the apparent truth that an agent can have beliefs only if they typically have certain causal powers relative to her behavior. That challenge can be met. For example, an alternative explanation would be that it follows from the independently plausible claim that only agents that are relatively rational can have beliefs.)
agency. The best answer, it would appear, is that it lacks that expressive power because there is nothing that, from the agent’s own point of view, would serve to rationalise it. After all, although the agent wants to have some ice-cream, she does not regard her speech-act as a way of furthering that desired end – she only entertained the thought that it would. So, performing the speech act does not make sense from her own point of view. But while it follows from this that belief can guide action expressive of the agent’s agency, whereas mere entertaining can’t, once again this difference presupposes rather than explains the difference in mode of intentionality between the two states.

These reflections strongly suggest that the functionalist is running two ideas together: the idea that it is constitutive of our conception of belief that belief is apt to rationalise belief and action; and the idea that a state is a belief in virtue of some causal power. Once these two ideas are clearly distinguished, functionalism loses much of its appeal as a theory of mental attitudes and modes of intentionality. To the extent that there is a clear functional difference between belief and entertaining, that difference itself presupposes a prior difference in the mode of intentionality. And to the extent that we aim to explain modes of intentionality in terms of independent causal powers, it remains dubious, at best, if there is a causal power that fits the bill.  

### 3.4 Modes of Intentionality and Consciousness

If talk of modes of intentionality isn’t just a slightly obscure way of talking about causal functional properties, what is it? The worry is that by raising doubts about the functionalist interpretation of modes of intentionality, one is effectively losing one’s grip on the notion altogether. What, in other words, do we mean by the notion of a mode of intentionality if not a causal functional property? Talk of modes of intentionality starts to sound very mysterious indeed.

---

74 As Brentano put it in a passage that has lost nothing of its pertinence:

> How come that the one conception of the object does have such an influence upon action, which the other one lacks? Merely asking this question suffices to highlight the oversight that Bain [to whom the functionalist doctrine is attributed] has committed. The special consequences would not be, if there would not exist a special reason for them in the constitution of the thinking itself. The disparity in consequences does not make the assumption of an inner difference between the mere presentation and the judgment dispensable. Quite the contrary, it emphatically points towards such an inner difference. (Brentano [1911]: Ch. 3, §2)

(By speaking of an 'inner difference', I take Brentano to refer to the difference in modes of intentionality between belief and supposition. For alternative interpretations see Russell [1921] and Braithwaite [1932-33].)
The claim I want to defend is that there are distinctive *phenomenal episodes* that anchor our grasp of the cognitive and affective mode of intentionality.\(^{75}\)

Recall the case of being happy about the blooming flowers. To be in that state is to be in a state with affective intentionality, that is, a state marked by a certain positive ‘force’ or pro-ness. But to be happy about the blooming flowers is also to be in a distinctive phenomenal state. Put another way, there is a distinctive character to what it is like to be in that state. And it is very natural to think that the positive ‘force’ of being happy about the blooming flowers consists in the distinctive phenomenal character of that hedonic state. In other words, the pro-ness of being happy about the blooming flowers may be thought to be ultimately just a matter of the way one *feels* about the blooming of the flowers.

Admittedly, this contradicts much theorising about the nature of hedonic experience. Many philosophers think there is something quite flawed in the attempt to understand pleasure as a feeling, preferring instead a desire-based analysis or an account of pleasure that treats it as *sui generis* attitude. But why should we accept that “pleasure is not a feeling” (Feldman [1997]: 463)? The standard argument is to point out that the phenomenal character of pleasure differs in kind from the phenomenal character of sensuous qualities of experience, such as the phenomenal character of seeing red, tasting something sweet, or feeling a sharp pain. Whereas we can strip away those sensuous qualities from the overall experience, the same does not apply to pleasure, as Sidgwick famously pointed out. That is to say, whereas we can look at a car and notice the same kind of reddish-quality that was present in our experience of a tomato at the supermarket, we can’t in the same way detect a pleasure-quality shared between our joy of driving the car and eating the tomato. And on this basis many

---

\(^{75}\) Two positions dominate the contemporary debate over the phenomenal character of experience, the ‘what-it-is-likeness’ of consciousness. *Representationalists* maintain that the phenomenal character of experience is entirely determined by what is being represented in experience; that the phenomenal supervenes upon content (E.g. Byrne [2001]; Chalmers [2004]; Harman [1990]; Jackson [2003]; Tye [1995]). *Phenomenalists*, in contrast, hold that the phenomenal character of experience outstrips its intentionality, that there are phenomenal elements of our experience that are wholly non-intentional (Block [1990]; Block [2003]; McGinn [1997]; Shoemaker [1982]). Against this backdrop the present thesis must sound entirely confused. On the one hand, I might be taken to contend that our conception of modes of intentionality is anchored in distinctive phenomenal episodes which themselves supervene upon the content of experience. This, of course, would obliterate the very distinction between modes of intentionality and content and thus can’t be right. On the other hand, I might be taken to suggest that our conception of modes of intentionality is anchored in phenomenal elements of our experience that are wholly non-intentional, in ‘raw feels’. But this can’t be right either. For one, modes of intentionality seem quite different from the usual candidates for ‘raw feels’ like after-images, pains and orgasms. For another, modes of intentionality are essentially linked to intentional mental states as they are manners of a content being had by a state. But we can’t explain this essential linkage if modes of intentionality were elements of experience wholly lacking an intentional nature. One possible conclusion to draw from this is that trying to elucidate modes of intentionality by reference to the phenomenal character of experience must end in disaster. But another possibility is that neither representationalism nor phenomenalism can be quite right as there are elements of our experience that are neither wholly determined by the content of experience nor utterly non-intentional. This, I think, is the proper conclusion to draw.
have concluded that hedonic experience has no distinctive phenomenal character. But this doesn’t follow. What follows is just that if there is a distinctive phenomenal character to hedonic experience, then it can’t be a sensuous quality but must be, as we might put it, a feeling tone — “a mode of consciousness distinct in nature and conditions from all sensations” (McDougall [1911]: 312).

There are good reasons to think that hedonic experience involves a distinctive feeling tone. First, there is clearly a world of a difference between what it is like to enjoy something and what it is like to find displeasure in it. These experiences differ in their phenomenal character. This raises a problem for those who want to deny that hedonic experience is marked by a distinctive feeling tone. How do they purport to explain this platitude? Their only option is to argue that the difference between taking pleasure in something taking displeasure in it is a difference in what one takes pleasure and displeasure in. Yet this is a desperate move, since it would seem that there need be no such difference in the object of pleasure and displeasure; that one and the same thing can be the object of pleasure or displeasure. If this was not so, then we would be faced with the queer consequence that pleasant or unpleasant objects are necessarily pleasant or unpleasant.

Second, a pleasant experience is an intrinsically good experience, and it is so in virtue of its pleasantness. An unpleasant experience is an intrinsically bad experience, and it is so in virtue of its unpleasantness. Indeed, this is why hedonic experiences have ethical significance. Now to make proper sense of hedonic experience being good or bad in virtue of their pleasantness or unpleasantness we must, it seems, talk about the way those experience feel — pleasant experiences are good because they feel good, and unpleasant experiences bad because they feel bad. We shall be at a loss to make sense of this if we deny that hedonic experience has a distinctive phenomenal character.

Third, the popular alternatives to feeling-based analyses of pleasure are untenable. Consider desire-based accounts, such as the theory that an experience is pleasant in virtue of a desire for it to continue and the theory that a pleasant experience is the

---

77 See also Broad [1962] p. 229 and Schlick [2002] Ch.2. Moore [1988], another eminent proponent of an account of hedonic phenomena in terms of conscious experience, was less careful, at one point even writing that “it is enough for us to know that ‘pleased’ does mean ‘having the sensation of pleasure’” (Moore [1988] p. 13). Part of the problem may lie in the common assumption of treating pleasure and pain as opposites. This, however, is a simplistic thought as careful philosophical reflection has amply shown. See, for example, McCloskey [1971]; Penelhum [1957]; Trigg [1970].
78 It might be said that we desire pleasant experiences and that this explains their ethical significance. This will not do. Pleasant experiences would be of ethical significance even if they would not be desired (e.g. even if the creature lacked the ability to desire. See below for the argument that this is possible). But note, too, that this option isn’t open for most critics of the view that pleasure is a feeling anyway, as they also deny that desire has any ethical significance (e.g. Feldman [1997]; Parfit [1997]).
upshot of satisfied desire. There are counterexamples to both claims. Against the first claim, we can point out that there can be too much of a good thing. “There are”, as Kenny reminds us

so many cases where the prolongation or repetition of what was enjoyed would ruin the enjoyment. The sweetest last to make the end most sweet might not be sweet were it not also last. Enjoying a play does not mean wishing that it had six acts instead of five, and one can enjoy the first movement of a symphony without being distressed that it is followed by the second. (Kenny [1963]: 135)

And with respect to the second proposal, on the other hand, we can ask

what of those cases in which we literally ‘run into’ a pleasure we did not know to exist where we found it? Children, for instance, ‘discover’ all sorts of pleasures which they then hold on to or seek again. Surely the pleasure could not have been caused by the success of the very desire which the pleasure itself had initiated. (Duncker [1941]: 394)

Indeed, there is also a more general problem with any analysis of pleasure in terms of desire. This is that pleasure is simpler than desire. We can unpack this intuition by noting that desire has certain logical presuppositions that pleasure lacks. Desire requires the capacity to conceive of a gap between the way the world is and how it could be. Pleasure does not. Consider a simple organism that lacked the capacity to represent past, future, or merely possible states and whose conception of things was strictly limited to its present perceptual input. It would seem possible for this organism to take pleasure in its perceptual states, but it could not possibly desire anything. Since pleasure has fewer logical presuppositions than desire, the attempt to analyse the former in terms of the latter must badly fail.

The alternative theory that is prominent amongst those who deny the phenomenal nature of hedonic phenomena is that pleasure is constituted sui generis attitude of liking. There is, I think, nothing wrong with this as long as one also maintains that our grasp of this attitude is anchored in distinctive phenomenal episodes. But since this is being denied, proponents of this approach must, it seems, have either of two things in mind. They must either think that our conception of pleasure is ultimately grounded in its being a state with a certain causal-functional profile or that there is simply nothing whatsoever that gives substance to our conception of pleasure - neither its phenomenal nature, nor its relation to desire, nor its causal-functional properties.

80 See also Gosling [1969].
81 Feldman [1997]. See also Broad [1962] and Trigg [1970].
The latter option makes pleasure objectionably mysterious. The former option seems scarcely more credible. What exactly is this functional causal property supposed to be? It can’t be the causal power to give rise to a desire for the experience to continue, but this would have seemed to be the most plausible candidate. At the very least, therefore, someone favouring this option faces the serious challenge of specifying the functional causal property that supposedly determines the nature of pleasure and hedonic states more generally.

So while it is true that the phenomenal character of hedonic experience differs from the phenomenal character of sensuous qualities of experience, we shouldn’t conclude that the pleasantness of an experience doesn’t add anything to the phenomenal character of the experience. It clearly does. And although this doesn’t prove that the affective intentionality of pleasure consists in the feeling tone of pleasant experiences, it does remove the most serious obstacle to that thesis, namely, that there just isn’t any distinctive phenomenal character to hedonic phenomena to begin with.

Whatever we end up saying about pleasure and affective intentionality, we can’t say quite the same about belief and cognitive intentionality. It would be patently false to maintain that what it is to regard a proposition as true is just a matter of the way the agent feels about the proposition since a person can believe something despite not feeling anything at all. We do not, after all, lose all our beliefs as soon as we drop into a deep dreamless sleep.

It would be foolish to deny that one can believe something without being in a distinctive conscious state. But it would be equally foolish to deny that belief can manifest itself in consciousness. You can believe Shakespeare wrote Hamlet without this manifesting itself in your consciousness, but you can also consciously apprehend that Shakespeare wrote Hamlet in a manner distinctive of belief. Just as in the case of hedonic experience, the phenomenology of conscious belief is distinct from the phenomenology of sensuous experience. We can strip-away a shared reddish quality from the experience of a red car and a red tomato, but we can’t do the same in the case of experiences of conscious belief. But, just as it would be false to conclude that there is no phenomenology to hedonic experience, it would also be erroneous to conclude that there is no distinctive phenomenal character to the experience of conscious belief. What it is like to consciously judge that Shakespeare wrote Hamlet does differ from what it is like to consciously entertain in thought the proposition that Shakespeare wrote Hamlet in thought. These experiences differ in their phenomenal character and this difference is one that is inextricably tied to the manner in which the proposition is being apprehended. They differ in their feeling tone.

This suggests that while we can’t simply equate what it is to regard a proposition as true with the feeling tone of an experience, some more complex reduction may be possible. Consider an analogy. A person can be in love for hours, days and months without
her love manifesting itself in consciousness. She can love without her experience being toned by love. It would be odd (though not untried) to infer from this that distinctive conscious experiences are somehow irrelevant or merely incidental to love. When we wonder if someone loves us we do not just wonder if they are disposed to engage in love-related behaviour or thought processes. That side is important too, but it does not even begin to exhaust what is at stake. Clearly, what is at the core of our inquiry is whether the other person feels a certain way about us – not at all times, of course, but surely sometimes or at the right times. So, while we can’t equate being in love with a feeling, it is plausible to maintain that being in love does bear an essential relation to feelings of love. And something similar may also be true in the case of belief. Here too a straightforward reduction of belief to conscious states is hopeless. All the same, belief may still stand in some complex (and no doubt difficult to explicate) relation to conscious states that is utterly essential to the nature of belief.

One reason for thinking that this is true is that it appears to offer us the best chance of making sense of what it is to regard something as true. The feeling tone of belief manifesting itself in consciousness can give substance to the idea of regarding something as true. Per contra, the most plausible alternative, causal functionalism, has been shown to be inadequate for the task. If we wish to avoid clothing the mental ‘force’ of belief in mystery, we should conclude that while we can’t equate belief with conscious belief, we also could not fully explicate what it is to regard a proposition as true without reference to the distinctive feeling tone of conscious belief. More generally, appeal to conscious experiences is not, as some would make out, a lavish extravaganza, but a much needed foundation for our conception of modes of intentionality. So, while it is fair to ask “what can be so important about the capacity to become occurrent?” (Braddon-Mitchell and Jackson [1996]: 139), the answer to this

82 Braddon-Mitchell and Jackson pose this challenge in the context of a discussion of Searle’s view that all genuine mental states are at least potentially conscious (E.g. Searle [1998] Ch. 7. It is also close to the heart of McDowell’s concerns in Mind and World. Note, that while I share Searle’s sense that the notion of an unconscious mental state has been treated in much too cavalier a fashion in contemporary thought, Searle ties his point to an argument about the fine-grained nature of mental content, whereas my point here is that there is something in the very idea of an intentional mode that requires appeal to the phenomenal). The full passage reads:

Such a view [Searle’s] must admit that most of our beliefs and desires do not have a distinctive phenomenology, yet insist that it is only because they can acquire such a phenomenology that they are genuine. It is hard to see what role the phenomenology is suppose to play. It isn’t claimed that beliefs and desires have this phenomenology at all times as part of their nature. Nor is this phenomenology required for the beliefs and desires to play their causal roles in regulating action (even internal actions such as non-conscious inference). What can be so important about the capacity to become occurrent? (Braddon-Mitchell and Jackson [1996] p. 139)

They go on to point out that if there is a belief that guides action but never rises to the surface of consciousness although it could have done so, then we have no reason to doubt that it is a genuine belief. Intended as a blow to Searle, it is, of course, a point wholly compatible with Searle’s thesis that the notion of an unconscious mental state is parasitic upon the notion of a conscious mental state and that it is only if a state is in principle accessible to consciousness that it counts as a mental state.
question should by now be clear: without conscious experience, we lose our grip on what mental ‘force’ comes to. With conscious experience, we can grasp the nature of the phenomenon.  

3.5 Cognitive and Affective Evaluation

What is the importance of all this for the nature of desire? The first two chapters showed that desire is a pro-attitude but that it isn’t a pro-attitude in virtue of an evaluative cognition. Together, this made for a perplexing result. How should we make sense of the idea of a pro-attitude if not in terms of an evaluative cognition? Distinguishing between the content of intentional mental states and the mode of intentionality allows us to provide an answer to this question as it allows us to explicate how there can be two different kinds of evaluation: cognitive evaluation and affective evaluation.

What is involved in making an evaluation? One thing that it can involve is representing an evaluative state of affairs. To judge that helping those in need is good, for example, is to make an evaluation because it involves representing an evaluative state of affairs. This answer hides an ambiguity, though. When we say someone represents \( p \) we might mean simply that she is in a state with the content that \( p \). Alternatively, we might mean that she is in a state with the content that \( p \) and the cognitive mode of intentionality. Now, being in a state with evaluative content doesn’t suffice for making an evaluation. For example, if I entertain the proposition that helping those in need is good but do not ‘go forward in judgment’, I am not making an evaluation. What is needed is not only having an evaluative content before the mind but also apprehending it as true. Call evaluation that is the result of a combination of cognitive mode and evaluative content, cognitive evaluation.

As long as we thought that intentionality was solely a matter of ‘evaluative representation’ it would have been impossible to see how there could be any other kind of evaluation. Once we recognise the importance of modes of intentionality in general, and affective intentionality in particular, this obstacle fades away because we realise that evaluation can also be a matter of apprehending something under an evaluative mode of intentionality. This kind of affective evaluation is not a combined function of (evaluative) content and cognition, but a function of a content being apprehended with an evaluative ‘force’.

---

83 Tying belief to conscious experience may have other desirable consequences as well. It may, for example, help to explain the epistemic authority with respect to our own beliefs. Interestingly, early critics of a functionalist style analysis of belief in terms of its power to guide behavior (e.g. Blake [1935]), argued that it could not provide a plausible epistemology of belief.
We are now in a position to understand how there can be *two distinct kinds of evaluation*: evaluation that involves regarding an evaluative proposition as true; and evaluation that involves regarding a proposition in an evaluative manner. This shows that it isn’t paradoxical to say that a state is both non-cognitive and evaluative in nature. And this is just the result needed to allow for a viable non-cognitivism about desire.

Much ground has been covered in this chapter but the main ideas are really quite simple. Mental states, I argued, can be characterised by intentional content and by modes of intentionality, the latter being ultimately anchored in distinctive feeling tones, that is, in the distinctive phenomenal character of some episodes of our conscious lives. At the very least we ought to recognise two fundamental kinds of intentional modes: cognitive intentionality that involves regarding something as true and affective intentionality that involves being for or against something. An exciting consequence is that there are two kinds of evaluation: cognitive evaluation that consists in regarding an evaluative proposition as true and affective evaluation that consists in apprehending a proposition in a distinctively evaluative manner. The relevance for a theory of desire is that this framework may allow for a viable form of non-cognitivism about desire, because the possibility that is now on the table is that desire, like hedonic states, is a pro-attitude precisely because it is constituted by an affective mode of intentionality. Consequently, by distinguishing between two dimensions of intentionality and two kinds of evaluation, a new option for theorising about desire has emerged – an option that explicates the pro-ness of desire by analogy to the cognitive-ness of belief. In the next chapter, I shall develop this idea in more detail.

---

84 This, as well, is an important theme in Brentano [1911]. See also, Brentano [1902].
Chapter 4

An Affective Theory of Desire

The previous chapter laid the conceptual foundations for a novel approach to desire. In this chapter I shall build upon these foundations and develop an affective theory of desire. This theory, I argue, explains how desire can be a pro-attitude while avoiding the assumption that desire is essentially cognitive.

Talk of an affective theory of desire perhaps most naturally calls to mind an account that centres upon the kind of hedonic affect that we encountered in chapter three. In the first section, I shall therefore consider whether we should develop the affective theory of desire in hedonic terms. My verdict will be that we should not. Hedonic states are paradigmatic affective states, but the best affective theory of desire is not a hedonic theory. Instead, the best affective theory explicates desire in terms of felt need. In the second section, I shall develop the felt need theory of desire in some detail. This will be the heart of the positive project of this thesis, where I shall establish what I regard as the best theory of desire. This affective theory has a number of virtues. Some of these will be reviewed in the third section, while others follow in the final two chapters. Nonetheless, there are also well-known objections to theories of desire that put conscious states centre stage. In the last section, I shall make a start at addressing these objections, and show that they fail to undermine the affective theory of desire.

4.1 Hedonic Theories of Desire

In a famous passage, John Stuart Mill wrote, “desiring a thing and finding it pleasant [...] are phenomena entirely inseparable, or rather two parts of the same phenomenon; in strictness of language, two different modes of naming the same psychological fact” (Mill [1991]: 173). Taken literally, this is a non-starter. Suppose I want to eat ice-cream. At the time of my desire, however, I am not eating ice-cream and, therefore, can hardly find doing so pleasant. There are two modifications that make the suggestion a great deal more plausible. (i) It can be claimed that for an agent to desire that \( p \) it is necessary and sufficient that she would find \( p \) pleasant were it to occur. (ii) It can be claimed that for an agent to desire that \( p \) it is necessary and sufficient that she takes pleasure in the thought that \( p \). These ideas form the core of pleasure-based hedonic theories of desire. Such theories have had considerable influence and they deserve to be considered more closely.  

To begin with, then, let us consider whether it is necessary for an agent to desire something that she would find it pleasant were it to occur. This can’t be correct. It is possible for a desire to be satisfied without the agent’s knowing that the desired end has come to obtain. Under these circumstances, the agent could not possibly take pleasure in the obtaining of the desired end. To get around this problem we might say that for an agent to desire that \( p \) it is necessary that she would take pleasure in \( p \) if she realised that \( p \) obtains. However, this claim too faces a problem: there are plenty of cases in which people know their desire to be satisfied yet fail to feel satisfied. For example, bringing about the desired end can leave an agent too exhausted to take pleasure in it; instead of enjoying the desired end, the agent can experience negative emotions, such as guilt, shame, or fear; some desired ends prove disappointing, and so on and so forth.

That one would take pleasure in \( p \) if \( p \) were to obtain is not necessary for desiring that \( p \). But is it sufficient? We can recall two points made in the previous chapter to show that this can’t be the case either. First, we can ‘run into’ pleasures, so to speak. We can enjoy an activity or a scent, for example, without any prior desire. It follows that, although at some earlier stage it was already true that we would enjoy the activity if we engaged in it, or that we would take pleasure in the scent if we smelled it, it wasn’t true, at that point, that we desired to engage in the activity, or that we desired to smell the scent. Consequently, the fact that we would enjoy something can’t be sufficient for desiring it. Second, desire has certain presuppositions that pleasure lacks. I have argued that desire requires the capacity to conceive of a gap between the way the world is and how it could be, whereas pleasure does not. An organism whose conception of things was strictly limited to its present perceptual input could take pleasure in its perceptions, but could not desire anything. It follows that it can’t be sufficient for desiring that \( p \) that one would take pleasure in \( p \) were \( p \) to obtain.

Let us turn to the second pleasure-based theory of desire, beginning with the thesis that taking pleasure in the thought that \( p \) is necessary for the desire that \( p \). Standing desires are a putative counterexample to this claim. That is, it seems natural to suppose that we can desire something without thinking of it at all, and thus, a fortiori, without taking pleasure in the thought of it. Christoph Fehige puts the point with characteristic wit:

Suppose that I’m a loving father and a plumber, and that right now I’m concentrating on repairing a dripping tap and hence not thinking of my children. Our theory should not force us to conclude that my children are, at this moment, not dear to me [or that I don’t desire for them to do well]. My children may well be dear to me all the time – I just can’t think of them all the time. Not for instance, while repairing a dripping tap in order to be able to afford their college fees. (Fehige [2001]: 50)
The challenge posed by standing desires is not unique to pleasure-based theories. Indeed, any theory of desire that puts conscious states centre stage, including the felt need theory I favour, will need to address the phenomenon. For this reason I shall defer discussion of this challenge until the next section, where I shall show how affective theories can address it.

A different objection is that if one thinks it unlikely that one will get what one wants, or sees a great amount of uncertainty in it, focusing one’s mind on the desired end will often result in displeasure, anxiety and stress. My present thought of water, as I sit on my desk with a glass of water easily at hand, is neither pleasant nor unpleasant, but the lost traveller who is dying of thirst will find the thought of water extremely unpleasant precisely because she so dearly wants a quenching drink and does not know how to get it. Or suppose it is the eve of the election and I am running for president. It has been a tight race; I desperately want to win and am nervously awaiting the results in the midst of my loyal supporters. In this case, my mind is very much focused on becoming president, not on repairing pipes, but saying that I am taking pleasure in the thought could hardly be further from the truth. It is difficult to avoid the conclusion that the Buddhist conception of desire as a cause of distress and sorrow contains at least a grain of truth.86

Taking pleasure in the thought of something is also not sufficient for desire. Consider Mark Johnston’s example:

Suppose that I feel pleased at the thought of my working in the local soup kitchen. Even so, I need not find my working in the soup kitchen appealing. I might only find the thought of it appealing. I could be a self-indulgent dreamer, one who particularly enjoys the costless contemplation of himself in a good moral light. I might be pleased at the thought of my working in the soup kitchen while knowing that I would not find my working there appealing. (Johnston [2001b]: 225)

**Mutatis mutandis** for having a desire to work in the soup kitchen.

Consider another case. Suppose I am reading a work of fiction and take pleasure in the description of a particular scene. Does this mean that I must have some desire for this scene to exist in actuality? Surely not. These two examples point to a more general problem with the sufficiency thesis: if you desire something, you want it to be the case. You do not just entertain it, but you entertain it in a way that includes a kind of demand that it should come to obtain. Not so in the case of pleasant thoughts. Our

86 V. Ehrenfels claimed that although desire need not be pleasant, it needs to make the thought of the desired end more pleasant (or less unpleasant) than it otherwise would have been (v. Ehrenfels [1982] §9, §79.). As the example of the traveler dying of thirst shows, this can’t be correct. The traveler’s desire surely accentuates the distress, rather than alleviates it.
pleasant thoughts can exemplify a contentment with the way things are that wanting
something to be the case must necessarily lack. Indeed, it seems conceptually possible
that a creature has the capacity to entertain thoughts, but lacks a conception of reality,
of something actually existing. This creature could take pleasure in the thoughts that
are before her mind, but it could not want anything to be the case.

Pleasure-based theories have a great deal of initial appeal. In part this is because it
is easy to conflate the fact that pleasure frequently accompanies desire (e.g. I enjoy a
game and desire to make my next move) with the idea that pleasure is necessary and
sufficient. Moreover, we usually think (though sometimes erroneously) that satisfying
our desires will bring pleasure. But most of the attraction of pleasure-based theories
derives from the stubborn thought that the connection between desire and pleasure is
more than merely contingent. That some of us desire to eat brussel sprouts is entirely
contingent. That we desire pleasure isn’t. Nor does feeling satisfied when one knows
one’s desire to be satisfied have the flavour of a wholly contingent relation. While this
seems right, we should be careful not to jump to the conclusion that the only explana-
tion of this consists in pleasure being necessary and sufficient for desire. Other facts
about these states, e.g. that they are both positive affective states, may explain why
there are connections between these states that appear to be conceptual in nature.

Historically, pleasure-based theories have been the most significant amongst hedo-
nic theories. Another central class of hedonic theories, however, are those that focus
upon displeasure. John Locke notoriously advocated such a theory. “Who is there”,
Locke asks,

that has not felt in desire what the wise man says of hope, (which is not
much different from it), that it being “deferred makes the heart sick”; and
that still proportionable to the greatness of the desire, which sometimes
raises the uneasiness to that pitch, that it makes people cry out, “Give
me children.” Give me the thing desired, “or I die.” Life itself, and all its
enjoyments, is a burden that can’t be borne under the lasting and unremoved
pressure of such an uneasiness. (Locke [1975]: Book 2, Chapter 21, §32)

His position culminates in the conclusion that desire is “nothing but an uneasiness in
the want of an absent good” (Locke [1975]: Book 2, Chapter 21, §31).

There is, I think, something theoretically problematic about the attempt to under-
stand desire in terms of displeasure, which is that desire is a pro-attitude, whereas
displeasure is a con-attitude. Now it might be said that a con-attitude towards a propo-
sition is the same as a pro-attitude towards its negation, but this view does not stand
up to scrutiny. Taking pleasure in something is not the same as taking displeasure in
its absence or negation, and these states do not necessitate each other. I can enjoy a
movie without finding not watching the movie displeasurable. I can take displeasure
in the thought of an accident without taking pleasure in the thought of there being no accident. So, a displeasure account does not seem capable of explaining the pro-attitudinal nature of desire.

Nonetheless, let us briefly consider the following position: it is necessary and sufficient for the desire that \( p \) that one is displeased by not \( p \). First, consider the necessity thesis. That thesis is untenable. I want to visit Madagascar. I want to win the game. Right now I am not in Madagascar and haven't won the game. There need be no uneasiness or distress in either of these experiences. It is true that we are often distressed by the world thwarting our innermost desires, but so long as we are reasonably confident that we will be able to satisfy them, we tend to carry our lot with ease and without displeasure.

Next, consider the sufficiency thesis. This doesn't seem to be quite right either. For example, there seem to be cases in which one is so engulfed by one's displeasure that one forgets to desire for things to change, so to speak. Severely depressed people offer an extreme illustration. It is part of their condition that they find most things unpleasant. At the same time, they appear to lack any desire for things to change. They are so focused on their misery that they don't conceive of, let alone desire, any alternative.

We can also fault the sufficiency thesis on a more abstract level. One problem is that desire requires a conception of the way the world actually is and a conception of how it could be. As noted before, hedonic states do not have equally complex presuppositions. Creatures that lack a conception of reality, or that can't form a conception of how things could be, can still take pleasure or displeasure in what is before their mind.

In response, it might be said that the displeasure central to desire is displeasure in things not being a certain way. Clearly, no creature that lacks a conception of the way the world actually is, or of how it could be, could take displeasure in \( that \). Even so, I think a related problem remains. I have also argued that desire involves a certain dynamic dimension such that one entertains the desired end in a way that includes a need for it to become reality. In desiring something one does more than conceive of something that is not, but could be, the case. One also conceives of it as having to become the case. This dimension can be absent even if one takes displeasure in things not being a certain way. As the example of depressed people highlights, people can be stuck in their displeasure at things not being a certain way without conceiving of things as having to be different, without wanting anything.

Again, this is not to downplay the deep connections between negative hedonic affects and desire. In particular, being displeased by the world and the way it afflicts one is a powerful source of desire and one that does not seem to be entirely contingent in nature. Similarly, there does seem to be more than a merely contingent connection between having one's desires thwarted and feeling frustrated and displeased.
Nonetheless, we should avoid the temptation to analyse desire in terms of displeasure. Hedonic theories are a natural and important class of the affective approach to desire, but ultimately they are not convincing.

4.2 Desire and Felt Need

How else might we understand an affective theory of desire if not in terms of hedonic states? Consider again Locke's powerful example of Rachel who yearns to have children. In her cry 'give me children or else I die,' she does not just express her frustration with a world that continually stymies the satisfaction of her innermost desire. She also gives expression to her felt need for the world to be a certain way. She feels that she must have children. This aspect is missing from hedonic theories of desire. Pleasure and displeasure are evaluative states, but they do not essentially embody the felt need for something to be a certain way. And it is this experiential state of felt need that should be at the centre of the best affective theory of desire.

Before going further, let me clarify the idea of a felt need.

(i) Felt needs are affective states. To have a felt need for something is to apprehend it under an affective mode of intentionality, that is, to apprehend it in a manner that includes a distinctive evaluative 'force', a distinctive manner of being for it.

(ii) We are creatures with biological needs, and our desires are often linked with our biological needs. In fact, it is presumably because of this link with biological needs that we have evolved to have desires in the first place. But despite these important connections between felt needs and biological needs, a felt need is not a biological need. A felt need is an experience, a biological need is not. Moreover, to have a biological need for x implies that the organism will be seriously harmed, unless x is secured in a timely fashion. Having a felt need does not imply any such thing.

(iii) It is extremely important to note that felt needs can differ in intensity. Often the most natural expression of a felt need will be "this must become reality", but at other times this exclamation will have too much of a sense of urgency to capture the phenomenal character of the experience, in which case the experience may be better expressed by saying "would only that this state become reality". While it will often be useful to illustrate points by reference to cases in which it is apt to say that the agent feels that the desired end must become reality, this is not itself part of what I mean by speaking of felt needs. As long as the way in which a state is before the mind includes some, however weak, felt force for it to become reality, it will be a felt need.

(iv) Felt needs often can be given linguistic expression. For example, 'this must become reality' is a natural linguistic expression of paradigmatic felt needs. However, felt needs are not linguistic phenomena.
A felt need is not a prediction of some kind. To give expression to a felt need by saying 'this must become reality' is not to imply that being in that state involves some conviction that the phenomenally needed thing will become reality. We come closer to the truth if we say it involves regarding the needed thing as something that must come about in the sense that it ought to come about. But while there is certainly some kind of affinity between felt needs and normative claims, it would also be a mistake to think that feeling that something must become reality implies believing that it ought to obtain. In the end, what is essential to felt needs is that they are a distinctive kind of conscious experience, an experience that is marked by a certain strain, an experience that something must come to obtain.

What exactly is the relation between felt need and desire? Consider this thesis: a subject S desires that p if and only if she is experiencing a felt need for p. This thesis embodies a grain of truth: a felt need for p is sufficient for the desire that p. If someone has a felt need, she must have a desire. Clearly, it need not be her sole desire nor her strongest desire. As we shall see below, she may also fail to know that she has that desire. Moreover, she may even misidentify the object of her felt need and, consequently, the object of her desire. But it does not seem possible that there could be a person who experiences a felt need, and yet does not have a desire which is manifested in that conscious episode. The suggestion that someone could have a felt need for something, that she feels that something must become reality – not in the sense that something will necessarily occur, or that she believes she has a biological need of some kind, or that she believes she is under an obligation to do something, but in the relevant sense that refers to that distinctive conscious state – yet lack a desire for the object of her felt need is, I submit, simply unintelligible.

But is experiencing a felt need also necessary for desire? To answer this question we must return to the matter of standing desires. Those desires, I said above, threaten to undermine the claim that any conscious state is necessary for desire, a point forcefully advanced by Michael Smith:

Consider [...] a long term desire: say, a father's desire that his children do well. A father may actually feel the prick of this desire from time to time, in moments of reflection on their vulnerability, say. But such occasions are not the norm. Yet we certainly wouldn't ordinarily think that he loses this desire during those periods when he lacks such feelings. (Smith [1994]: 108-109)

Put another way, a person can desire that p whilst not experiencing the felt need for p. It follows that experiencing a felt need can't be necessary for desire.
If there are standing desires then we can’t maintain that a subject S desires that \( p \) if and only if she is experiencing a felt need for \( p \). However, some people deny the existence of standing desires. This is not as outrageous as it might seem at first, since one can attempt to explain away the impression of standing desires by appeal to a disposition to desire. This point has recently been made by Graham Oddie. His argument deserves to be quoted in full:

The evidence for subconscious desires is not entirely unambiguous. You ask me whether I desire that my children’s lives go well. I reply: Yes, naturally. You ask, did I have the desire fifteen seconds ago, or did it spring into existence when you asked me about it? Again, it might be fairly natural to think that I had the desire fifteen seconds ago, and that I was not at that stage conscious of it. But do I desire that my children’s lives go well when I am in a deep dreamless sleep? Or suppose I am run over by a bus, and go into a coma, a fairly serious although not irreversible coma – a coma of the sort that I have a chance of emerging from psychologically intact. In that comatose state do I still desire that my children’s lives go well? It seems rather a stretch to say that I really did desire that, while totally comatose. It seems much more plausible to say that when I am in a coma or asleep, what I have is some disposition to desire, a disposition that will be triggered under suitable conditions. If my character and personality have a chance of surviving through the coma, then there must be something like that disposition present there as well. So I have that disposition when in the coma, and when I am asleep. Don’t I also have the disposition when I am not attending to the matter at all? Presumably. Now, since I have that standing disposition, do we really need to add that (when I am awake but not attending to the matter) I also have the fully-fledged desire? That seems otiose. The disposition to so desire when the matter is brought to my attention is sufficient to explain the phenomena. (Oddie [2005]: 55-56)

I think we should agree with Oddie that people who are in a ‘fairly serious coma’ do not have any desires. Instead, they are disposed to acquire desires under suitable conditions. However, I fail to see why this should show that when one is not in a coma but, say, merely focused upon repairing the dripping tap, one can’t actually desire that one’s children are doing well but must at most be disposed to acquire this desire under suitable conditions.

Now it might be said that Oddie’s argument is that we should always refrain from attributing a standing desire in addition to a disposition to acquire a desire because doing so is ‘otiose’, that is, without any explanatory pay-off. However, this argument

---

87 Note, in section 2.7, for example, Oddie clearly endorses the claim that desire is a conscious state. However, in section 3.3, he does appear to be willing to allow that there ‘maybe’ are some desires that “have a life below the threshold of conscious experience” (Oddie [2005] p. 55).
would work only if it could be shown that in all purported cases of standing desires people actually have a disposition to acquire a desire. Yet, whether this is the case is precisely what is at stake. Clearly, purported cases of standing desires involve a disposition of some kind, namely, a disposition to undergo certain conscious states under suitable conditions. But what needs to be shown is that this disposition is always best characterised as a disposition to acquire a desire and not—as would often seem more natural to say—as a disposition that makes up part of what it is for the agent to have a (standing) desire.

Perhaps, however, the thought is this: being in a ‘fairly serious coma’ is like being in a ‘deep dreamless sleep’ in so far as in both cases one is unconscious. And both are like putative cases of standing desires in so far as one is not conscious of the relevant end. Therefore, one can’t have a standing desire for \( p \) but at most be disposed to acquire this desire, if one is in a ‘deep dreamless sleep’ or doesn’t think of the matter at all. However, the argument goes through only if we assume that the impossibility to desire that \( p \) when being in a coma is due solely to the fact that being in a coma makes it impossible to consciously think of \( p \). This assumption, however, can be reasonably denied because a ‘fairly serious coma’ is a condition that differs in many important respects from the condition of normal human agency, each of which may explain our reluctance to ascribe desires to people who are in a ‘fairly serious coma’. This finds further support in the observation that it might also be said that if one is in a ‘fairly serious coma’ one does not believe anything but is only disposed to acquire various beliefs. Yet, it seems fallacious to infer that, therefore, there can be no standing beliefs.

Another supporting argument is also unpersuasive. Suppose a person is in a coma but would undergo an episode of desirous consciousness if she were to think of \( p \). Since people who are in a coma do not have desires, it follows that being disposed to undergo an episode of desirous consciousness is not sufficient for having a desire (a conclusion I shall attempt to corroborate further below). Moreover, since purported cases of standing desire are cases in which people are disposed to undergo an episode of desirous consciousness, and since having this disposition is not sufficient for having a desire, these cases can’t in fact involve desire. Rather, they must only involve a disposition to acquire a desire. Again, this argument fails because it assumes that the sole reason for ascribing a standing desire must be that the person is disposed to undergo an episode of desirous consciousness under suitable conditions. This, however, is not the sole reason. Indeed, as I shall argue below, part of what justifies the ascription of a standing desire is that a person does have experiences of desirous consciousness from time to time. So, even if being disposed to undergo episodes of desirous consciousness should not suffice for having a desire, it is possible that in some cases in which one is disposed to undergo an episode of desirous consciousness this involves having a (standing) desire rather than merely the disposition to acquire a desire.
Oddie is certainly right in claiming that the case for standing desires is not ‘entirely unambiguous’. Still, on balance, I think that our intuitions are in favour of there being standing desires, Oddie’s claims notwithstanding. Consequently, I think, we should reject the thesis that an agent S desires that $p$ if and only if she is experiencing a felt need for $p$. However, it might be argued that we can reconcile the case of standing desires with the intuition that felt need is essential to desire by maintaining that a subject S desires that $p$ if and only if she is disposed to experience a felt need for $p$.

This is an important thesis. Indeed, it may well be true that being disposed to experience a felt need for $p$ is necessary for desiring that $p$. However, I doubt that being disposed to experience a felt need for $p$ is also sufficient for desiring that $p$. Suppose Chris is a dairy farmer in the Alps, living in a tight-knit community. The daily life revolves around tending to the cows and spending time with family and friends. Now suppose that Chris has in fact long been disposed to experience a strongly felt need to sail the oceans – it’s just that the conditions under which that disposition would be manifested haven’t come up yet. For example, if he were to watch a sailing documentary, he would become enthralled and think he must learn how to sail. He would also experience the need to go sailing if a friend were to recount her adventures or if he considered the option of sailing the seas. Alas, whenever the local TV station is running a spot on sailing, Chris is out milking the cows and while many topics feature prominently in the conversations he enjoys within his social circle, sailing ain’t one of them. As a result, Chris goes through life without ever giving any thought to sailing. Under these circumstances, I believe, it is highly implausible to say Chris desires to go sailing and much more plausible to say he is disposed to desire to go sailing. The reason for this is clear: desire need not dominate our mental lives, but it does need to have some kind of actual effect. A mere disposition of some kind, never actualised and orthogonal to all concerns and interests that have been displayed in the subject’s life, seems too insubstantial to capture our intuitions about desire.

The case of standing desires shows that it can’t be necessary for desire that one experiences a felt need throughout the entire time that the desire exists. But the possibility of standing desires is also consistent with the thesis that one can desire that $p$ only if one experiences a felt need for $p$ from time to time. To illustrate this, consider once again the analogy with love. There are distinctive conscious episodes through which love manifests itself; we can feel love for someone. Moreover, while feeling love for someone is to love, one can love someone without feeling love throughout. Nonetheless, feeling love is essential to love. No-one could truly love another person without from time to time feeling love for the other person. The same, I contend, applies to desire. Felt need is the manifestation of desire in consciousness, and to be in a state of desirous consciousness is to have a desire. At the same time, one can desire without experiencing a felt need throughout. Nonetheless, felt need is utterly essential to desire. No-one could desire without from time to time experiencing a felt need for the desired end.
It might be thought that this simply can't be correct. Suppose from $t_1$ to $t_2$ David has a standing desire to become king. Since it is a standing desire, David does not need to experience a felt need throughout $t_1$ to $t_2$. So assume that he only experiences a felt need at $t_2$. But if this is so, then he already has a desire at $t_1$ even though he experiences no felt need at $t_1$. Consequently, it can't be true that one can have a desire only if from time to time one experiences a felt need – after all, the desire had already sprung into existence prior to any experience of felt need.

This assumes that it is possible to have a desire prior to any experience of felt need. Paradigmatic cases of standing desires do not bear this out. They involve a person continuing to have a desire subsequent to an episode of desirous consciousness. Therefore, it is one thing to say David could have continued to desire to be king once his longing thoughts had given way to concentrating on his encounter with Goliath. It is another thing entirely to say David could have desired to be king prior to ever experiencing any felt need for this to be the case.

Still, it might be said that there are also cases in which we count a person as having a desire prior to any experience of felt need. Does this mean that the objection is successful after all? Only if we assume that we would have also counted the person as having a desire even if the initial state was never followed by an episode of desirous consciousness. In some cases we might wish to count certain events as the early onset of love even if these early events unfold without the person experiencing feelings of love. However, unless these early phases are eventually followed by feelings of love there would be no justification for describing them as states of love at all, rather than, say, as mere dispositions to fall in love. Likewise, I wish to claim that even though there might be some cases in which, looking back, we describe a person as desiring something prior to any experiential manifestation of this desire, this is appropriate only if this early state is followed by episodes of desirous consciousness. Without any such link to felt need, it seems there is no justification for describing the earlier state as a genuine desire rather than a state of being disposed to desire.

The claim that the appropriateness of counting something as a desire at some point in time might depend upon subsequent events will be distasteful to some and mysterious to others. Notwithstanding this, I think this indeed forms part of our thinking about desire. More generally, I think we do frequently interpret earlier events through the lens of later events. Thus, we may describe the writing of a few sentences as the beginning of a popular masterpiece, even though this description is clearly dependent upon events unfolding in a certain manner. Likewise, I believe, we sometimes count certain states as instances of a particular mental state only because of their relation to later states.
It is a further implication of my account of desire that it is not always a fully determinate issue whether a state does stand in a close enough relation to experiences of felt need for it to be a state of desire. That is, it is no fully determinate question what counts as an agent from time to time experiencing felt need. Again, some will find such indeterminacy objectionable in a theory of desire. On the contrary, however, I would say that this is a strength of the theory. It is sometimes indeterminate whether a desire has already come to exist (or ceased to exist). It is a virtue of a theory that it does not impose artificially strict boundaries in this regard.

Still, as it stands, the claim that one can have a desire only if from time to time one experiences a felt need is too simple. *Instrumental* desires – desires that we have in virtue of a non-instrumental (intrinsic) desire and the belief that the instrumentally desired state promotes the non-instrumentally desired state – can exist as pure dispositions. Suppose Rachel has a non-instrumental desire that her children do not die in childhood. Even if she can’t have this desire unless from time to time she experiences a felt need for her children not to die in childhood, she surely can have the instrumental desire that her children do not die in childhood by getting eaten by a camel even if she never has and never will bring this possibility to mind (and *ipso facto* even if she never has, and never will, experience any felt need for this not to happen).

Instrumental desires are parasitic upon non-instrumental desires. They owe their existence to the inferential integration of our non-instrumental desires with our beliefs. Respecting the fact that instrumental desires can exist without ever being brought to mind, therefore, does not commit us to the view that any desire can be had in the absence any experience of felt need. Indeed, the case of instrumental desires does nothing to undermine the more general thesis that desire is, necessarily, a state individuated by its relation to episodes of felt need – even if, in the case of instrumental desire, the relevant relation is one that is inferentially mediated.

In part, my defence of the claim that one can have a non-instrumental desire if and only if from time to time one experiences a felt need rests upon its intuitive appeal. It is, I think, extremely intuitive to think that if a person experiences a felt need then she has a desire. Furthermore, it seems plausible to think that where felt need is not just momentarily lacking but entirely absent, desire must be absent too. Desire without any relation to felt need would be like Hamlet without the prince; the crucial ingredient is missing. A parent’s desire for the welfare of her child need not occupy her consciousness at all times, but if the parent never experiences the felt need for the child’s life to go well, what reason is there to think the parent desires this, rather than, say, is merely disposed to desire this or just believes it would be good if the child’s life goes well? However, there is also another reason for accepting the necessity claim. Making felt need essential to desire promises substantial theoretical benefits, including being able to make sense of the idea of desire as a pro-attitude and, consequently, an explanation of the role of desire in rationality. This does not make the necessity claim immune
from all counter-examples, but its explanatory usefulness does create a presumption in favour of the necessity thesis when it comes to resolving those indeterminate cases lying at the outskirts of our conception of desire where intuitions may reasonably diverge.

We are finally in a position to state the felt need theory of desire: a subject S (non-instrumentally) desires that p if and only if from time to time S has a felt need for p to become reality. This, I contend, is the best affective theory of desire.

4.3 Benefits of the Felt Need Theory of Desire
In the introduction I said that a successful theory of desire must unify our pre-reflective conception of desire. That is, it must capture what underlies our intuitions about specific cases. Now, according to the felt need theory we judge that a person has a desire because we judge that from time to time the person experiences a felt need for something to become reality. On the other hand, the theory implies that our judgment that someone lacks a desire for something is based on the judgment that the person altogether lacks any felt need for it. The considerations surveyed in the previous section appear to support these claims. The felt need theory arguably really does make explicit the structure of our pre-reflective intuitions about specific cases.

Indeed, in this context it is well worth to recall that, in chapter one, I argued that N-desires refute the causal engine thesis, but I also accepted that the refutation is successful only if we can show that there is an underlying unity between N-desires and other desires. The felt need theory passes this test with flying colours. Whether one N-desires a state of affairs or has a more ordinary desire, both states are characterised by the experience that the desired state of affairs needs to come to obtain. Somewhat metaphorically, we may say in both cases desire aims for the desired end to become reality, even though in the case of N-desires this can’t become the person’s aim.

In the introduction I also said that a successful theory of desire must respect our intuitions about the kind of state desire is. Again, the felt need theory scores strongly. For example, the felt need theory coheres well with the intuition that desire is a state that can be unconscious yet that there are also genuine episodes of desirous consciousness. Moreover, it seems to be consistent with the plausible view that we are in a privileged epistemic position with respect to our own desires but also that judgments about our own desires aren’t infallible. And the felt need theory also allows for the possibility of recalcitrant desires and it is consistent with there being desirous creatures that have very limited cognitive capacities.

What about instrumental desires? We might say that a subject S has an instrumental desire that p if and only if she has a non-instrumental desire that q, believes that p promotes q, and rationally combines these two states.
Importantly, the felt need theory, in contrast to cognitive theories of desire, also respects the deep-seated intuition that desire is not subject to the norm of truth. A related point is that the felt need theory fits nicely with the wide-spread intuition with the widespread intuition that, just as belief aims at truth, so desire aims at satisfaction: as the intuition is commonly expressed, desire is a state with which the world must fit. Common attempts to analyse what this means in terms of motivation or normative truths have arguably failed. But the felt need theory offers an alternative account. Desire involves the felt need for the desired thing to become reality. It is because of this ‘force’ of desire that it seems so fitting to say that desire aims at satisfaction or that desire is a state with which the world must fit.

However, perhaps the most important pay-off that springs from the felt need theory of desire is that it offers a credible explanation of the role of desire in rationality. To explain how desiring an end can make it rational to be motivated to act in ways one believes will promote the end, we must explain the pro-attitudinal nature of desire. One possible explanation of this is in terms of cognitive evaluation, but in chapter two I argued that cognitive accounts are subject to serious objections. The other possible explanation is in terms of affective evaluation. According to this account, desire is a pro-attitude because it involves an affective mode of intentionality. The felt need theory offers such an account of desire because to have a felt need for something, to feel that it must become reality, is to apprehend it in a manner that includes being for it. Consequently, the felt need theory *qua* affective theory of desire offers an account of the pro-attitudinal nature of desire and, consequently, an explanation of its role in rationality.

This account of the pro-attitudinal nature of desire fits our intuitions remarkably well. To illustrate, consider the desire to become president. This desire includes a positive appraisal of becoming president, a way of finding it appealing or being attracted by that option. Reflection suggests that this pro-ness of desire just consists in a distinctive feeling tone. More precisely, reflection suggests that this pro-ness of desire consists in an experience wherein it seems that one *must* become president, that this is how the world needs to be. Consequently, felt need appears to be well suited to explain the intuition that desire is a pro-attitude. Accordingly, the felt need theory offers a convincing affective and non-cognitive account of this crucial dimension of desire.

---


90 As noted above, sometimes it will better capture the phenomenal character of the experience if one says that one has an experience wherein it seems to one that would only that one be a president.

91 One objection to making conscious states necessary to desire is that doing so has no theoretical pay-off. (Braddon-Mitchell and Jackson [1996] pp. 138-140. See also the general discussion of this point in the previous chapter). This criticism, I contend, is based on a false premise. No adequate explanation of the pro-attitudinal nature of desire can go past the distinctive nature of desirous consciousness.
An Affective Theory of Desire

This is an impressive catalogue of strengths. In due course we shall add two more to the list: that the affective theory can account for the role of desire in the explanation of, and our (normative) reasons for action. First, however, let us turn to consider some objections.

4.4 Objections to the Felt Need Theory of Desire

It is frequently said that "desire *per se* has no special feel or phenomenology" (Braddon-Mitchell and Jackson [1996]: 123). One thing this might mean is that one can have a desire whilst being unconscious of it. However, this is, of course, entirely consistent with the felt need theory. On the other hand, it might also mean that desirous consciousness isn't necessary for desire at all. This clearly contradicts the felt need theory. But are there any good reasons for thinking that in this more radical sense 'desire has no special feel or phenomenology'? In the following I shall discuss four arguments aiming to establish this.

4.4.1 Is There a Phenomenology of Desire?

The dominant view in modern analytic philosophy is that desire can be caused by phenomenal states, such as pangs of hunger. Desire can also be the cause of phenomenal states, such as suffering if the fulfilment of the desire is continually thwarted. But it is assumed that there is nothing that can reasonably be labelled the phenomenology of desire itself. There are conscious states contingently associated with desire, but there is no genuinely desirous consciousness, no way for desire itself (or a constitutive part of desire) to be part of our conscious experience.93

I think this influential position gains much of its attractions from two shaky assumptions. First, it is observed that there can be standing desires; that desiring something does not imply being in an experiential state. And from this it is inferred that there is no phenomenology of desire. However, this reasoning seems flawed. Clearly, it is possible to believe something without being in an experiential state. Still, as I pointed out in the previous chapter, it nonetheless seems very plausible to think that there is a distinctive phenomenal character to episodes of consciously regarding a proposition as true. Likewise, one can love another person without being in an experiential state. But it certainly seems as if there is a phenomenology of love; that there are conscious episodes in which love manifests itself in consciousness. The same seems true in the case of desire as well. Even though one can desire without being in an experiential state there nonetheless seem to be genuine episodes of desirous consciousness.

The other assumption that is underpinning the influence of the view that there is no phenomenology of desire is, I think, the supposition that if there is a phenomenology of desire, then this must consist in the experience of a bodily sensation. But, as is often pointed out, it is really quite implausible to think that any bodily sensation should be thought of as the phenomenology of desire. However, this line of reasoning falters once we realise that the phenomenal character of desire may be the phenomenal character of a conscious state inextricably linked to an affective mode of intentionality, rather than the phenomenal character of an experience of a bodily sensation. Consequently, we may concur that there is no bodily sensation that can be thought of as constituting the phenomenology of desire yet insist that nonetheless there is such a thing as the phenomenology of desire.

However, this response must appear unconvincing if one is sympathetic to an analysis of felt need in terms of bodily sensations, such as muscle tension. Let me say why I think this is mistaken. The most obvious problem is this: it seems possible to have a felt need for something and not experience any bodily sensations at all. A fully paralysed person who has lost all introspective awareness of her body provides a vivid illustration, as it seems that this person can still experience the felt need for, say, a glass of water. Second, no bodily sensation is sufficient for the experience of a felt need. One reason for this is that a felt need is directed towards states of affairs. You feel that you must have some water, that you must have children, or must become president. Bodily sensations, on the other hand, are not similarly directed towards states of affairs. The tensing up of your muscles may be the object of your introspective awareness and may cause you to act so as to bring about the state of affairs you feel is needed, but the tensing of your muscles is not directed towards anything. So, no matter what bodily sensations are at stake, they do not add up to a felt need, as they lack the direction towards an object characteristic of felt need. Third, the experience of felt need seems to capture the pro-attitudinal nature of desire; it seems to capture what the evaluative endorsement of desire comes to. The suggestion that the experience of felt need is the experience of body tension makes nonsense of this. Finally, feeling that one must have something is sufficient for desiring that thing. It would be mysterious how this could be so if felt need were constituted by bodily sensations like muscle tension. All told, therefore, we should reject the attempted reduction of felt need to bodily sensations.

94 It might be said that a felt need has a sensuous component and an intentional component constituted by some cognitive state. Thus, to feel that one must become president is, say, just to think of being president and for this thought to be accompanied by, or the cause of, some bodily sensation. But that does not seem right either. When one feels that one must become president then one apprehends this possibility with a particular force. This, intuitively, is something quite different to thinking of something and this thought to be accompanied by the experience of muscle tension. See also the discussion of affective intentionality in chapter three.

95 A related point: we can see the point of ethical criticism of people's desires if desire involves the experience that something must become reality. That is to say, we can see how it makes sense to regard people as vile who experiences the felt need for their colleague's downfall. But we can't make any sense of the suggestion that people should be the subject of such criticisms in virtue of their bodily sensations.
4.4.2 The Argument from Content

The felt need theory offers a phenomenological conception of desire in so far as it considers a conscious state to be necessary and sufficient for desire. Michael Smith maintains, however, that no phenomenological conception of desire can be adequate. “Ascriptions of desires”, Smith says,

 unlike ascriptions of sensations, may be given in the form ‘A desires that p’, where ‘p’ is a sentence. Thus, whereas A’s desire to φ may be ascribed to A in the form ‘A desires that she φs’, A’s pain can’t be ascribed to A in the form ‘A pains that p’. (Smith [1994]: 107)

This, Smith argues, provides a reason to reject the phenomenological conception of desire. He writes:

For, according to this conception, there is no difference between desires and sensation. Each is a state that simply and essentially has phenomenological content. The strong phenomenological conception of desires is thus unable to account for the fact that desires have propositional content at all. (Smith [1994]: 107-108)

As it stands, this is not much of an objection to the felt need theory. After all, the felt need theory does not identify desire with a kind of sensation. Both experiences of felt need and sensations are states with phenomenal character. But an experience of felt need involves a mode of intentionality, a way of grasping a content that is before one’s mind. Sensation is not.

Smith acknowledges that this onslaught is an attack against a ‘straw man’. However, he thinks that once it is seen that desires can’t be identified with sensations, one is also in a position to argue against all versions of the phenomenological conception, even the more plausible weaker conceptions according to which desires are like sensations in that they have phenomenological content essentially, but differ from sensations in that they have propositional content as well. (Smith [1994]: 108)

The crucial point, Smith says, is that such conceptions “in no way contribute to our understanding of what a desire as a state with propositional content is, for they can’t explain how it is that desires have propositional content.” (Ibid.).

96 See also Smith [1998].
Evidently, Smith assumes that any adequate theory of desire must give us a theory of content determination. However, this seems to demand too much. It seems perfectly legitimate to accept a modest amount of division of labour. We may need one theory to explain in virtue of what a state has its intentional content and another theory to explain in virtue of what a state is a desire rather than, say, a belief. The possibility that two theories are needed to answer two questions in no way undermines the credibility of one theory as an answer to one question. To be sure, a theory of desire must not be inconsistent with the correct theory of content. But this does not mean it must itself provide a theory of content as well.\(^{97}\)

### 4.4.3 The Argument from Epistemology

Descartes famously thought the mind to be wholly transparent to the subject. In the ‘Cartesian theatre’ nothing can be hidden from the subject nor can the subject possibly be mistaken about the contents of her mind. This conception of the epistemology of the mind has since come into disrepute. Almost everyone nowadays accepts that people can fail to know some of their mental states and that error can creep into one’s own beliefs about ones mental life. That is, peoples’ understanding of their own mind can be incomplete and erroneous. That being said, most philosophers also think Descartes was on to something. For example, we do enjoy a privileged epistemological standpoint with respect to our beliefs and desires. At the very least, we have a greater authority when we speak about our beliefs and desires than we have when speaking about someone else’s.\(^{98}\) Moreover, when restricted to consciousness alone, the Cartesian doctrine is also much more inviting. Even if the mind is not wholly transparent to the subject, the subject’s own conscious states may well be.

These intuitions work both in favour of, and against, the felt need theory of desire. On the one hand, we can argue that any explanation of the peculiar first-person authority we have with respect to our desires must ultimately be grounded in the link...

---

\(^{97}\) Smith thinks that a disposition to be motivated to act is necessary and sufficient for desire. He also thinks that this has the benefit of providing us with a theory of content: to desire that \(p\) is to be disposed to act in ways that would bring about \(p\) if one’s beliefs were true. (Smith [1994]; Smith [1998]. This theory of content was developed by Stalnaker [1987] and Stame [1977]). However, the case of N-desires provides a serious problem for this theory because in the case of N-desires the relevant belief could not be true. Consequently, the antecedent of the counterfactual ‘if one’s beliefs were true, then ...’ is necessarily false and the counterfactual true regardless of its consequent. The theory, therefore, has the false implication that a N-desire has every possible content. In short, it would be a virtue if Smith’s theory of desire also provided us with a theory of content, but it does no such thing.

\(^{98}\) Dretske is a lone dissenter. He writes “we are all very good – in fact absolute authorities – about what we think and experience, but not very good (in fact, I think, very bad) about the attitudinal aspect of these mental states” (Dretske [1999], p. 55). This seems to be doubly mistaken. It does seem, as I shall suggest below, that we can be mistaken about the content of our mental states. But, more importantly, it seems rather bizarre to think that we are somehow bad at telling whether we desire, imagine or believe that \(p\). Errors may be possible, but it is surely false to claim that we are “very bad” at telling this.
between desire and desirous consciousness. At the very least, cutting off any essential links between desire and distinctive conscious states will make it much more difficult to explain our peculiar first-person authority.\textsuperscript{99}

At the same time, it might be thought that the fact that we can be ignorant of our desires and harbour false beliefs about them also raises a problem for affective theories. After all, how could there possibly be errors of this kind if consciousness is wholly transparent to the subject and if desire is essentially linked to distinctive conscious states? We may argue that such errors are impossible, but that is counterintuitive. Suppose, for example, there is a person who constantly goes out of her way to catch a glimpse of herself in reflective surfaces. At the same time, however, let us suppose she sincerely denies having the desire to contemplate her appearance. Isn’t it possible that she is mistaken in this? And if it is possible, how can the felt need theory account for this? Or suppose that a person professes a deep and non-instrumental desire to be a musician but an impartial observation of her life makes it much more plausible to think she simply has a deep and non-instrumental desire to please her mother who has great hopes for her becoming a famous musician. Again, how could someone, if the felt need theory were true, mistakenly think she has a non-instrumental desire when, in fact, she does not?\textsuperscript{100}

The difficulties are real but far from fatal. Indeed, we can take them to refute the felt need theory only if we assume that knowledge of our own conscious states must always be complete and shielded from all error. That assumption, however, seems too strong. Various factors can prevent the contents of consciousness from becoming the objects of belief and knowledge. Put another way, what a person is conscious or aware of is one thing, what she comes to believe is another. Cognitive factors, such as limits to attentional spans and processing power, are one culprit. Looking at a scene, there is a sense in which we are conscious of many more details than we come to know about, because limits of attention span and processing power interfere with the integration of all the information contained in our perceptual states into our web of beliefs. Motivational influences are at least as important. J.C.B. Gosling, for example, invites us to consider the case of a man who sits on many committees which are, of course, a great trial to him and he is constantly reminding us of his burdens. It is noticeable, however, that whenever the prospect of such work arises there are always good reasons why he should accept it, and that if committee pressure slackens off, during holiday times say, or because of the dissolution of committees, then he tends to grow restive. Further, the business of the committees which he chairs always seems to extend for a considerable time even when the agenda is of the shortest. (Gosling [1969]: 45)

\textsuperscript{99}See, for example, Goldman [1970] p. 98. Goldman criticises an account of desire purely in terms of dispositions to act on the grounds that it is incompatible with our privileged knowledge of our own desires.
\textsuperscript{100}See Smith [1994] for these two examples. See also Hulse et al. [2004] for related claims.
Gosling concludes, “the fact is that he clearly enjoys committee work” (Ibid.). Nonetheless, it would be a mistake to infer that this man’s enjoyment of committee work does not consist of a conscious state. Rather, we must acknowledge that despite enjoyment being a specific conscious state, it does not follow that the man necessarily knows that he is in that conscious state. The man fails to integrate his experience into his web of beliefs, and he fails because otherwise he would have to concede a “certain pettiness and self-importance” (Gosling [1969]: 46), something he is motivated not to concede.

The same pattern is familiar from many other emotions. We think a person can have experiences through which anger, hate, jealously, or love manifest themselves but also be subject to motivational influences that prevent her from properly owning up to having these emotions. We are also familiar with the way in which motivational influences can cause the agent to form more acceptable beliefs about the objects of her emotions. A person may believe she is happy that justice has been done, when she is actually happy about seeing the defendant being humiliated. Knowledge of our desires is subject to the same distorting motivational influences. We can self-deceptively fail to acknowledge desires that have manifested themselves in our experiences, or self-deceptively form erroneous beliefs about what it is that we ultimately desire, simply because owning up to these desires would be too painful. And this, it appears, is fully compatible with a conscious experience being necessary and sufficient for desire. We must only reject what is in any case untenable: that our knowledge of our own conscious states is necessarily complete and free from error.

4.4.4 The Argument from Motivation
Perhaps the most prominent objection to making an experience of felt need necessary for desire, though, goes like this: according to a venerable tradition, motivation and intentional action implies desire. But it seems implausible to maintain that motivation and intentional action implies a state individuated by its relation to felt need. Consider, for example, the case of Peter who is tending to the cats at his sister’s place because he has promised to do so. According to the received account of the nature of motivation and intentional action, this would have been impossible without a relevant desire, such as the desire to keep his promise. Yet, while his action reveals this desire, his conscious life need not. As he goes about feeding the cats he may experience disgust towards cats and an intense yearning for rest after an arduous day at work. What he does not experience is a felt need to keep his promise. Nor does the possibility of the scenario appear to require that he must be subject to this felt need at other times. Peter may never have felt the prick of desire to keep his promises yet still keep them and therein demonstrate that he has the desire.
I think the argument stands scrutiny if we accept that motivation and intentional action implies desire. But should we accept this? In the next chapter I shall address this question in much greater detail, and argue that intentional action does not imply desire. If my argument in the next chapter succeeds, the objection will be undermined.

In this chapter I have developed an affective theory of desire. The core idea of an affective theory is that just as belief is constituted by a cognitive mode of intentionality, desire is constituted by an affective mode of intentionality — a mode of intentionality inextricably linked to distinctive conscious episodes and in virtue of which desire is a pro-attitude. There are many ways to develop the specifics of an affective approach to desire. These primarily concern different ways of understanding the relevant conscious episodes that anchor our conception of the affective mode of intentionality at stake. I argued that putting experiences of felt need centre stage makes for the best affective theory of desire, and that having a felt need for \( p \) from time to time is both necessary and sufficient for the (non-instrumental) desire that \( p \).

This analysis of desire has many benefits, chief among them is that it offers a pertinent account of the pro-attitudinal nature of desire and, consequently, an explanation of its role in rationality — an explanation that hinges upon the experiential pro-ness of desire and that is, therefore, an explanation that remains fundamentally non-cognitive. I have also argued that, despite the opposition of the mainstream analytic to the idea of a phenomenology of desire, the felt need theory has nothing to fear from a number of well-known objections. However, the task of defending this theory is not done. We must still show that the felt need theory can illuminate the role of desire in the explanation of, and reasons for action. In the final two chapters, I shall address these two issues.
David Hume famously stated that “reason alone can never be a motive to any action of the will ...Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them” (Hume [2003]: 2.3.3). Inspired by this, many latter-day Humeans maintain that motivation requires desire. If they are right, then the felt need theory is in trouble because, as I conceded at the end of the last chapter, it is implausible to suppose that motivation implies a state distinguished by its relation to experiences of felt need. However, Humeans are not right. Motivation does not imply desire. Or so I shall argue.

My defence of Anti-Humeanism will take up the greater part of this chapter. In section one, I will begin by clarifying the Humean and Anti-Humean positions. This is important as there are a number of distinct Humean and Anti-Humean theses that are frequently run together. In section two, I shall provide an argument for one important Anti-Humean thesis, namely, that belief alone can constitute a motivating state. In section three, I shall argue that Smith’s influential argument to the contrary is unpersuasive. In section four, I shall turn to defend another Anti-Humean thesis, namely, that motivation does not imply desire. And in section five, I shall seal my defence of the Anti-Humean position by rebutting the general Humean argument that there is “a perfectly good sense in which everything which we do, meaning to do it, is what we want to do” (Armstrong [1968]: 153).

Establishing that the Humean thesis is flawed serves the important purpose of undermining the remaining objection to the felt need theory. However, if desire is not necessary for motivation and action, what role does it play in the explanation of action,
and can the felt need theory account for this role? The final section of this chapter will
turn to these questions, and I shall argue that the felt need theory does indeed have
the resources to account for the role of desire in the explanation of action.

5.1 Clarifying the Dispute

There are a number of theses frequently advanced under the umbrella of Humeanism.
Equally, there are a variety of theses called Anti-Humean. To avoid confusion, we
must draw some distinctions and put some clarifications in place.

One Humean thesis is that being motivated to perform an intentional action
implies desire. Call this Humeanism about motivation. This claim is opposed by Anti-
Humeanism about motivation, that is, the thesis that being motivated to perform an
intentional action does not imply desire.

Many philosophers who argue against Humeanism nonetheless accept Humeanism
about motivation. In his seminal critique of Humeanism, Thomas Nagel, for example
concedes, “whatever may be the motivation for someone's intentional pursuit of a goal,
it becomes in virtue of his pursuit ipso facto appropriate to ascribe to him a desire for
that goal” (Nagel [1978]: 29). Similarly, both John McDowell and Philippa Foot voice
dissatisfaction with Humeanism while at the same time accepting Humeanism about
motivation. If Humeanism about motivation were the only pertinent Humean
thesis, it would be hard to make sense of this.

The air of paradox dissolves if we draw a distinction between motivation and
motivating states. Motivation is a state individuated by its connection to action. To
be motivated to act is to be in a state that will carry one towards action, at least if one
is able to act and no countervailing motivations are at play. Motivating states, on the
other hand, are individuated by their role in rationalising other states. Let us say that a
state S rationalises some other state S* if and only if (i) it is a requirement of rationality
that one be in S* if one is in S (ii) an agent is in S* because she is in S. Furthermore,
let us say that a motivating state is a state that is apt to rationalise some other state (e.g.
the state of intending to act). Thus, a state S is a motivating state if and only if (i) it is
a requirement of rationality that one be in S* if one is in S and (ii) it is possible that an
agent is in S* because she is in S.

---

104 McDowell [1998]; Foot [1978b].
105 I shall follow Davidson in assuming that this 'because' marks a (complex) causal relation (Davidson
[1980a]; Davidson [1980b]. For an opposing position see, for example, Sehon [1994]). Nothing of substance
depends on this. Note further, it might be said it is also necessary that the agent is in S* partially because the
agent recognises the requirement of rationality to which she is subject. If this should be correct, we would
need to add the additional clause throughout. Again, nothing of substance depends on this.
If a motivating state $S$ rationalises a person’s being in $S^*$, then it is natural to say that $S$ is the person’s *reason* for being in $S^*$. This is why motivating states are also commonly referred to as *motivating reasons*. We must, however, be careful to distinguish motivating reasons from *normative reasons*. Motivating reasons are tied to requirements of rationality and, as I said in chapter one, what rational requirements a person is subject to is determined solely by her mental properties. In contrast, normative reasons, such as moral or prudential reasons, are determined, at least in part, by non-mental properties. Consequently, mental duplicates will share all motivating reasons but need not have the same normative reasons.\(^{106}\)

Drawing this distinction between motivation and motivating states/reasons suggests a possible interpretation of philosophers such as Nagel, McDowell, and Foot. For it might be suggested that these philosophers endorse the thesis that motivation implies desire but reject the thesis that motivating states imply desires. However, this suggestion does not yet provide us with a fully adequate interpretation of all of these philosophers. McDowell, for example, seems to think motivating states imply motivation.\(^{107}\) Since he also accepts that motivation implies desire, he could not consistently deny that motivating states imply desire. To do justice to McDowell’s position we must, therefore, draw a further distinction between what *constitutes* a motivating state and what is *implied* by a motivating state. This allows us to read McDowell as embracing the claims that being in a motivating state implies being motivated and that being motivated implies having a desire, but as denying that motivating states must be constituted by desire.\(^{108}\)

It is unclear whether Nagel or Foot also think being in a motivating state implies being motivated. On the other hand, it is clear that both agree with McDowell that motivating states need not be constituted by desire. Indeed, it seems to be their shared commitment to the thought that a specification of an agent’s motivating reason need not involve any reference to desire that has earned them the badge of Anti-Humeanism. Thus, in addition to distinguishing between Humeanism about motivation and Anti-Humeanism about motivation, we must also distinguish between *Humeanism about motivating states* – the thesis that motivating states are necessarily constituted, at least in part, by desire – and *Anti-Humeanism about motivating states* – the thesis that motivating states need not be constituted, not even in part, by desire.

\(^{106}\) In chapter six I shall say more about normative reason. For the distinction between motivating reasons and normative reasons, see also Baier [1965]; Darwall [1983]; Smith [1994].

\(^{107}\) This, at least, is suggested by McDowell’s discussion of the virtuous person in McDowell [1998].

\(^{108}\) See also Smith [1994]: 92-93.
5.2 An Argument for Anti-Humeanism about Motivating States

One way for Anti-Humeanism about motivating states to be true is for there to be motivating states that consist of belief alone. However, motivating states can consist of belief alone only if there are beliefs such that rationality requires that one intend to act if one has those beliefs.\(^{109}\) In the next subsection, I shall argue at least one belief satisfy this criterion.

5.2.1 Belief, Intention, and Rational Requirements

Moral belief is one candidate. Suppose Maria believes she is morally required to take care of her ailing parents. It might be said that if Maria holds this belief and lacks the intention to take care of her ailing parents, then she is irrational. More generally, it might be said that rationality requires that one intend to \(\varphi\) if one believes that one is morally required to \(\varphi\).

However, it is contentious whether this is true. Let us say that an *amoralist* is a person who has beliefs about what she morally ought to do but who is wholly indifferent towards the demands of morality. That is, an amoralist is much like the person who believes it is a requirement of etiquette that she does not talk while chewing but who does not care at all about doing what etiquette requires. It is plausible to think the amoralist ought to care about morality and that her failure to do so is condemnable. Indeed, some may even wish to go so far as to claim that the amoralist is rationally required to care and that her lack of concern is irrational. But given her indifference, a failure to intend to do as she believes she morally ought to do can’t itself be seen as subject to rational censure. Put another way, there is no lack of inner coherence between the absence of concern for morality and the absence of an intention to act in accordance with what one believes to be required by morality. Consequently, rationality can’t simply require that one intend to \(\varphi\) if one believes that one is morally required to \(\varphi\).\(^{110}\)

It is difficult to know what to make of this argument because it is very difficult to know whether amoralists are really possible. Many philosophers are adamant that it is impossible to have moral beliefs and be wholly indifferent towards morality, while others insist that this is a genuine possibility.\(^{111}\) It would be good if we could side-step this perennial debate. To do this I will focus upon a different kind of

---

109 Actually, this is a bit too strong since motivating states can consist of belief alone as long as there are beliefs such that rationality requires that one is in some mental state \(S\) if one has those beliefs.

110 Amoralists figure importantly in the Humean arguments of Blackburn [2000]; Brink [1989]; Foot [1978a]; Harsanyi [1958]; Nowell-Smith [1954]; Sumner [1968]; Svavarsdottir [1999]. It should be emphasised that most discussions fail to distinguish between the question whether it is possible to believe that one morally ought to \(\varphi\) and lack any motivation to \(\varphi\) and the question whether it is a requirement of rationality that one be motivated/intend to \(\varphi\) if one believes that one morally ought to \(\varphi\).

111 For some important discussions see Falk [1986]; Frankena [1958]; Hare [1952]; Smith [1994]).
belief: *the belief that, all things considered, one has most reason to* \( \varphi \) (for simplicity’s sake I shall henceforth omit the qualification ‘all things considered’). I shall argue rationality really does require that one intend to \( \varphi \) if one holds this belief.

Suppose we have overheard Fred saying that he believes he has most reason to finish his project before the looming deadline but then run into Fred at the beach just as he is about to hit the waves. Surprised, we ask him whether he has already finished his project, to which he replies he has not and he does not intend to do so. Does that mean, we continue asking, he has revised his opinion about what he has most reason to do? Fred answers that he has not. Surely, in that case Fred is irrational. More generally, it seems that a person is irrational if she believes she has most reason to \( \varphi \) but does not intend to \( \varphi \). Put another way, it really seems as if rationality requires that one intend to \( \varphi \) if one believes that one has most reason to \( \varphi \).

However, it might be thought this thesis is subject to a similar objection as the claim that rationality requires that one intend to \( \varphi \) if one believes that one is morally required to \( \varphi \). Let us say that an *arationalist* is a person who has beliefs about what she has most reason to do but who does not care about doing what she has most reason to do. Thus, the arationalist’s attitude towards doing what she thinks she has most reason to do is the same as the amoralist’s attitude towards doing what she thinks she is morally required to do. Again, we might think the arationalist’s indifference is condemnable, indeed, even subject to rational reproof. However, it might be said that, given the arationalist’s lack of concern, the failure to intend in line with what she believes she has most reason to do is not itself subject to rational criticism. Put another way, it might be said there is no lack of coherence between the absence of concern for doing what one believes one has most reason to do and the absence of an intention to act in accordance with what one believes one has most reason to do. If that is right, then rationality can’t require that one intend to \( \varphi \) if one believes one has most reason to \( \varphi \). (Note: the objection is not that some people know it is irrational not to intend what they believe they have most reason to do yet also do not care about being irrational. Rather, the objection is that being indifferent to doing what one believes one has most reason to do absolves one from the charge of irrationality if one fails to intend in line with that belief.)

This objection is considerably weaker than the analogous amoralism objection – even if we grant the possibility of arationalists. Let us ask what it means to be indifferent towards, or not to care about, doing what one believes one has most reason to do. Perhaps the most plausible interpretation is that it means that believing one has most reason to \( \varphi \) is not something that weighs with one in one’s decision making process.

---

112 Blackburn [2000] pp. 90-91 might be confusing these two points. Note also that the possibility of people not caring about being irrational does not help the Humean. As we shall see, the argument for Anti-Humeanism about motivating states depends only on the premise that there are rational requirements that govern the relation between belief and intention.
This, in turn, can be understood in three ways. First, it can mean simply that one’s intentions are not sensitive to what one believes one has most reason to do. However, this hardly disproves that they ought to be. A person does not absolve herself from the charge of irrationality by failing to do what rationality requires. Second it can mean that even though one believes one has most reason to φ one does not think that one has any reason to φ. This would make the appeal to the arationalist incoherent. Perhaps we can make sense of the amoralist saying “I know I morally ought to φ, but I don’t think I have any reason to φ”. But the exclamation “I know I have most reason to φ, but I still don’t see any reason to φ” is simply unintelligible. Third, it can mean that one does not regard the fact that one believes one has most reason to φ to provide one with a reason to φ. This may be quite sensible, but it does not protect one from the charge of irrationality if one fails to intend to φ. Let me explain.

Suppose Miranda gets free tickets to the theatre. She expects the play to be thought-provoking and enriching. Reflecting upon her options she concludes that she has most reason to see the play. Suppose further, however, Miranda fails to intend to see the play and we criticise this as irrational. Finally, suppose Miranda defends herself by saying she is not irrational because the play will also be attended by the prime-minister, and she does not think this provides a reason for seeing the play. We would be stunned by this response. If Miranda had pointed to considerations that made her revise her judgment that she has most reason to see the play, we would have been sympathetic to her defence. Surely though, her thinking that some consideration does not provide her with an additional reason can’t absolve her from the charge of irrationality. Given that she already believes she has most reason to see the play the fact that she does not accept yet further considerations to speak in favour of seeing the play is simply irrelevant. The same point applies to the arationalist. The arationalist does not regard believing one has most reason to φ to be the kind of thing that can provide a reason for φ-ing. Even so, since she already believes she has most reason to φ this can hardly absolve her of irrationality if she does not intend to φ. (This simple point may be obscured by the analogy with the amoralist. Given the amoralist’s indifference towards morality the amoralist can believe she is morally required to φ and at the same time also believe she does not have any reason whatsoever for φ-ing – and it is this belief that explains why we can’t criticise the amoralists failure to intend to φ as irrational. As I pointed out, above, since the arationalist ex hypothesi believes she has most reason to φ she can’t coherently believe that she has no reason whatsoever to φ.)

I conclude that there are no good reasons to deny that the person who believes that she has most reason φ but lacks the intention to φ is irrational. Rationality really does seem to require of us that we intend to φ if we believe that we have most reason to φ.
5.2.2 A Non-Cognitivist Objection

This suggests a simple yet very powerful argument for Anti-Humeanism about motivation: it is a requirement of rationality that one intend to \( \varphi \) if one believes one has most reason to \( \varphi \). Moreover, sometimes a person intends to \( \varphi \) because she believes she has most reason to \( \varphi \). Believing that one has most reason to \( \varphi \) is, therefore, apt to rationalise intending to \( \varphi \). But a motivating state just is a state apt to rationalise some other state. It follows that motivating states need not consist of desire; Humeanism about motivating states is false, Anti-Humeanism about motivating states true.

In the next subsection I shall discuss the objection that it is impossible to intend to \( \varphi \) simply because one believes one has most reason to \( \varphi \). Instead, a full explanation of why an agent intends to \( \varphi \) must always involve include a reference to one of her desires. In this subsection, however, I shall discuss the objection that ‘believing’ that one has most reason to \( \varphi \) is a state constituted by desire.\(^{113}\) To show that this state can constitute a motivating state, therefore, falls short of establishing that belief alone can constitute a motivating state. Let us consider this idea in a bit more detail.

Ordinarily we think of moral statements as truth-evaluable. For example, we think if someone says keeping promises is morally right, she says something that is true if keeping promises is in fact the morally right thing to do and false otherwise. Non-cognitivists deny this. They claim that moral statements are not truth-evaluable and that when we make moral statements we are not doing something that falls within the domain of aiming to represent the way the world is. Rather, moral statements express emotions, desires, conduct-prescriptions, or other non-cognitive states. What about moral beliefs? Strictly speaking, the non-cognitivist denies that there are such things as moral beliefs. According to her, our moral mental life is essentially non-cognitive.

Ordinarily we also think of the statement that one has most reason to \( \varphi \) as truth-evaluable. For example, we think if someone says she has most reason to finish the paper, she says something that is true if she indeed does have most reason to finish the paper and false otherwise. A non-cognitivist might deny this. She might claim

\(^{113}\) This must be distinguished from the related claim that being in this state implies desire, though it is not a state constituted by desire. Two comments are in order with respect to this claim. First, it is consistent with Humeanism about motivating states being false. Second, as far as I can see there are only two motivations for thinking that believing one has most reason to \( \varphi \) implies desire. (i) It might be said that believing one has most reason to \( \varphi \) implies desire because our reasons for action are tied to our desires. In chapter six, I shall question having a reason to act implies desire. Even if it did, though, it would not follow that one can hold this belief only if one has certain desires, but rather that one can hold this belief and this belief be true only if one has certain desires. (ii) It might be said that believing one has most reason to \( \varphi \) implies being motivated to \( \varphi \) and that being motivated implies having a desire. I shall argue below that believing one has most reason to \( \varphi \) does not imply being motivated to \( \varphi \). Moreover, as I will argue in discussing the second non-cognitivist’s argument, it is hard to see how this response could be successful anyway. If believing that one has most reason to \( \varphi \) implies being motivated to \( \varphi \), we can maintain that being motivated can consist of belief alone. On the other hand, in order to deny that being motivated can consist of belief alone we must presuppose that no belief implies being motivated. Either way, the argument collapses.
that such statements are not truth-evaluable and that when we make them we are not doing something that falls within the domain of aiming to represent the way the world is. Rather, these statements express non-cognitive states. What about believing that one has most reason to $\varphi$? Again, the non-cognitivist might deny that, strictly speaking, there are such beliefs. That is, she might say it is impossible to regard as true the normative proposition that one has most reason to $\varphi$. Being in a state that we would normally describe that way really involves having some suitable desire.\(^{114}\)

I have said that one way for Anti-Humeanism about motivating states to be true is for there to be motivating states that consist of belief alone. Moreover, I said that motivating states can consist of belief alone only if there are beliefs such that rationality requires that one intend to act if one has those beliefs. It seemed as if we had identified a belief that satisfies this criterion. However, if the non-cognitivist is right, this is an illusion. We would still not have identified a rational requirement that governs the relation between belief alone on the one hand and intention on the other. Moreover, the non-cognitivist says there are three arguments that prove her analysis must be correct. Let us consider these more closely.

### 5.2.2.1 The Case for Non-Cognitivism

The first argument is that we can combine any belief with any practical attitude without irrationality but that we can’t ‘believe’ we have most reason to $\varphi$ and lack the intention to $\varphi$ without irrationality. Therefore, ‘believing’ that we have most reason to $\varphi$ must consist, at least in part, of desire.\(^{115}\)

This simply begs the question. Rationality requires that one intend to $\varphi$ if one believes that one has most reason to $\varphi$. Unless, the non-cognitivist provides a reason for why we must accept a non-cognitivist analysis of this ‘belief’, we can maintain that this requirement of rationality provides a counterexample to the assumption that we can combine any belief with any practical attitude without irrationality.

The non-cognitivist might try to defend her first argument by saying that there certainly are many beliefs we can combine with any practical attitude without irrationality. Yet, then it is “unacceptably hybrid” to also think that there are some beliefs that, on pain of irrationality, require an intention to act. After all, what makes these beliefs so special?\(^{116}\) Intended by the non-cognitivist as a knock-out question, the cognitivist

---

114 Non-cognitivists may also analyse ‘believing’ that one has most reason to $\varphi$ in terms of some other non-cognitive state. Perhaps some such analysis would allow them to avoid the problems I raise below. This, in turn, would show that my argument for the claim that belief alone can constitute a motivating state will remain vulnerable to a certain sort of non-cognitivist objection. Even so, it would do nothing to safeguard Humeanism about motivation. And all I need in order to defend the felt need theory against the objection from motivation is to show that Humeanism about motivation is false.

115 For this style of argument see for example Nowell-Smith [1954] and Urmson [1968].

116 See Dancy [2000]: 93-94 for a related suggestion. Note, though, Dancy’s argument also ties in with the argument from explanation which I shall consider separately in section 5.2.3.
actually has a simple answer. Rational requirements supervene in part upon content. And it is in virtue of their content that some beliefs stand in rational relations to intentions. Other beliefs lack the required content.

At this point the non-cognitivist may appeal to a second argument in support of her analysis: 'believing' that one has most reason to $\phi$ implies at least some motivation to $\phi$, but being motivated never simply consists of belief. Rather, it always consists, at least in part, of desire.¹¹⁷

I shall argue below that the first premise of this argument is false: that is, I shall argue it is possible to think one has most reason to $\phi$ and lack any motivation to $\phi$. If this is correct, this non-cognitivist’s argument fails as well. Let us, however, grant this premise for the moment. What about the second premise? If ‘believing’ one has most reason to $\phi$ implies being motivated to $\phi$, why not conclude that motivation can be internal to belief, that is, that at least in some instances of being motivated to act, the agent’s motivation consists in nothing other than a belief?

If being motivated to act and desiring were one and the same state (and the desire-as-belief thesis false), we could straight away set this possibility aside. However, as I argued in chapter one, we can’t identify desire with being motivated. Still, it might be said that being motivated can’t consist of belief alone because being motivated implies having a goal, and having a goal just is having a desire. The argument is powerful as it is very plausible to think that being motivated to $\phi$ implies having a goal, namely, the goal to $\phi$. Moreover, it seems that having a goal must always just consist of having a desire. After all, having a goal is being in a state with which the world must fit. But being in a state with which the world must fit is having a desire.¹¹⁸

However, things are a bit more complicated. Goal representations are states with which the world must fit, that is, they have the world-to-mind direction of fit and so does desire. But it isn’t obvious that desire is the only state that has the world-to-mind direction of fit. In particular, it isn’t obvious that belief can’t have the world-to-mind direction of fit. Of course, belief is the paradigmatic example of a state that has the mind-to-world direction of fit in as much as the belief that $p$ will manifest some kind of failing if $p$ is not the case. This rules out that a belief that $p$ could have the world-to-mind direction of fit with respect to the proposition that $p$. It is, however, consistent with the possibility that this belief has the world-to-mind direction of fit with respect to the proposition that $p^\ast$. Thus, while it would be clearly false to claim that believing one has most reason to $\phi$ has the world-to-mind direction of fit with respect to the content that one has most reason to $\phi$, it does nothing to undermine the thesis that it is a state that has the world-to-mind direction of fit with respect to the content that one $\phi$.

¹¹⁷ See Blackburn [2000]; Darwall et al. [1992]; Hare [1952]; Miller [2003].
This complexity is clearly appreciated by Smith in his influential discussion of this matter. Smith nonetheless concludes that having a goal just is having a desire, because he thinks for every belief it is always possible that an agent has that belief but fails to be in a state with which the world must fit. Consequently, no belief could itself have the world-to-mind direction of fit with respect to any content.\textsuperscript{119}

Smith's claim that one can believe anything and not be in a state with which the world must fit is controversial. A possible counterexample is that believing one has most reason to \( p \) implies being motivated to \( p \) and, consequently, implies being in a state with which the world must fit. To defend Smith's thesis one must, therefore, do either of two things. One must show that the counterexample fails because 'believing' one has most reason to \( p \) is not really an instance of belief alone but involves, at least in part, some non-cognitive state like desire. Alternatively, one must show that the counterexample fails because, even in the case of this belief, it is possible to have the belief and not be in a state with which the world must fit. Smith opts for the latter, but the non-cognitivist can't do the same because her argument presuppose that this 'belief' does imply being motivated to \( p \). However, in order to make the first reply the non-cognitivist owes us an argument for accepting a non-cognitive analysis of 'believing' one has most reason to \( p \). And if the non-cognitivist is not to beg the question at this point, the argument can't be that 'believing' one has most reason to \( p \) implies being motivated to \( p \) but being motivated to \( p \) can't consist of belief alone but implies desire. Either way, the second non-cognitivist argument fails to persuade.

The third and final argument in the non-cognitivist's toolbox is that a cognitive analysis of believing one has most reason to \( p \) implies an unacceptable commitment to the existence of 'queer' properties, properties that, as Simon Blackburn mockingly puts it, are "intensely magnetic" (Blackburn [2000]: 90) and draw those who cognise them "irresistibly" (Ibid.). Therefore, the non-cognitivist concludes, we must accept her analysis if we do not want to take on unpalatable metaphysical commitments.

It is not entirely clear how this argument is supposed to go. Perhaps the most plausible interpretation is that the non-cognitivist assumes there are no properties that must motivate an agent if the agent is aware of their instantiation but that the cognitivist's account entails a commitment to such properties. However, this is not very convincing because a cognitive analysis of believing that one has most reason to \( p \) does not entail there are properties that must motivate an agent if the agent is aware of their instantiation. (i) One might reject the claim that believing one has most reason to \( p \) implies being motivated. Consequently, even on a cognitive analysis there will be no pressure to accept the existence of properties that must motivate if an agent is aware of their instantiation. (ii) One might claim that some beliefs imply being

\textsuperscript{119} Smith [1994]: 119-120.
motivated, but it is also possible to represent the properties that make these beliefs true without being motivated. That is, one might think that the connection between believing one has most reason to \( \varphi \) and being motivated to \( \varphi \) is due to the deployment of normative concepts and that it is possible to represent the relevant properties by deploying non-normative concepts. (iii) One might think that normative beliefs imply being motivated but also that normative beliefs are false because there are no normative properties. Once again, one's cognitivism would not commit one to the thesis that there are properties that must motivate an agent if the agent is aware of their instantiation.

5.2.2.2 The Case Against Non-Cognitivism

The non-cognitivist has no compelling arguments for her claim that believing one has most reason to \( \varphi \) must be analysed in terms of desire. In this section I shall argue that there are also very good reasons for rejecting any such analysis.

To start with, consider the thesis that holding this ‘belief’ consists of some non-normative belief and desire. This can’t be right. Failing to intend in line with what one ‘believes’ one has most reason to do is irrational. Desiring something without intending to act so as to bring it about need not be.

Suppose that the non-cognitivist replies by clarifying her position is that ‘believing’ one has most reason to \( \varphi \) involves the agent’s strongest desire. This can’t be right either. Suppose Eddy ‘believes’ he has most reason not to eat another piece of that delicious chocolate cake. Must this really involve his strongest desire? Surely not. Surely, his strongest desire might be to eat another piece or even two.

The non-cognitivist is obliged to say that although it may seem to Eddy that his strongest desire is to eat another piece of cake, he must in fact desire something else, like keeping to his diet, even more strongly. This is problematic. People can make mistakes about what they desire or desire most strongly. But we ordinarily think whether they make a mistake can only be determined by careful consideration of all the complexities involved in a particular case. The non-cognitivist’s claim on the other hand must depend upon a much more general justification. And, as far as I can see, the only justification available is to appeal to the claim that strength of desire is a function

---

120 There are a variety of well-known problems for non-cognitivism about moral discourse and moral ‘judgments’. For example, like ordinary assertions, moral statements are generally in the indicative mood. Next, we also typically appear to treat moral discourse as truth-evaluable; if someone said to us that it is morally right to torture innocent people, we would normally consider her as having said something that is false. Moreover, we often seem to engage in the serious attempt of trying to discover what is the morally right thing to do – an endeavour that seems to presuppose a commitment to the idea of moral truth and falsity. Finally, as has frequently been pointed out, we commonly assess moral arguments as valid or invalid, but there are serious problems making proper sense of this within a non-cognitive framework. (For some good discussions see Brink [1989]; Geach [1960]; Geach [1965]; Miller [2003]; Smith [1994]; Urmson [1968]). Though I shall not discuss these points in the text, I believe similar difficulties also arise for non-cognitivism about the statement or ‘belief’ that one has most reason to \( \varphi \).
of its role in the psychological organisation of the agent. Thus, the non-cognitivist might say that a desire, which influences one to 'believe' one has most reason to \( \varphi \) must thereby be one's strongest desire.

This will not do, though. Suppose that despite 'believing' he has most reason to abstain from eating more cake, Eddy grabs another piece and quickly devours it with guilty pleasure. If strength of desire were determined by its role in the psychological organisation of the agent, this desire would be the strongest as it proves most effective. Thus, even if we grant that strength of desire is determined by its role in the psychological organisation of the agent, it does not follow that believing one has most reason to \( \varphi \) must involve one's strongest desire. In short, the claim that 'believing' one has most reason to \( \varphi \) must involve one's strongest desire violates our ordinary experience and the only possible justification is undermined by the fact that people sometimes intentionally do what they 'believe' they have most reason not to do. More generally, I conclude that not only are the non-cognitivist's arguments in support of her view rather weak, there are also substantial problems with a non-cognitivist analysis of 'believing' that one has most reason to \( \varphi \) in terms of desire.

### 5.2.3 Motivating States, Explanation and Desire

I have argued that believing one has most reason to \( \varphi \) is an instance of belief. I have also argued that it is a requirement of rationality that one intend to \( \varphi \) if one believes one has most reason to \( \varphi \). If one can also intend to \( \varphi \) because one holds this belief, motivating states can consist of belief alone. However, the Humean might say it is impossible to intend to \( \varphi \) simply because one believes one has most reason to \( \varphi \). Instead, the Humean might insist, a full explanation of why an agent intends to \( \varphi \) must always include a reference to one of her desires.

At first sight this Humean claim is not very plausible. It certainly seems that people sometimes intend to \( \varphi \) simply because they believe they have most reason to \( \varphi \). For example, suppose that Emma intends to quit her job. We would not be surprised to hear her explain that she intends to quit her job because she thinks, all things considered, this is what she has most reason to do. Or suppose we convince Emma that she has most reason to quit her job. We would not be surprised if as a direct result of coming to hold this belief, Emma also came to intend to quit her job. Moreover, it seems that causal relations are conceptually and metaphysically contingent. If that is true – that is, if "anything can cause anything" (Lewis [1997]: 144) – then it must be
possible for the belief that one has most reason to \( \phi \) to cause the intention to \( \phi \) and this should, at least if the causal relation is of the right kind, be sufficient for it to be true that an agent intends to \( \phi \) because she believes she has most reason to \( \phi \).\(^{121}\)

The Humean thinks that this is mistaken. She maintains that a full explanation of why we intend to \( \phi \) must always include a reference to some desire of ours. She might argue as follows: two people can believe they morally ought to \( \phi \) but only one be motivated to \( \phi \). Therefore, believing one morally ought to \( \phi \) can’t by itself explain being motivated to \( \phi \) even in the case of a person who holds this belief and is motivated. Something else needs to be added. And that is desire.\(^{122}\)

\textit{Mutatis mutandis}, the Humean might say two people can believe they have most reason to \( \phi \) but only one be motivated to \( \phi \). Moreover, since intending to \( \phi \) implies being motivated to \( \phi \), it follows that believing one has most reason to \( \phi \) can’t by itself explain the intention to \( \phi \) even in the case of a person who holds this belief and intends to \( \phi \). Something else needs to be added. And that is desire. Call this the \textit{argument from explanation}.

The argument from explanation is built upon a fairly strong assumption about the nature of explanation since it presupposes that only if the \textit{explanans} implies the \textit{explanandum} can an explanation succeed. Even if it is granted, though, two problems remain. First, suppose we accept that an intention to \( \phi \) can’t be fully explained by belief alone and that ‘something else needs to be added’. We should not automatically infer this ‘something else’ must be \textit{desire} or, indeed, any other \textit{additional mental state}. Instead, the ‘something else’ may just be the agent exercising her capacity for rational agency. After all, if an agent believes she has most reason to \( \phi \) and exercises her capacity for rational agency, she must be motivated to \( \phi \) and she must intend to \( \phi \).

Second, if we can’t explain an agent’s intention to \( \phi \) in terms of belief, then we can’t do so by appeal to desire either. For while it is possible to believe one has most reason \( \phi \) and lack the intention to \( \phi \), it is also possible to desire that \( \phi \), have an appropriate

\(^{121}\) One might deny that ‘anything can cause anything’. A sunburn, for example, can only be caused by the sun. Nothing else will do. However, while this is correct it is also true that anything could cause an irritation of the skin that is \textit{intrinsically indistinguishable} from a sunburn. It is just that this irritation would not be a sunburn because our concept of a sunburn is that of a state with certain intrinsic properties as well as a particular causal origin. Consequently, appealing to this complication would help the Humean only if our concept of an intention also is a concept of a state with certain intrinsic properties as well as a particular causal origin. I doubt that this is true. That is, in contrast to our concept of a sunburn, our concept of an intention does not appear to include a commitment to a particular causal origin. It seems that any state intrinsically indistinguishable from an intention to \( \phi \) is an intention to \( \phi \). If that is correct, then anything can cause the intention to \( \phi \), including belief. Furthermore, the only remotely plausible explication of the claim that intentions require a particular causal origin is that an intention to \( \phi \) must be caused by something that can \textit{rationalise} it. Thus, \textit{perhaps} it could be maintained that the state caused by a knock on the head could not be an intention to buy a new dress even if the state is intrinsically indistinguishable from this intention. Nonetheless, it would not follow that an intention could not be caused by belief. This is because given that it is a requirement of rationality to intend to \( \phi \) if one believes one has most reason to \( \phi \) there is no reason to deny that at least this belief could rationalise the intention to \( \phi \).

\(^{122}\) See Dancy [2000]; Goldman [1970]; Stocker [2004b]; Svavarsdottir [1999].
bridging-belief, but nonetheless lack the intention to $\varphi$. Applying the same line of reasoning, we should, therefore, say that if it is possible to have a desire and not intend to $\varphi$, the desire can't provide a complete explanation of the intention even in the case when the desire and the intention are both present. Something else needs to be added. But on pain of a vicious regress, this something can't be desire.

I have argued that it is possible that one intends to $\varphi$ because one believes one has most reason to $\varphi$. I have also argued that rationality requires that one intend to $\varphi$ if one believes one has most reason to $\varphi$. Finally, I have argued that believing one has most reason to $\varphi$ really is an instance of belief. Taken together, these claims imply that there are some beliefs that are apt to rationalise the intention to $\varphi$. Since a motivating state just is a state apt to rationalise another state, believing that one has most reason to $\varphi$ is a motivating state. It follows that Humeanism about motivating states is false, Anti-Humeanism about motivating states true.

5.3 Smith's Argument for Motivating State Humeanism

On a number of occasions, Smith has influentially argued that “having a motivating reason is, inter alia, having a goal” (Smith [1994]: 116) and that “having a goal just is desiring” (Ibid.). Since ‘motivating reason’ is but another term for ‘motivating state’, it follows that having a motivating state is, inter alia, having a desire – that is, that motivating states must consist in part of desire. Does this mean we must reconsider our conclusion that Humeanism about motivation states is false?

Believing one has most reason to $\varphi$ is, I have argued, a motivating state. Suppose this belief implies having a goal. In that case we can challenge Smith's second premise (as Smith himself concedes). Suppose, however, this belief does not imply having a goal. In that case we can challenge Smith's first premise. After all, if believing one has most reason to $\varphi$ is a motivating state and if holding this belief does not imply having a goal, then having a motivating state can't imply having a goal.

However, Smith might reply that – despite last section's detailed argument – we should be more confident that having a motivating state implies having a goal than we should be about believing one has most reason to $\varphi$ being a motivating state. This,

123 Note, too, that the same point applies with respect to merely being motivated to $\varphi$. As I have argued in chapter one, it is possible to desire that $p$ hold an appropriate bridging-belief without being motivated to act. Thinking otherwise is to confuse rational necessity with metaphysical or conceptual necessity. That is, it is to confuse the idea that being motivated would be rational given the belief-desire pair with the idea that it is metaphysically or conceptually necessary.

124 Another problem for the argument from explanation is this: it is possible that one believes one has most reason to believe that $p$ but fails to believe that $p$. This does not imply that if one believes one has most reason to believe that $p$ and also believes that $p$, this has to be explained by a suitable desire. Yet, if there is no need for a suitable desire in the theoretical case, why should it be required in the practical case? What feature of the practical case should make desire utterly essential?

125 The 'inter alia' is necessary because Smith thinks that having a motivating reason also involves having a conception of the means to attain that goal.
Smith might say is because his first premise really “has the status of a conceptual truth” (Ibid.). After all, “we understand what it is for someone to have motivating reason in part precisely by thinking of her as having some goal” (Ibid.). This is a surprising claim. I have explicated the notion of a motivating state in terms of it being a state apt to rationalise another state. Moreover, I have said that a state S is apt to rationalise another state S* if and only if (i) it is a requirement of rationality that one be in S* if one is in S and (ii) it is possible that an agent is in S* because she is in S. At no point is there a mention of agents having a goal.

Smith might respond by saying that my account of motivating reasons is flawed. A better account of motivating reasons, he could say, would have to take into account the fact that our conception of someone having a motivating reason is in part the conception of someone having a goal. One problem with this response is that it is not supported by his own explication of the notion of a motivating reason. Smith says, “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]: 96. Smith goes on to say motivating reasons are “potentially explanatory” (Ibid.)). Moreover, Smith says in light of this “motivating reasons would seem to be psychological states that play a certain explanatory role in producing action” (Ibid.). Finally, Smith also emphasises that citing a motivating reason would allow us to render an agent’s action intelligible. This is essential. For there is an a priori connection between citing an agent’s reason for acting in a certain way and making her acting in that way intelligible: that is, specifying what there is to be said for acting in the way in question. (Smith [1994]: 95)

At no point in his explication of the notion of a motivating reason does Smith deem it necessary to appeal to the notion of having a goal. In addition, it also seems that Smith’s claims are fully compatible with the explication of a motivating state/reason that I offered above. His insistence that motivating states must be (potentially) explanatory appears to cohere well with my claim that some state S is a motivating state only if it is possible that one is in some state S* because one is in S. Likewise, his emphasis on the capacity of motivating states to render other states ‘intelligible’ appears to cohere well with my claim that motivating states must be apt to rationalise other states. And, of course, my characterisation of motivating states is fully consistent with the requirement of motivating states being psychological states. So, given what Smith actually says about the notion of a motivating reason, he could hardly object to my characterisation of motivating reasons as fundamentally flawed.

---

126 Wallace [1990]: 374 gives another reason for being sceptical about this claim.
However, Smith could also respond by arguing that even though reference to having a goal need not be explicit in the characterisation of a motivating reason, it must nonetheless be implicit. One way of developing this suggestion would be based upon the idea that rationality tells us only what goals we ought to have relative to the goals and beliefs we already have. A motivating state would, therefore, have to consist, at least in part, of having a goal if it is to rationalise other goal states. Since having an intention consists, at least in part, of having a goal it follows that something could be a person’s motivating reason for having an intention only if that motivating reason consisted, at least in part, of having a goal.

This is not a route open to Smith. This is because three of Smith’s core claims are that “if an agent believes that it is right for her to φ, then she rationally should desire to φ” (Smith [1994]: 119), that “moral judgments of the form ‘it is right that I φ’ express a subject’s beliefs” (Smith [1994]: 12), and that having a belief never implies having a goal (Smith [1994]: 119-121). Consequently, Smith is committed to denying the thesis that by way of rationality one can only get to a goal-state from a starting point that already includes having a goal.  

Smith might also argue that the idea of a person being motivated to act is implicit in our conception of someone having a motivating reason. Yet, surely ‘it is a conceptual truth’ that ‘we understand what it is for someone to be motivated in part precisely by thinking of her as having some goal’. Indeed, I think this is exactly what Smith would say if pressed. As is apparent throughout his writings, Smith implicitly assumes that having a motivating reason is the same as being motivated. To give but one vivid illustration, consider the following passage:

He may well be motivated to do what he is required to do (that is, he may have a motivating reason to do what he has a normative reason to do), he may be motivated to so something that there is no normative requirement for him to do (that is, he may have a motivating reason to do what he has no normative reason to do), and there may be a normative requirement that he do what he has no motivation to do (that is, he may have a normative reason to do what he has no motivating reason to do)” (Smith [1987]: 39; my emphasis.)

---

127 Nor would it help other Humeans to go down that route. Rationality requires that one intend to φ if one believes one has most reason to φ. Suppose having this belief does not involve having a goal. In that case it would not be true that rationality only tells us what goals we ought to have relative to the goals and beliefs we already have. Suppose, on the other hand, that having this belief does necessarily involve having a goal. In that case we might conclude motivating states necessarily are, at least in part, constituted by having a goal. Yet, this would still not show that Humeanism about motivating states is true because in that case we could reject Smith’s second premise that having a goal just is having a desire. This is because – as we have seen above and Smith explicitly concedes – if there are beliefs that imply having a goal, we can maintain that having a goal sometimes consists in nothing over and above than having a belief.
Moreover - as I conceded in the previous section - it is plausible to think that being motivated implies having a goal. Since being motivated involves having a goal, and since Smith equates having a motivating reason with being motivated, it is not surprising to see that Smith thinks that having a motivating reason involves having a goal.128 However, it is a mistake to equate having a motivating reason with being motivated - or indeed, even to think that having a motivating reason implies being motivated. Smith thinks, and I agree, that having a desire and an appropriate bridging-belief constitutes a motivating state. But, as I argued in the first chapter, one can have any desire and any bridging-belief yet fail to be motivated. Thinking otherwise is to confuse rational necessity with metaphysical or conceptual necessity. That is, it is to confuse the idea that being motivated would be rational given the belief-desire pair with the idea that it is metaphysically or conceptually necessary. Moreover, I argued that believing one has most reason to (p also constitutes a motivating reason and as I shall argue in the next section - and as Smith agrees - it seems possible to hold this belief and at the same time lack any motivation to (p. So, unless Smith provides an independent reason for thinking that believing one has most reason to (p can't be a motivating state, this appears to provide a further counterexample to the general claim that having a motivating state necessarily involves being motivated.

Smith's argument that having a motivating reason is having a goal and that having a goal just is having a desire has been very influential. Upon closer examination, however, it turns out that there is surprisingly little force to the claim that having a motivating reason must necessarily involve having a goal.129 It can't, therefore, overturn the conclusion of the previous section's argument: that belief alone can constitute a motivating state; in short, that Humeanism about motivating states is false and Anti-Humeanism about motivating states true.

5.4 In Defence of Anti-Humeanism about Motivation

I have argued that a motivating state can be constituted by belief alone because one can be in a motivating state in virtue of believing that one has most reason to (p. Other beliefs may also be motivating states. Perhaps believing one has a reason to (p provides one with a motivating reason to be motivated to (p. Or perhaps believing one is morally required to (p provides one with a motivating reason to be motivated to (p, or even with a motivating reason to intend to (p. Alternatively, it might be argued that while these beliefs do not directly provide motivating reasons, they do so indirectly in so far as they can be a person's reason for believing she has most reason to (p.

128 Critics of Smith's argument often follow Smith in equating having a motivating reason with being motivated. Being critical of Smith's premise that having a motivating reason involves having a goal, they consequently find themselves forced to claim that "one may be motivated without having a goal" (Shafer-Landau [2003]: 135). This is, I think, rather implausible. Being motivated to (p really does appear to involve having (p-ing as a goal. But this does not mean that having a motivating reason must equally involve having a goal.

129 As we shall see in the next section, we should also reject Smith's second premise that having a goal just is having a desire.
However, even if Humeanism about motivating states is false, Humeanism about motivation might still be true. Indeed, it might be thought that Humeanism about motivation must be true because being motivated implies having a goal, but having a goal is having a desire.

Now, as we have seen above, if there are beliefs that imply being motivated, this argument fails. However, I doubt that there are beliefs that imply being motivated. Believing one has most reason to \( \varphi \) is as good a candidate as any for a belief that implies being motivated. Let us, therefore, consider whether it is possible to hold this belief and not be motivated to \( \varphi \). Patently, it is possible to believe that one has most reason to \( \varphi \) and not be most strongly motivated to \( \varphi \). Fred believed that he had most reason to finish his project, but he was most strongly motivated to go surfing. Still, it might be said that if Fred genuinely believed that he had most reason to finish his project, then he must have had at least some motivation to \( \varphi \). On this view, believing that one has most reason to \( \varphi \) would have some in-built motivational power, albeit a motivational power that is liable to being overridden.

It is difficult to prove that this must be false, but it does sit uneasily with some deep-seated intuitions about motivation. This is because it seems that in addition to recognising the possibility of motivation being overridden we also recognise the possibility of motivation being extinguished. The point has been admirably stated by Michael Stocker:

> Through spiritual or physical tiredness, through accidie, through weakness of body, through illness, through general apathy, through despair, through inability to concentrate, through a feeling of uselessness or futility, and so on, one may feel less and less motivated to seek what is good. One's lessened desire need not signal, much less be the product of, the fact that, or one's belief that, there is less good to be obtained or produced, as in the case of a universal Weltschmerz. Indeed, a frequent added defect of being in such 'depressions' is that one sees all the good to be won or saved and one lacks the will, interest, desire or strength. (Stocker [1979]: 744)

In this vein we can observe that if one is severely depressed, one might believe one has most reason to \( \varphi \) but lack any motivation to \( \varphi \). In that case it does not seem as if one is motivated to \( \varphi \), only to a lesser extent than, say, one is being motivated to rest and do nothing. Rather, it seems that the depression has the effect of leaving one without any motivation to do what one believes one has most reason to do. If that is right, and it certainly seems to capture an important aspect of the human condition, then believing one has most reason to \( \varphi \) does not imply having any motivation at all.
Nonetheless, it does not follow that Humeanism about motivation is true. Indeed, I think Humanism about motivation is false. The reason I think this is that having a goal is sufficient for being motivated, but having a goal does not imply having a desire. Consider, for example, a person who suffers from Alien Hand Syndrome. Sergio Della-Salla and his colleagues describe the condition as follows:

The phenomenon called ‘alien hand’ (AH) consists of the occurrence of movements of an upper limb that are unintended although clearly directed to some purpose. Such a person will often find his ‘alien hand’ seemingly autonomously perform tasks, such as buttoning up his shirt. That is to say, a person suffering from Alien Hand Syndrome experiences certain behaviors of one of his hands not to be under his control and often to be quite contrary to what he wants. (Della-Sala et al. [1991]: 1113)

Many tasks performed by an ‘alien hand’ are “clearly goal-directed” (Della-Sala et al. [1991]: 1114). But despite the fact that performing these tasks involves goal states, persons suffering from Alien Hand Syndrome will frequently complain that these behaviors do not conform to their desires, as when, for example, their ‘alien hand’ unbuttons the shirt they had just buttoned up. Moreover, they will sincerely deny wanting the shirt to be unbuttoned and it seems that a denial of this kind need not be mistaken. For although the task performed by their ‘alien hand’ is clearly goal-directed, it need not be a task that they have any pro-attitude towards. But desire involves having a pro-attitude. It follows that being in a goal state can’t consist of desire.

Here is another way to make the same point. A goal state is a state that plays a certain role in the regulation of behaviour. But a person can be in a goal state without this involving any kind of positive appraisal. A desire, on the other hand, involves an evaluative endorsement. It is impossible to desire something and be wholly negative or even wholly neutral towards it. Consequently, being in a goal state can’t simply consist of having a desire.

One might respond to this argument by challenging the claim that having a goal is sufficient for being motivated. Many philosophers are attracted to the view that intentional action and being motivated to perform an intentional action imply a rationalising account in terms of motivating states. They might, therefore, try to defend Humeanism about motivation by saying that it is not quite correct to describe cases of ‘alien hand’ behaviour as springing from motivation. This is because the goal state involved in these cases does not stand in the right kind of relation to a motivating state for it to count as a state of being motivated to act. Suppose this is true. The defence of Humeanism about motivation would still fail. This is because pointing out that a goal counts as a state of being motivated only if it stands in the right kind of relation to a motivating state does not mean that it is not a motivating state. It means that it is not a motivating state in the right kind of relation to a motivating state. The claim that having a goal is identical to having a desire is also subject to the counterexample of N-desires, though this case does not refute the thesis that goal states are necessarily constituted by desire.
of relation to a motivating state would support Humeanism about motivation only if Humeanism about motivating states were true. Yet, as I have argued, Humeanism about motivating states is false.

Alternatively, one might try to respond to this argument by challenging the claim that desire is a pro-attitude. Thus, the Humean might say while desiring something is typically accompanied by a pro-attitude, this isn't part of the nature of desire. This would allow one to insist that tasks performed by the 'alien hand' do involve desire and that the sincere denial by a person suffering from Alien Hand Syndrome must be mistaken.

This is quite an implausible position to take. Desire really does seem to be a pro-attitude (as I have also argued in chapter one). And it seems that the person who suffers from Alien Hand Syndrome need not be mistaken when she claims that she does not desire what her ‘alien hand’ is doing. However, suppose that it could be argued that there is a sense of desire according to which having a goal just is having a desire – that is, one usage of the term that simply refers to a goal representation that plays a distinctive role in the regulation of behaviour. If so, this sense would need to be sharply distinguished from the other, philosophically more interesting sense, in which desire is a pro-attitude and does have a range of normative roles to play, including being the kind of state that has a certain rational significance and that can partially constitute an agent's motivating reasons (see the final section of this chapter). For our purposes it is sufficient to note that it can't be said that having a goal just is having a desire in this philosophically more interesting sense.¹³¹

Belief, I argued, does not imply being motivated. Nonetheless, some beliefs can give rise to motivation without this involving desire. This is because a person who believes she has most reason to \( \varphi \) could come to have the goal to \( \varphi \) simply because she holds this belief. Moreover, having the goal to \( \varphi \) does not imply desire, but it does imply being motivated – at least, if one has the goal because of a suitable motivating state, such as believing one has most reason to \( \varphi \). I conclude, Humeanism about motivation is false as well.

5.5 The Argument from Wanting

In the last three sections I have argued that Humeanism is false. But it might be said that there is also a powerful argument that shows that Humeanism is true. The argument runs as follows: it is a necessary truth that intentionally \( \varphi \)-ing implies wanting to \( \varphi \). Clearly, there may be aspects of \( \varphi \)-ing that one does not want. Equally clearly, the sense in which one must want to \( \varphi \) if one is to \( \varphi \) intentionally is not so narrow as to pick out a phenomenally salient state. But there is "a perfectly good sense in which everything which we do, meaning to do it, is what we want to do" (Armstrong [1968]:

¹³¹ In this context, see also the discussion of Quinn-style cases towards the end of chapter one, as well as the discussion of conceptual pluralism in the introduction.
A desire in *that* sense, the Humean says, really is necessary for intentional action. Indeed, understood in this inclusive way, Humeanism is surely true – “maybe a trivial truth, but a trivial truth is still a truth” (Lewis [1988]: 323).

A well-known reply to this *argument from wanting* is that it fails to establish Humeanism about motivating states. In his seminal discussion, Thomas Nagel, for example, argued that the argument rests “on a confusion between two sorts of desires, motivated and unmotivated” (Nagel [1978]: 29). A motivated desire is one that can be rationalised in terms of a motivating reason. An unmotivated desire is one that can’t be rationalised. But as Nagel points out, from the fact that intentionally \( \varphi \)-ing necessarily involves wanting to \( \varphi \) it does not follow that it must involve an unmotivated desire to \( \varphi \). That is, although Nagel thinks “whatever may be the motivation for someone’s intentional pursuit of a goal, it becomes in virtue of his pursuit *ipso facto* appropriate to ascribe to him a desire for that goal” (Ibid.), he also thinks “the issue is whether another desire always lies behind the motivated one, or whether sometimes the motivation of the initial desire involves no reference to another, unmotivated desire” (Ibid.).

Nagel’s response, if successful, would undermine the force of the argument from wanting as it pertains to Humeanism about motivating states. However, the argument would still appear to support Humeanism about motivation. (As we have seen Nagel explicitly concedes that motivation implies desire). This is why some, like Jonathan Dancy, dismiss Nagel’s response as an “irrelevant refutation” (Dancy [2000]: 82). Given that Nagel’s response shows that the argument does not establish Humeanism about motivating states and since this doctrine has been the traditional focus of the debate between the Humean and the Anti-Humean, this judgment is too harsh. Nonetheless, it is undeniable that Nagel’s response would hail a victory of sorts for the Humean and that the Anti-Humean would, therefore, be well advised to consider whether there isn’t a better reply to the argument.

Instead of claiming that wanting to \( \varphi \) can refer to motivated as well as unmotivated desire, the Anti-Humean should reply that it refers to having a *pro-attitude* towards \( \varphi \)-ing. On this account, to say you must have wanted to \( \varphi \) if you \( \varphi \)-ed intentionally is just another way of saying there must have been something about \( \varphi \)-ing that appealed to you, something that attracted you and that you saw in a positive light, if your \( \varphi \)-ing is to count as an intentional action. It follows that the fact that intentionally \( \varphi \)-ing

132 See also Goldman [1970]; Lewis [1988]; Meikle [1974]; Smith [1994].
133 Nagel also writes that unmotivated desires ‘simply assail’ us whereas motivated desires are ‘arrived at by decision and after deliberation’. He does not seem to notice that this is an orthogonal distinction with much less philosophical depth (for an interesting discussion, see Schueler [1995] and Wallace [1990]).
134 See also McDowell [1998]; Wallace [1990]. A closely related response has been made by Philippa Foot. She writes that “the argument depends upon a ‘use of ‘desire’ which indicates a motivational direction and nothing more” (Foot [1978b]: 149). Likewise, Rowland Stout writes that “to say that I wanted to do it in that sense is to say no more than that I was motivated to do it” (Stout [2005]: 21).
implies wanting to \( \varphi \) supports Humeanism only if having a pro-attitude implies having a desire. Yet, having a pro-attitude just is being in a state that involves a positive evaluation. Yet, as I argued in chapter three, there are two kinds of evaluation: cognitive and affective evaluation. Consequently, one can have a pro-attitude simply because one is in a certain kind of cognitive state, such as believing that one has most reason to \( \varphi \).

The Anti-Humean, therefore, maintains that it is the intuition that intentional action implies the agent seeing something in a positive light – and not the intuition that intentional action implies desire – that underpins the force of the claim that intentional action implies wanting. Moreover, seeing something in a positive light can be a matter of belief alone. Yet, of course, a sense of wanting which is broad enough to be inclusive of some beliefs is too broad to be of any help for the Humean. Humeanism built on this foundation wouldn’t be a ‘trivial truth’ but a mere verbal one. The substantive issue is whether motivation and action imply desire, and not whether we have a term that refers equally to desires and some beliefs.\(^{135}\)

To defend the argument the Humean must deny that “wanting to \( \varphi \)” simply means “having a pro-attitude towards \( \varphi \)-ing” in the context of the claim that intentionally \( \varphi \)-ing implies wanting to \( \varphi \). However, such a denial faces three difficulties. First, the claim that intentional action implies having a pro-attitude is weaker than the claim that intentional action implies a specific pro-attitude, namely, desire. Moreover, once the two claims are distinguished it is far from obvious that the stronger claim is true and the Humean ought to provide an extra argument to establish this claim. Second, understanding wanting in terms of having a pro-attitude would seem to allow for an explanation of why it should be true that intentionally \( \varphi \)-ing implies wanting to \( \varphi \). Roughly, the explanation is that intentional action is action that is genuinely expressive of an agent’s agency and that an action can only be expressive of an agent’s agency if the agent herself sees something positive in it. The Humean, on the other hand, can’t explain why intentional action should have to imply wanting – that is, \( \text{desiring} \) – in the same way: for that explanation only supports the more general thesis that intentional action implies having some pro-attitude or other. Moreover, it is far from clear that the Humean could provide an alternative account to explain the intuitive power of the claim that intentionally \( \varphi \)-ing implies wanting to \( \varphi \) that would specifically draw upon desire. Finally, suppose a person does what she promised to do but says she did it despite having no desire at all for doing it. The Humean must say this person is mistaken. If she kept the promise intentionally, then she must have wanted, that is, \( \text{desired} \) to keep the promise. Interpreting ‘wanting to keep the promise’ in terms of the generic ‘having a pro-attitude towards keeping the promise’, on the other hand, has the advantage of allowing us to avoid having to ascribe false beliefs to the person.

\(^{135}\) In more or less explicit terms this response can be found in Locke [1982]; Nagel [1978]; Schueler [1995].
The upshot is that the Anti-Humean can accept that intentionally \( \varphi \)-ing implies wanting to \( \varphi \) but deny that it follows that intentionally \( \varphi \)-ing implies desire.

5.6 Does the Felt Need Theory Accommodate the Role of Desire in Action?

We started out with this objection: Desire is necessary for motivation and action. No state individuated by its relation to episodes of felt need is necessary for motivation and action. So, desire can’t be a state individuated by its relation to episodes of felt need. We are now in a position to reject this objection as being based upon the false premise that desire is necessary for motivation and action.

This is not to deny, though, that desire plays an important role in the explanation of action. Many actions really are explained by reference to an appropriate desire. Paul drank a glass of water because he desired to quench his thirst. Paula went on a holiday to Fiji because she desired a beach holiday. The Humean’s mistake is not to accord desire a significant role in the explanation of action but to think it plays a unique and necessary role.

What exactly is the role of desire in the explanation of action? We explain intentional actions by appeal to the reasons for which they were done, that is, by appeal to the agent’s motivating reasons. Desire can explain action because desire, in concert with an appropriate bridging-belief, can constitute a motivating reason. Put another way, we explain actions by revealing the favourable light in which the agent saw her action and desire can play an important role in the explanation of action because it involves a pro-attitude; to desire something is to see it in a positive light. By showing an action is aimed at securing the desired state we reveal in what way it was aimed at promoting something the agent saw in a favourable light.

This account of the role of desire in the explanation of action clearly depends upon the role of desire in practical rationality and its pro-attitudinal character. To explain the latter is to explain the possibility of action explanation by appeal to desire. Moreover, in chapter four I argued that it is a prime virtue of the felt need theory of desire that it can offer a fitting explanation of the pro-attitudinal character of desire and its rational significance. It follows that the felt need theory of desire can also explain the role that desire plays in the explanation of action.

The previous chapter ended with the objection that the felt need theory fails to do justice to the role of desire in the explanation of action. I conceded that if desire is necessary for motivation and action, this objection goes through. As we have seen, though, motivation and action don’t imply desire; desire isn’t the sole source of motivation nor is desire implied by motivation. The role desire plays is more limited. In concert with an appropriate bridging-belief it can constitute a motivating reason. And it can do so
because to desire something is to have a pro-attitude towards it. Since the felt need theory offers a forceful account of this, it has all the resources needed to do justice to the role of desire in the explanation of action.
Chapter 6

Desire and Reasons for Action

The rational agent looks after the things she sees in a positive light as well as she can. Inevitably, this means that she will concern herself a great deal with satisfying her desires. But does the significance of desire extend beyond the sphere of rationality? Do desires also give us (normative) reasons for action?

Many writers do not draw a firm distinction between talk of rationality and talk of reasons. However, as I shall use the notions, there is an important difference. Rationality supervenes upon the mind. In contrast, what we have reasons to do depends, at least in part, upon non-mental facts.

Evidently, what is rational need not be supported by reasons. It is rational for a person who thinks that the clear liquid is gin and who desires to drink a gin and tonic to drink that stuff, but she has no reason to do so if it is in fact petrol.136 Even when the two overlap, however, they may not have the same source. It is rational for someone to apologise if she believes that she has most reason to do so. If this belief is true, she will really have a reason to apologise. But while the act is rational in virtue of the belief, the reason is provided by whatever makes the belief true.

It has long been a corner stone of ethical theorising that desire is not only relevant to the rationality of action but that it can also provide us with reasons for action. In fact, many have maintained that desire is the sole source of practical reasons. On the other hand, in recent years the view that desire never provides us with reasons has gained a lot of traction and is fast becoming the new conventional wisdom.137 One aim of this chapter is to clarify the dispute and to argue for the existence of desire-based reasons. The other aim is to show that the felt need theory has the explanatory resources to account for these reasons.

The plan for this chapter is to begin by considering the case for denying desire-based reasons. In the first section, I shall acknowledge a couple of distinctions that are frequently appealed to by critics of desire-based reasons. Drawing these distinctions lends a measure of credibility to the denial of desire-based reasons, but it does not establish that position. In the second section, I shall, consequently, look at three

136 Williams [1981].
137 E.g. Audi [2002]; Bond [1983]; Cullity [2003]; Dancy [2000]; Heuer [2004]; Parfit [2001]; Quinn [1993]; Schueler [1995]. Some, like Raz [2001b] and Scanlon [1998] share the general tenor of this, though feel pressed to admit that there are some rather exceptional and ‘peculiar’ cases in which desires can give us reason. For the view that desires can give us reasons see, for example, Brandt [1998]; Chang [2004]; Fehige [2001]; Lewis [1989]; Stampe [1987].
arguments that can be made to bolster the claim that desires never give us reasons. Close examination of these arguments will show that none is successful. In the third section, I shall argue the positive case for desire-based reasons. Once the plausibility of desire-based reasons has been restored, it will be an easy undertaking in the fourth and final section to show that the felt need theory can explain the truth of desire-based reasons.

6.1 Desire and Reasons: Two Distinctions

On the face of it, the denial of desire-based reasons is barely credible. A person who desires to spend more time gardening surely has a reason to do so. Someone who wants to learn Italian surely has a reason to take lessons or to go and live in Italy. Someone who longs to single-handedly sail around the world surely has a reason to take steps towards fulfilling this dream. Who could possibly want to deny this? Upon reflection, however, a more nuanced picture emerges. If the right distinctions are drawn, the denial of desire-based reasons need not be too offensive. In this section, I shall explain what these distinctions are that form the foundation for the denial of desire-based reasons.

The first distinction is between (i) having a reason and having a desire, as opposed to (ii) having a reason because one has a desire. Derek Parfit, for example, admits “we usually have some reason to fulfill our desires” (Parfit [2001]: 19) but adds “we would have this reason even if we didn’t have this desire” (Ibid.). No doubt, this distinction is easily overlooked because we are naturally drawn to those things that give us reasons. For example, we desire pleasure and seek to avoid pain. Yet, it can be said that if we would lack these desires we would still have reasons to seek pleasant experiences and to avoid painful ones. Consequently, these reasons can’t be provided by desire.

Not everyone accepts this distinction. Practical reasons must have a genuine claim over the agent and internalists maintain that we can’t understand how this could be possible unless having a reason is, at least in part, a matter of facts ‘internal’ to the agent; that is, a matter of what the agent identifies with or cares about. Desire-internalists hold that what an agent identifies with, or cares about, is determined by her desires. They, therefore, maintain that an appropriate desire is necessary for having a reason. If desire-internalism is true, we will have failed to make the denial of desire-based reasons any more palatable. But we may well doubt that we would have succeeded in doing so even if desire-internalism were false. The problem is that, while it is difficult to say whether all reasons depend upon desire, it should be uncontroversial that some do. For example, it would be very odd to say everyone has a reason to learn Italian, but there is nothing odd in supposing that someone who desires to do so has such a reason.

---

138 See esp. Foot [1978a]; Williams [1981]; Williams [1995]. Internalism is often equated with desire-internalism, but it is unclear if anyone would want to genuinely espouse the latter. Williams [1981], for example, makes it clear that he is using ‘desire’ in a mere ‘formal’ sense, meant to include such things as ‘dispositions of evaluations’ or ‘believing that a particular consideration is a reason to act.’ (See also Foot [1978b]).
Indeed, even a staunch critic of desire-based reasons like Tim Scanlon concedes "many of the reasons people have depend on their desires" (Scanlon [1998]: 41) and "different people have different reasons for action because their desires are different" (Ibid.). Therefore, if the dependence upon desire were at stake in the dispute over desire-based reasons, a denial of desire-internalism would win the battle, but lose the war.

This shows that the case against desire-based reasons must be built upon a different foundation. This is where the second distinction comes into play: the distinction between (i) what a reason depends upon and (ii) what provides us with a reason. That there should be such a distinction has a great measure of plausibility. Suppose that, Albert has a reason to listen to *Cosi Fan Tutti*. This reason depends upon all sorts of factors. It depends upon Albert’s existence, on Albert’s being a rational agent and not a dog, on Albert’s ability to hear music, on *Cosi Fan Tutti* being an opera of great musical value and not a recording of the fantasies of a paedophile, and so forth. Despite the reason’s dependence upon all these factors, it is implausible to say the reason is provided by each or all of these.

The general shape of the distinction is, in fact, ubiquitous in our thought. We often find it important to distinguish between a role-playing feature and an enabling condition. Even though the role-playing feature can play its role only against the background of the enabling condition, it commands a special role. It was the impact of the stone, we think, that caused the glass to break even though a great many other conditions needed to be in place for the impact of the stone to play its causal role. Similarly, snow being white makes ‘snow is white’ true, but does so only against a background of a set of enabling conditions. And likewise, we think that although something can play the reason-providing role only if other conditions obtain, this does not undermine the special importance of whatever it is that does play the reason-providing role.

Now it might be said that while we frequently make the distinction between a role-playing feature and an enabling condition, the distinction does not correspond to a distinction in the nature of things. If so, the case against desire-based reasons is doomed, since some reasons evidently depend upon desire. Still, while it is perhaps difficult to offer a defence of the distinction between a role-playing feature and an enabling condition, it is natural to assume that some such defence should in principle be possible. If critics of desire-based reasons need nothing more than this distinction, their case would begin to look much more credible indeed.

Then again, to establish that desire never provides us with reasons, we need to go beyond drawing the distinction between what our reasons depend upon and what provides us with reasons. We must also show that desires inevitably fall on one side of the divide. There are three major arguments upon which one might draw to establish this. I shall consider these in turn.
6.2 Three Arguments Against Desire-Based Reasons

6.2.1 Desire-Based vs Value-Based Reasons

Those who think desire never provides us with reasons for action typically think the stakes are high and the choices stark: either, they say, all reasons are provided by desire, or they are all provided by value. Parfit, for example, writes, “according to desire-based theories, practical reasons are all provided by our desires or aims. According to value-based theories, these reasons are provided by facts about what is relevantly good, or worth achieving” (Parfit [2001]: 38).\(^{139}\) Parfit himself goes on to argue that not all practical reasons are provided by desire and, therefore, none are. Many would want to agree with Parfit that not all practical reasons are provided by desire (though—perhaps—not everyone would want to do so for the reason Parfit gives). For instance, it is intuitive to think we have reason to seek pleasure even when we lack the desire to do so. Likewise, even when our reasons depend upon desire, it is often intuitive to think the ultimate ground for them lies elsewhere. Returning to the case of Albert, even if we assume that Albert would not have had a reason to hear *Cosi Fan Tutti* if he lacked any desire to do so, it is natural to say that the ultimate ground or basis for his reason is not that he *desired* to hear it, but rather that the music is of great beauty or that he would take great pleasure in listening to it. Or consider the following case: Suppose that Albertina has a reason to throw a life-raft to a drowning person. Surely, it is much more plausible to think this reason is provided by, or based on, the fact that doing so will save the person’s life rather than to think it is provided by some desire of Albertina, even if some such desire should turn out to be necessary for her to have the reason. But given the stark choices guiding the dispute, these intuitions seem to push us right into the camp of value-based theories, and, therefore, towards the conclusion that desires never provide us with reasons.

Perhaps we should not accept the assumption that either all reasons are based on desire or on value. Certainly not everyone agrees that there is a need for this stark setting of choices. Ruth Chang, for example, argues for a hybrid position according to which some reasons are based on desires while others are based on values.\(^{140}\) At first blush, this move can seem very reasonable. Faced with the extremes, it is tempting to think that “the truth may lie somewhere in the middle” (Chang [2004]: 57). Moreover, Chang’s ecumenical approach to the phenomena does afford considerable explanatory flexibility. For all that, the proposal is really unacceptable. The assumption that either all reasons must be based on desire or in value does reflect the genuine insight that we require a unified account of the source of practical reason. Per contra, if Chang were right, we would need to tell two radically distinct stories about practical reasons. Some reasons, it would turn out, have their source in one thing (desire) whereas others have their source in something completely different (value). It is hard to say whether this

---

\(^{139}\) See also Audi [2002]; Bond [1983]; Dancy [2000]; Quinn [1993]; Raz [2001b].

\(^{140}\) Chang [2004].
suggestion is even intelligible. It boggles the mind to understand how there could be a unified domain of practical reasons if some reasons required one kind of theoretical explication whereas others stood in need of a totally separate, and radically distinct, theoretical account.\footnote{141 Chang [2004] recognises the challenge but argues that her hybrid view is compatible with ‘unity at a deeper level’ because “all practical reasons are provided by facts with a single source – they are provided by facts about the action or its object” (Chang [2004]: 90). This equivocates on what is at stake. We can describe all reason-providing facts as ‘facts about the action or its object’ (e.g. we can say what provides the reason is the fact that the action satisfies a desire or that it is valuable), but this does not alleviate the need to tell two radically distinct stories about the source of these reasons, that is, two radically distinct stories in virtue of what these facts can play their reason-providing role. On Chang’s account, in one case this story must end with desire and in the other case it must terminate in value without a bridge between these.\footnote{142 See Jackson [1998] for an interesting argument in favor of the identity view and Moore [1988] and Ross [2002] for forceful expositions of the supervenience account. Another view that has gained a lot of traction lately is Scanlon’s buck-passing account (Scanlon [1998]: 97). On this account, goodness is identified with the higher-order property of having some properties that provide reasons of the relevant kind.}

If we should nonetheless resist the conclusion critics of desire-based reasons try to press upon us, it is due to a more subtle point. Let us consider the claim that all reasons are provided by ‘facts about what is relevantly good’ a bit more closely. It is a conceptual truth that nothing can be good without having some property in virtue of which it is good. A thing can’t – logically can’t – just be good. There must necessarily be something that makes it good. But what is the relation that holds between something being good and its having a good-making feature? On one view, this relation is one of \textit{identity}. On another weaker view, it is one of \textit{supervenience} alone.\footnote{142 Suppose we accept the identity thesis. In that case, it makes at most a terminological difference to say reasons are provided by goodness or by good-making features. If the two are not identical, though, there are (at least) two options for what might provide reasons: the good-making feature or the goodness it determines.} Suppose we accept the identity thesis. In that case, it makes at most a terminological difference to say reasons are provided by goodness or by good-making features. If the two are not identical, though, there are (at least) two options for what might provide reasons: the good-making feature or the goodness it determines.

It would take us too far afield to try to resolve the question of whether goodness is identical to, or merely supervenient upon, good-making features. But if the two are distinct, it seems much more plausible to suppose that reasons are provided by the good-making features. This conclusion, in any case, is reached by a number of critics of desire-based reasons. Parfit, for instance, says that “reasons are provided by facts about what is relevantly good or worth achieving” (Parfit [2001]: 18) but also that he is “inclined to believe” (Parfit [2001]: 20) that it is “natural properties” (Ibid.) that provide us with reasons and that “goodness is not itself a reason-giving property” (Ibid.)! And we must admit that there is a good measure of plausibility to this. On the one hand, taking good-making features to provide us with reasons better matches our linguistic intuitions: it is more natural to say that what provides Albert with a reason to hear \textit{Cosi Fan Tutti} is that it has great beauty, not that it is good. Further, it is plausible to think since good-making features provide the basis for goodness, they should also provide the ultimate foundation for our reasons.
This is significant because it shows that an extra premise is needed to draw the conclusion that desire never provides us with reasons from the assumptions that all reasons are based on desire or in value and that some reasons are not provided by desires. What it means to say that all reasons are based on value is that all reasons are based on good-making features. Patently, that could be true, and, at the same time, some reasons could be based on desire if desire is one of the determinants of value. Until this possibility has been ruled out, we can't infer that no reasons are based on desire just because not all reasons are so-based and there must be a unified account of the sources of practical reasons.

I shall conclude my discussion of this argument by briefly considering two objections. First, it might be said that taking good-making features to be the ultimate basis of our reasons for action violates the demand, called 'non-negotiable' not too long ago, that there be a unified source of practical reasons. So, we must conclude that it is goodness that provides us with reasons and that desire could not do so even if it were a determinant of value. In response, we should note that this would only follow if goodness and good-making features were distinct from one another. Moreover, we should note that this would also have the consequence that no good-making feature – whether pleasure, beauty, or justice – could ever provide us with reasons. But in any case, the objection misconceives what is involved in requiring a unified source of practical reasons. Most people would admit that there could be distinct kinds of reasons –, e.g. pleasure-based reasons and beauty-based reasons, or reasons based on the experiential value and reasons based on aesthetic value. Even so, this need not undermine the possibility of a unified account of practical reasons, as there would indeed be a common thread tying these reasons together – they would all be united by their relation to goodness. On the other hand, contrast this with a true hybrid proposal, where reasons are taken to stem from two radically distinct sources that are not united by a common denominator. The latter would require providing two non-overlapping accounts of the sources of reasons, but the former would not. This is what makes the suggestion that reasons could have their source in two radically distinct things unacceptable in a way that the idea that there are multiple sources of a common kind is not.

Second, it might be thought that this attempt to salvage desire-based reasons is missing the point. Thus Chang writes that

there is a sense in which all reasons could be 'provided' by desires even on a value-based view if, for example, all reasons are provided by evaluative facts about what makes one's life go best, and those evaluative facts in turn reduce to facts about what one wants. But the claim that certain evaluative facts reduce to facts about what one wants is an extraneous substantive claim that holds independently of the thesis that those evaluative facts
provide reasons. Put another way, when we ask whether a desire can provide reasons, we are asking whether a desire can provide reasons in virtue of being a desire, and not in some less direct way. (Chang [2004]: 59)

This does not seem to be correct. It is true that we can consider whether it is the good-making feature or goodness itself that provides us with reasons and that we can settle that question without considering the ‘extraneous substantive’ issue of what those good-making features are. But that does not mean that in order to show that desire can provide us with reasons we must show that it is not a determinant of value. If it meant that, we would also have to show that, say, pleasure is not a determinant of value in order to show that it can provide us with reasons. Surely, this demand is too onerous in either case. To be fair, historically, many have combined allegiance to desire-based reasons with enmity towards all talk of good-making features and values. But it would truly be ‘extraneous’ to burden the champion of the possibility of desire-based reasons with this further thesis.

6.2.2 Reason-Based Desires and Desire-Based Reasons

Critics of desire-based reasons frequently charge their opponents with getting things the wrong way around. Desires, they say, are themselves ‘had for a (perceived) reason’. They are an ‘endorsement’ of the reasons we already take ourselves to have. But they do not add to these.143

It is plausible to think some of our desires are had for a reason, but is this true of all desires? It might be said that this must be true on pain of severing what a person desires from what she has a pro-attitude towards. Thus, critics sometimes allow that there could be ‘desires’ that are not had for a reason but insists that these are ‘mere urges’ that the agent ‘entirely disavows’ and in which ‘the agent herself sees no good whatsoever’.144 This does not seem right, though. Suppose we accept that desire implies seeing some good in the desired end. It does not follow that desire must always be had for a reason. After all, one might desire something and believe it to be good, but not have the desire because of the belief. Indeed, one might even have the belief because of the desire. Suppose that there are desire-based reasons. Then one could desire a state and believe one has a reason to promote it because it will satisfy the desire. Further, some of our reasons are causally dependent upon our desires. For example, sometimes we only reap enjoyment from activities that we also desire to do. Even if our reason should be provided by enjoyment, having that reason depends upon having the desire. Again, in that case it would be natural enough to believe one has a reason to engage in the activity only because one introspects a desire for it.

143 See Bond [1983] Ch.2&3; Cullity [2003]; Dancy [2000] pp. 35-38; Quinn [1993]; Raz [2001a]; Raz [2001b].
144 Dancy [2000]: 36 puts this point especially vividly but see also Quinn [1993] and Raz [2001c]). Needless to say, such ‘desires’ would be bad candidates for reason-providing states.
The argument also has another flaw. Suppose that all desires are had for a reason. It does not follow that desire never provides any reasons. Desire could, after all, be both a response to and a source of reasons. One objection is that this would be double-counting, or worse, simply inconsistent. But this objection gets a grip only if desire were to be thought of as a response to, and a source of, one and the same reason. It loses its force once we recognize the distinct possibility that desire is a response to one kind of reason and a source of another. To illustrate, consider the case of pleasure. Suppose my pleasure in reading a book is based on the belief that the book is of great literary value. This does not mean that my pleasure in reading the book does not provide me with any reasons to read the book over and above the reasons I have in virtue of the great literary value of the book. Clearly, my pleasure matters too, even if it is a response to a cognised value. The reason my pleasure matters is that it contributes to my welfare in a way that the mere value of the book does not. In the next section, I shall argue that desire likewise matters to our welfare. If that is correct, then desire could give us reasons even if desire would spring forward in response to the recognition of certain values. The argument that desire can’t provide us with reasons because it is itself ‘had for a reason’ consequently fails to convince.

6.2.3 Desire is Not Sufficient for Having a Reason

Many people, I suspect, are hostile to the idea that desire can provide us with reasons because they think it so glaringly obvious that there are cases in which a person desires something but has no reason whatsoever to promote the desired end. Many people, for example, have this response when asked to consider a person who desires to count all the blades of grass in her garden. Surely, they say, that person has no reason to do so. Others are moved to embrace the same conclusion by reflecting on cases involving a desire for something morally repugnant. Suppose, for instance, that someone desires to torture an innocent person. I would say this person has no reason to do so, her desire notwithstanding. Cases like these seem to prove that desire itself can never provide us with reasons.

Not everyone shares these intuitions. However, even if one does accept the starting point of this argument, one need not embrace the conclusion. To say that desire itself can’t provide us with reasons is ambiguous. It can mean that desire by itself can’t provide us with reasons; that is, that desire alone is not sufficient to provide us with reasons. Or it can mean that it is never desire that provides us with reason. The example shows that desire alone is not sufficient to provide us with reasons. Put another way, it shows that to have a desire is not necessarily to have a reason. But it does not show that it is never desire that provides us with reasons. It is all too easy to conflate these ideas, but they really must be kept apart: I have no reason to torture the innocent even if I would get immense pleasure from doing so. This shows that pleasure by itself does not provide reasons. But it does not show that pleasure can never provide us with reasons. Clearly, sometimes our reasons are grounded in pleasure. Nor is there anything paradoxical in this. Recall, in the first section, we distinguished between
an enabling condition and a role-playing factor in our thought about reasons. This distinction is just as relevant here as it was there. For it shows that we can’t infer that something is incapable of playing the reason-providing role from the fact that it does not do so necessarily.

So, even if desire does not always give us reasons, it may well sometimes do so. Put another way, desire may play the reason-providing role only against a background of enabling conditions, but this is consistent with desire itself sometimes occupying the reason-providing role. Finally, note that even though the validity of the distinction between an enabling condition and a role-playing feature can be called into question, doing so would not help critics of desire-based reasons. For as we saw in the first section, without this distinction in place, scepticism about desire-based reasons is doomed to fail. I conclude, therefore, that we can’t infer that desire never provides us with reasons from the observation that it does not always do so.

6.3 The Case for Desire-Based Reasons

It is time to make the positive case for desire-based reasons. My argument is that desire can provide us with reasons because having our desires satisfied can in and of itself add to our welfare.

Our thought about value has two dimensions: a world-centred and a subject-centred one. Some things are valuable in the sense that they make the world a better place. Others are valuable in the sense that they make a subject’s life go better. Although there are deep connections between world-centred and subject-centred value, neither can be reduced to the other. Some things, such as the satisfaction of a killer, are good for a subject while adding nothing and perhaps even taking away from the value of the world. Other things add to the overall value of the world without benefiting any subject. An act of justice, for example, need not make anyone better off but, for all that, can still have world-centred value. Making a sacrifice to help the deprived will also add to the value of the world in a way that can’t be fully captured merely by looking at subject-centred values. If it did, then a world in which a sacrifice is made to help the deprived would have to be worse than one in which the deprived simply stumble upon the same benefit, yet, intuitively, it is a better. Again, this is not to deny the deep connections that exist between world-centred and subject-centred value, but only to caution against a straightforward reduction.

Now it seems clear that if desire is a determinant of value, it will, at least primarily, be a determinant of subject-centred value. Let us, therefore, consider the idea of subject-centred value a bit more closely.

145 Nor is this entirely due to the harm suffered by the victims. Suppose Richard feels deep satisfaction because he thinks that he has killed his brother. This, it seems, suffices to make his life better but the world worse even if Richard is mistaken in thinking that he has killed his brother.
Subject-centred value is the value of a life going well – not morally, aesthetically, or epistemically, but well for the person whose life it is. Subject-centred value is irreducibly subjective in the following sense: It is necessary for something to have subject-centred value that it has some value for the subject. A person can live amongst rubbish or beauty but if she is wholly indifferent to this, it will be irrelevant to how well her life is going. She can make great scientific discoveries or prevent nuclear war. In doing this, she will have made the world a better place but, once again, if none of this makes any difference to her, then it will be irrelevant to the quality of her life.  

A paradigm of subject-centred value is experiential value. Some experiences have a positive quality – they feel good or pleasant. Having these qualities makes the having of these experiences good for the subject. All other things being equal, a life richer in such experiences will be a life that is greater in welfare; that is, better for the person whose life it is. Other experiences have a negative quality. Consider, for instance, being depressed or being in agony. All other things being equal, having such experiences makes for less welfare.

Experiential value isn’t the only source of subject-centred value. Another source is getting what is of importance to us. Often these two coincide and we get what is important to us and are happy about that. But they need not coincide. Suppose Leon suffers from an incurable disease. In the shortness of time that is still left to him, he wants to fulfil his dream to single-handedly sail around the world. The voyage itself has its experiential ups and downs. There is much beauty to be enjoyed, but there are also vast stretches of fear, loneliness and pain. Moreover, upon reaching the end of his journey Leon is too exhausted to feel happy about having realised his dream and even a little saddened by the knowledge that he will never set sail again. But for all that, it seems Leon’s life is better for his having completed the voyage and those who care about him will feel gratitude that he still had the opportunity to accomplish what was so important to him. Feeling joy and happiness would have been a “welcome dividend” but in the end what matters is that he got what was of importance to him and not that he got an experiential pay-off.
The distinctive relevance of getting what is of importance to us is also salient in cases when a person doesn’t get what is important to her but is happy nonetheless. Suppose that Amie is a sensitive young woman who places considerable importance upon having authentic rather than fake friends. Suppose further that those who seem closest to her in fact secretly ridicule and loathe her. As Amie is unaware of this, it does not impact upon her happiness. Nonetheless, it seems to make her life go much worse. Of course, to say this is not necessarily to say it would be better for her if she had known the truth about her supposed friends – that knowing an unpleasant truth is always better than ignorant bliss. Whether this is so is a distinct issue. The relevant point is just that her life would have been better if she’d had real rather than fake friends.

For all this, it may be too simple to say that securing what is important to us necessarily makes us better off. Various complications come to mind, and the thesis might have to be qualified to take these into account. For example, it might be said that only what is suitably central to a person’s life can impact upon its welfare, but that it is possible that something is important to us and yet not particularly central to our life. If so, then we will need to restrict the thesis so that it applies only when what is important to us is also suitably central to our life. Another problem that some people see is that getting what is important to us might leave us feeling deeply disappointed and unsatisfied and that in these cases we would be worse off, not better. Again, if we accept this, we should qualify our thesis even further to exclude cases in which such undermining negative emotions are present. Finally, there are the controversial cases of the worthless or even morally repugnant being of importance to a person. Could securing that ever make a person better off? Once more, we need not take a stance on this but can simply note that if it can’t, then securing what is important to us will only play the good-making role relative to a set of enabling conditions that excludes the utterly worthless and the morally repugnant.

At this point, however, some will object that by building in a requirement of minimal world-centred value we are effectively getting rid of the relevance of getting what is important to a person altogether; that the enabling condition itself has come to do all the work. I don’t think this is right. For one, only if something has value for a subject can it have subject-centred value. The emphasis on what is important to a subject satisfies this criterion; an account purely in terms of world-centred value

149 For some important discussions of these complications see Griffin [1988]; Kagan [1994]; Parfit [1987]; Sumner [1996].

150 Note, this doesn’t imply that getting what is important to us will only make us better off if we also feel better. Note further, there are a number of complications here. First, it may be said that disappointment is evidence that the agent misidentified what it was that really mattered to her. Second, it may also be evidence that a change in what is important occurred. Finally, a person may feel unsatisfied for extrinsic reasons, such as, having a disposition to feel unsatisfied. Here, it strikes me, getting what is important to the person still contributes positively to her welfare, although this contribution can be outweighed by the negative contribution of this extrinsically related experiential quality.
doesn’t. For another, all other things being equal, a subject will be better off if she gets what is more important to him. For example, if two things are of equal world-centred value, then, assuming no difference in experiential pay-offs, a person will be better off by getting that which is of greater importance to him. This shows that, even if a requirement of minimal world-centred value should apply, what is important to a subject has an irreducible role to play as a source of welfare.

How does all this relate to desire? Desire is connected to both these sources of subject-centred value, though not in the same way. Desire can play a causal role in bringing about experiential value. We often only enjoy something if we had a prior desire for it. Also, getting what we want can itself be something in which we take pleasure. But desire does not constitute experiential value. Per contra, desires invest things with importance for us; they come to matter to us in virtue of our desire for them. Here the relationship is constitutive and not merely causal or intentional.

It would be misguided to respond by saying that desire doesn’t always invest things with importance for us, that some desires that are too weak or fickle to do so and that we can be alienated from some of our desires. This is true but beside the point. What matters is not that desire necessarily makes things important for us, but that it can do so. What matters is that desire is the kind of state that can take an object to which we were hitherto indifferent and make it have great importance for us. (I am also not committed to the claim that everything that has importance for us does so because of our desires).

This is the key to desire-based reasons. Because desire can invest things with importance for us, having a desire satisfied can make us better off – not just indirectly by making us feel better as a result, but also directly, simply because our desire has been satisfied. To be sure, having a desire fulfilled will not always make us better off; it will do so only if a range of enabling conditions obtain. But because desiring something can make it have importance for us and because getting what is important for us can in and of itself make us better off, the satisfaction of desire can in and of itself add to our welfare. In this way, desire, or the fulfilment of desire, functions as a determinant of subject-centred value. And it is because desire can play this role that it can provide us with reasons.

Critics of desire-based reasons frequently contrast desire with pleasure. The latter they think can provide us with reasons, the former can’t. If what I argued is correct, then there is no such asymmetry between desire and pleasure. Both can function as a determinant of subject-centred value; both can be a source of our welfare. And it is because they can play this good-making role that both can provide us with reasons for action.
Finally, note how this account of desire-based reasons accommodates the insights of the preceding sections. First, the case for desire-based reasons does not rest solely on the observation that we often have reason to fulfill our desire or that our reasons for action commonly depend upon desire. Rather, it is built upon the claim that desire is a source of subject-centred value. Second, the present account of desire-based reasons allows us to deny that all reasons are based on desire but also to maintain that there is a single unified source of practical reasons as it allows us to uphold that all reasons for action are based on good-making features. Third, because of the significance of desire to subject-centred value and the irreducible subjective character of subject-centred value, desire (and pleasure) can provide additional reasons for action even when the desire itself is had for a reason. Finally, not every desire will provide us with reasons. Only a subset of desire invests things with importance for us, and getting what is important for us determines subject-centred value only relative to certain enabling conditions. Yet, while this shows that desire can only play the reason-providing role relative to a set of enabling conditions, the significance that desire has for our welfare ought to persuade us that it is the kind of state that genuinely can play that role.

6.4 The Felt Need Theory Explains Desire-Based Reasons

In *What We Owe to Each Other*, Scanlon explains that he rejects the 'standard desire model' according to which desires “are states which simply occur or not, and when they do occur they provide the agent with reason to do what will promote their fulfillment” (Scanlon [1998]: 43) because he thinks “none of the candidates for the role of desire has these properties” (Ibid.). Following Quinn, for example, Scanlon argues that the state of being disposed to act does not provide us with reasons for action and that, consequently, if desires are just functional states that dispose an agent towards action, they can’t be a source of reasons. Of course, on Scanlon’s own view, desires are more than functional states that dispose an agent towards action. Thus, Scanlon writes that a person has a desire that *p* “if the thought of *p* keeps occurring to him or her in a favorable light, that is to say, if the person’s attention is directed insistently toward considerations that present themselves as counting in favor of *p*”. (Scanlon [1998]: 39). Yet, Scanlon argues that even on this construal, desires do not provide us with reasons because a state that insistently directs our attention towards considerations that present themselves as counting in favor of something is not the kind of state that can provide us with reasons for action.

Suppose that, as sometimes happens, I am beset by a desire to have a new computer. What does this involve? For one thing, I find myself looking eagerly at the computer advertisements in each Tuesday’s *New York Times*. I keep thinking about various new models and taking their features to count in favor of having them. This is what I called above a desire in the directed-attention sense. It has clear normative content, since it involves

Desire and Reasons for Action 118

a tendency to judge that I have reason to buy a new computer. But does my being in this state make it the case that I have a reason to buy a new computer (because doing this would satisfy my desire)? It seems to me clear that it does not. Such a state can occur (indeed, it often does) even when my considered judgment is that I in fact have no reason to buy a new machine, since I believe (correctly, let us suppose) that the features of the newer models would be of no real benefit to me. In such a case the fact that I have this desire gives me no reason to buy a new computer (aside, perhaps, from the indirect one that it would put an end, for a time, to my being nagged by the desire and wasting time reading computer advertisements). It is not just that the reason provided by the desire is outweighed by other considerations. I would not say, “Well, I do have some reason to buy the computer, since it would satisfy my desire, but on balance it is not worth it”. The desire, even if it persists, provides no reason at all (except possibly the indirect one just mentioned).

Now suppose that I endorse the judgment to which the desire involves a tendency, and take myself to have good reason to buy a new machine. Even in this case, the reason that I have for buying a computer is not that it will satisfy my desire, but rather that I will enjoy having it, or that it will help me with my work, impress my friends and colleagues, or bring some other supposed benefit. (Scanlon [1998]: 43-44)

Even if Scanlon were right in concluding that none of the theories of desire that he considers allow us to make sense of desire-based reasons, it would not follow that ‘none of the candidates for the role of desire’ allows for a vindication of the commonsensical thought that desiring something does indeed, at least sometimes, provide us with a genuine reason for action. This is because Scanlon fails to appreciate the possibility of affective theories of desire. In particular, he overlooks the possibility of characterising desire in terms of episodes of felt need.

The felt need theory of desire does seem exceptionally well placed to explain the deep-seated intuition that desire can provide us with reasons for action. Desires can provide us with reasons for action, I argued, because desires can invest things with importance for us. To explain desire-based reasons, a theory of desire must, therefore, allow us to understand how desiring something can invest it with importance for the subject. The felt need theory does this because it characterises desires in terms of experiences of felt need for the object of desire, and it seems that having such experiences

---

152 Scanlon does argue that pleasure can provide us with reasons for action and that we frequently desire things that we will find enjoyable. This, it might be thought, would allow him to press the same kind of objection against some hedonic theories of desire. Perhaps, but even so this would only extend his claim so as to cover some affective theories of desire.
can make it the case that getting the desired end is something that has importance for the subject. That is, it seems that some things being of importance for the subject need not consist in anything more than the subject undergoing experiences of the kind that stand at the centre of the felt need theory of desire.

This is not to say that having an experience of felt need is sufficient for something to be of importance to a subject. If these experiential episodes are too fickle or weak, or if the agent is alienated from them, then they may not suffice (desires that are fickle and weak, or desires that the agent is alienated from, however, may likewise fail to provide us with reasons for action). Nor is it to say that experiences of felt need are necessary for some things being of importance to a subject. Perhaps, for example, some things are of importance to a subject in virtue of the subject’s believing them to be of great value. Still, it does seem plausible to suppose that when something is of great importance to a subject this frequently involves nothing more than the subject undergoing experiences of felt need. To illustrate, consider once more the case of Leon. It seems that we can come to understand how single-handedly sailing around the world is of such great importance for Leon simply by coming to think of him as a person who experiences the need to do so; that is, as a person to whom it seems that his single-handedly sailing around the world is something that must become reality.

Desire can provide us with reasons for action because desiring something can invest it with importance for us. And desiring something can invest things with importance for us precisely because desire is essentially linked to episodes of felt need. If this is correct, the felt need theory does not just offer a neat account of the role of desire in rationality and the explanation of action, but also of how desire can be the kind of state that can provide us with reasons for action. This is a truly impressive result and further supports the main conclusion of this thesis that the felt need theory of desire offers the best explication of the nature of desire.
I have defended the thesis that desire is a state individuated by its relation to distinctive conscious episodes. More precisely, I argued that a subject S (non-instrumentally) desires that $p$ if and only if from time to time S has a felt need for $p$ to become reality.

Theories that aim to analyse desire in terms of conscious experiences face a number of objections, and the felt need theory is no exception. However, I argued that most of these objections are based on misunderstandings, such as the assumption that the felt need theory implies that desire necessarily involves certain sensations, the assumption that the felt need theory identifies desire with a conscious state, or the assumption that the felt need theory implies that we can't have false beliefs about our own desires. The most forceful objection to the felt need theory, on the other hand, fails for a different kind of reason. This is the objection that all actions are motivated by desire, but that not all actions are motivated by a state individuated by its relation to episodes of felt need. This objection, I argued, fails because not all actions are motivated by desire.

In the introduction I said that a successful theory of desire must unify our pre-reflective conception of desire. That is, it must capture what underlies our intuitions about specific cases. Now, according to the felt need theory we judge that a person has a desire because we judge that from time to time she experiences a felt need for something to become reality. On the other hand, the theory implies that we judge that someone lacks a desire for something because we judge that she altogether lacks any felt need for it. I argued that it is indeed plausible to regard these claims as making explicit the structure of our pre-reflective intuitions about specific cases.

The felt need theory also fares well with respect to the second desideratum singled out in the introduction: namely, that a theory of desire must respect our intuitions about the kind of state desire is. For example, we intuitively think desire is not subject to the norm of truth, and the felt need theory is consistent with this. A related point is that the felt need theory also fits nicely with the common intuition that desire is a state with which the world must fit. Indeed, it is plausible to think that for desire to be a state with which the world must fit is nothing more than for desire to be a state that involves the felt need for the desired thing. Furthermore, the felt need theory coheres well with the deep-seated intuition that desire is a state that can be unconscious, and that there can also be genuine episodes of desirous consciousness. Moreover, it seems to be consistent with the plausible view that we are in a privileged epistemic position with respect to our own desires but also that judgments about our
own desires aren't infallible. And, of course, the felt need theory does allow for the possibility of recalcitrant desires and is consistent with there being desirous creatures that have very limited cognitive capacities.

The final desideratum for a successful theory of desire is that it helps us understand how desire can play important theoretical roles in practical rationality, the explanation of action, and (normative) reasons for action. I argued extensively that the felt need theory passes this test with flying colours. For one, I argued that in order to explain the role of desire in practical rationality we must explain the pro-attitudinal character of desire, that only an affective theory of desire can adequately explain this, and that the felt need theory is the best affective theory on offer. Moreover, I argued that the role of desire in the explanation of action is parasitic upon the role of desire in practical rationality and that, therefore, to explain the latter is to explain the former. Last, but not least, I argued that the felt need theory can also explain that desires can provide us with (normative) reasons for actions.

Given this impressive record the felt need theory of desire clearly must be reckoned with. Indeed, I think we are well-justified to conclude that a subject S (non-instrumentally) desires that $p$ if and only if from time to time S has a felt need for $p$ to become reality.
I have said that if it is impossible to hold beliefs with contradictory content, then there is another argument to the conclusion that N-desires can't involve a disposition to be motivated in ways believed to promote the desired end. Now it is surely possible for a person to believe that $p$ and also to believe that not-$p$ while remaining unaware of the tension between these beliefs. It is also possible for a person to think a sentence is expressive of what she believes and the sentence is, in fact, contradictory, that is, meaningless. But I doubt that it is literally possible to have a (single) belief with contradictory content. The problem with this is that we just can't get our head around what could possibly be believed in that case. It is easy to say a person can believe that two plus two is five or that she can draw a round square. But when we reflect upon what this person is said to believe we must fail because we can't conceive of what the person is said to regard as true. Put another way, the point is that contradictions resist being entertained. They resist being entertained because to entertain is to entertain something. But a contradiction is not something. Rather it is a pure nothing. When put to the task of grasping a contradiction, our mind must necessarily draw a blank. Indeed, this makes plain that the problem has nothing in particular to do with belief at all. It is in general exceedingly hard to make sense of the idea that anyone could possibly be in any state with impossible content.\textsuperscript{153}

Scepticism about the possibility of states with contradictory contents is closely associated with the Stalnaker-Lewis possible worlds analysis of content,\textsuperscript{154} and it may be thought that one can avoid the conclusion by opting for a more fine-grained analysis of mental content that takes the objects of thought to be sentence-like entities. This is a misunderstanding. The above argument does not depend upon some relatively technical account of propositions and possible worlds. Rather, it depends upon the intuitive point that such states could not have any content. This also applies if we take the objects of thoughts to be sentence-like entities. To see this, consider Field's analysis of believing: "X believes that $p$ if and only if there is a sentence or sentence-analogue $S$ such that X believes* $S$ and $S$ means that $p$" (Field [1994]: 42). This account of having a belief has two parts. First, to have a belief is to stand in some relation to a sentence and, second, for that sentence to have a certain meaning. Suppose the sentence has no meaning. Then there is also nothing that is being believed (ditto for other attitudes). So, since there can be no believing without believing something, there can be no

\textsuperscript{153} Many think that perceptual states can have contradictory content (e.g. Crane [1988]). I reject this for the very same reason as I reject the possibility of belief having contradictory content. That is not to say that I deny the phenomena that tempt people to say, for example, that perceptual states can have contradictory content. I merely insist that these phenomena can, and indeed must, be interpreted in a different way.

\textsuperscript{154} See Lewis [1986] and Stalnaker [1987].
belief unless there is an appropriately related meaningful sentence. But contradictory sentences are meaningless. Emphasising sentence-like entities rather than propositions, therefore, is not going to bring one any closer to making sense of states with contradictory content.  

Another common objection is equally misguided: Some contradictions are unobvious, and people may come to believe them by overlooking the fact that they believe a contradiction. Surely this must be possible. Consider this neat argument by Mellor: “Suppose I believe that all even numbers are the sum of two primes, and you don’t. One of us has a contradictory belief” (Mellor [1988]: 149). However, I maintain that, even if a contradiction is unobvious, it can’t be believed, simply because it can’t be grasped. Making the contradiction more complex and harder to parse does not miraculously make it possible for the contradiction to be grasped. Indeed, in the case of a complicated contradiction, it would at best have to be even harder to grasp the contradiction. What we can do is to believe something which we think is expressed by a certain sentence – a sentence which is in fact contradictory and meaningless. But believing something which we think is expressed by a certain sentence is not the same as believing a contradiction.  

Some people accept that one can’t believe what is a priori impossible. However, they think it is possible to believe something that is a posteriori impossible, e.g. that water is HCL. It is hard to see how they could be right. To believe is to believe that something is the case. Belief without content can’t exist. But the content of a belief is what is believed to be the case, and in the case of the impossible – whether a priori or a posteriori – there is nothing that could be believed to be the case. If we seriously consider the matter, I think we must concede that there is an ineliminable tension between the thought that a belief must have content – that it must be about something – and that the impossible could be the object of belief. In the end, the suggestion that someone could believe something impossible simply makes no sense.

---

155 The failure to see this is, perhaps, based on the mistaken assumption that these accounts say that what is believed is not that things are thus-and-so but literally a sentence. This is absurd. We do not believe or desire sentences, although it may well be that we can analyse believing or desiring something as standing in a certain relation to a sentence and for that sentence to have a certain meaning.
Hume was the first philosopher to explicitly recognise that there is “a great difference betwixt the simple conception of the existence of an object, and the belief of it” (Hume [2003]: 1.3.7.2) and also that “this difference lies not in the parts or composition of the idea” (Ibid.). If the difference does not consist in what is conceived, Hume reasoned it “must lie in the manner, in which we conceive it” (Ibid.). Specifically, Hume argued that it must consist in belief being a state with greater ‘force and vivacity’ than mere thought.\footnote{See e.g. Hume [2003]: 1.3.7; 1.3.8; Appendix.}

There are at least three interpretations of what Hume had in mind.\footnote{See, for example, Bennett [2001]; Everson [1988]; Govier [1972]; Pears [1990]; Stroud [1977].} (i) Just as two images can depict the same scene, but one have more vivacity because it is richer in contrast, saturation and detail, so Hume might have thought that believing that $p$ somehow involves a representation of that $p$ that is in some sense richer than the representation involved in merely entertaining that $p$ in thought. (ii) Just as one punch can have more force than another in virtue of its greater causal efficacy, so Hume might have thought that believing that $p$ is set apart from entertaining that $p$ in thought in virtue of its superior causal efficacy. (iii) Perhaps we should eschew an interpretation guided by the notions of ‘force and vivacity’, given that Hume himself didn’t think them particularly adequate. (Hume said that he really was “at a loss for terms to express my meaning” (Hume [2003]: 1.3.7.7)). We might then stress his insistence that the difference must lie in ‘the manner in which a content is conceived’ and that this is, ultimately, a matter of feeling. As Hume says when I wou’d explain this manner, I scarce fine any word that fully answers the case, but am oblig’d to have recourse to every one’s feeling, in order to give him a perfect notion of this operation of the mind. An idea assented to feels different from a fictitious idea, that the fancy alone presents to us (Ibid.).

Each interpretation has its strengths, but on balance I think the final one captures Hume’s thoughts best. The first interpretation is problematic because it ascribes a thesis to Hume which is obviously not credible but also because it is inconsistent with Hume’s avowed claim that one can’t distinguish between belief and mere thought on the basis of content, that “when after the simple conception of any thing we wou’d conceive it as existence, we in reality make no addition to or alteration on our first idea” (Hume [2003]: 1.3.7.2). The second interpretation would align Hume with
causal functionalism. But there is reason to believe he would not have welcomed such company unreservedly. True, Hume does emphasise the different causal profiles of judgment and mere thought, but he does not say this is what constitutes their difference. Per contra, he quite explicitly says their causal difference is itself a consequence of the different manners in which a content is conceived by these states. The third interpretation has largely been overlooked in the literature, but it strikes me as the best interpretation overall, and it has the added bonus of showing that the seeds of the main ideas of chapter three were already present in Hume’s thoughts.

158 Thus he speaks of “that act of the mind, which renders realities more present to us than fictions, causes them to weight more in the thought, and gives them a superior influences on the passions and imagination” (Hume [2003]: 1.3.7.7; my emphasis). He also says

in philosophy we can go no farther, than assert, that it is something felt by the mind, which distinguishes the ideas of the judgment from the fictions of the imagination. It gives them more force and influence; makes them appear of greater importance; infixes them in the mind; and renders them the governing principles of all our actions (Ibid.; underlining is my emphasis).


