

# **The Rational Significance of Desire**

Adrian Avery Archer

Submitted in partial fulfilment of the  
requirements for the degree of  
Doctor of Philosophy  
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

# CONTENTS

---

<b>1. Introduction</b>	<b>1</b>
1.1. Providing Reasons	1
1.2. Cognitivism about Desire	4
1.3. The Guise of the Good Theory	8
1.4. Overview	13
<b>2. Direction of Fit and the Aim of Desire</b>	<b>17</b>
2.1. Introduction	17
2.2. The Anscombe Interpretation	19
2.3. The Independence Thesis	24
2.4. Motivating Illocutionary Content	27
2.5. Desires, Wishes, and Wants	32
2.6. Presuppositions and Illocutionary Implications	35
2.7. Desires, the Doable, and the Good	41
2.8. Conclusion	46
<b>3. Desire as Belief or Perception</b>	<b>48</b>
3.1. Introduction	48
3.2. The (Broadly) Anscombean View	49
3.3. The Desire-as-Belief Thesis	55
3.4. The Desire-as-Perception Thesis	59
3.5. The Anscombe Interpretation	61
3.6. The Smith Interpretation	65
3.7. Diagnosing the Direction of Fit Analysis	67
3.8. The Two-Content Interpretation	73
3.9. Conclusion	76
<b>4. Two Kinds of Rational Significance</b>	<b>77</b>
4.1. Introduction	77
4.2. Rational Significance	78
4.3. The Dual-Significance Thesis	83
4.4. Objections to the Dual-Significance Thesis	89
4.5. Belief as a Justification-Receiving Attitude	95
4.6. Implications of the Dual-Significance Thesis	101
4.7. Conclusion	106

<b>5. Cognitivism and the Aim of Desire</b>	<b>107</b>
5.1. Introduction	107
5.2. The Objection to Simple Cognitivism	108
5.3. The Objection to Sophisticated Cognitivism	113
5.4. Reply to the Constitutive Aim Argument	119
5.5. The Inconsistency Test	124
5.6. Reply to the Perverse Desire Argument	128
5.7. Conclusion	132
<b>6. Are Desires Rationally Significant?</b>	<b>133</b>
6.1. Introduction	133
6.2. Desire, Perception, and Coin-flips	134
6.3. The Juxtaposition of Desire and Coin-flip	141
6.4. The Juxtaposition of Coin-flip and Coin-flip	143
6.5. The Juxtaposition of Coin-flip and Desire	148
6.6. Passive versus Active Selection Procedures	152
6.7. The Zero-Sum Structure of Action	155
6.8. Conclusion	163
<i>Bibliography</i>	165

## ACKNOWLEDGEMENTS

---

Many thanks to my dissertation supervisors, Christopher Peacocke, Susanna Siegel, and Katja Vogt, for their invaluable help and advice throughout the writing of this thesis, and to my examiners, Achille Varzi and Elliot Paul, for their generous comments and suggestions. Thanks also to Nomy Arpaly, Kyle Barrett, Mark Berger, Michael Brent, John Brunero, Melissa Bobe, Nicholas Engel, Alison Fernandes, Patricia Kitcher, Stephanie Leary, Shen-yi Liao, Karen Lewis, Antonella Mallozzi, Wolfgang Mann, Katharine McIntyre, Kian Mintz-Woo, Usha M. Nathan, Christiana Olfert, Ariadna Pop, Jesse Prinz, and Christopher Whalin for their helpful suggestions, discussions and support along the way. (Any errors or inaccuracies remain my own.)

I dedicate this thesis to my mother, Averlyn Archer, who always encouraged me to pursue my passions, and my grandfather, Peter Archer, to whom I owe my uncompromising curiosity.

## 1

## INTRODUCTION

---

### 1.1. Providing Reasons

What (if any) is the rational significance of desire? The present investigation will go some distance towards answering this question. The answer I ultimately arrive at is that a desire constitutes a type of selection procedure for deciding between competing courses of action or between action and inaction. As such, desires may be said to guide our actions. However, I claim that it remains unestablished that the type of guidance that a desire provides is appropriately described as “rational”. On the contrary, I argue that we have good reason to think that, in terms of its rational significance, a desire is no better than a random selection procedure, such as a coin-flip or guess. The significance of this conclusion, and how I arrived at it, will become clearer in the chapters that follow. In this introduction, I will attempt to provide some background and context for what is to come by defining some key terms, circumscribing my primary areas of concern, and offering a brief overview of each chapter.

Let us begin by getting clear on what it means for an attitude to provide an agent with a reason. We may distinguish between two different senses in which an attitude, *X*, may provide us with a reason to adopt a belief or intention, *Y*:

(A1) *X* provides us with a reason to adopt *Y* by putting us in touch with a consideration that speaks in favour of adopting *Y*.

(A2) *X* provides us with a reason to adopt *Y* by constituting a consideration that speaks in favour of adopting *Y*.

The difference between (A1) and (A2) is the difference between the perceptual appearance that it is raining outside giving one a reason to believe that it is raining outside, and the fact that one has the perceptual appearance that there is a pink elephant in the corner of the room giving one a reason to check oneself into a hospital. In the first case, the perceptual appearance provides one with a reason by putting one in touch with a consideration that is independent of the perceptual appearance itself and which speaks in favour of adopting the belief that it is raining outside—namely, the fact that it is raining outside. The fact that it is raining outside is independent of the perceptual appearance of it raining outside in the sense that it would obtain (and constitute a reason to believe that it is raining outside) even if one did not have the perceptual appearance in question. By contrast, the fact that one has the perceptual appearance of a pink elephant in the corner of the room does not put one in touch with some independent

consideration in favour of adopting the intention to check oneself into a hospital. Rather, the perceptual appearance itself constitutes the relevant consideration.

Significantly, when an attitude provides one with a reason in the sense described in (A1), it does so in virtue of the fact that it is an intentional state with objective purport. The perceptual appearance that it is raining provides one with a reason to believe it is raining because it purports to put one in touch with a mind-independent fact—namely, the fact that it is raining. If, following Franz Brentano, we assume that psychological states are distinctive because they have the property of intentionality or aboutness,<sup>1</sup> then we can say that when a psychological state provides one with reasons in the sense described in (A1), it is doing so *qua* psychological state (i.e., in virtue of being a state with the property of aboutness). By contrast, when an attitude provides one with a reason in the sense described in (A2), it does not do so *qua* psychological state. Rather, it provides one with a reason in the very same way that any old fact might. For example, the fact that I'm bleeding profusely gives me a reason to check myself into the hospital in the very same sense that having the perceptual appearance that there is a pink elephant gives me a reason to check myself into the hospital. If this is right, and given that the former lacks the property of aboutness, it follows that the latter does not provide reasons in the relevant sense in virtue of the fact that it has the property of aboutness. Hence, only (A1) describes a type of rational significance that is unique to psychological states.<sup>2</sup>

I take it to be relatively uncontroversial that a desire may provide reasons in the sense described by (A2). For example, the fact that I have a compulsive desire to wash my hands for a sixth time may give me a reason to wash my hands for a sixth time if I believed that doing so would free me of the unpleasant burden my compulsive desire embodies, even if I believe that there is nothing to be gained from washing my hands for a sixth time (i.e., apart from the fact that it would satisfy my compulsive desire). As with the perceptual appearance that there is a pink elephant in the corner of the room, my compulsive desire does not provide me with a reason to wash my hand by putting me in touch with a consideration that speaks in favour of my washing my hands that is independent of the desire itself. What remains controversial is the thesis that a desire provides reasons in the sense described by (A1). This is the sense of providing reasons with which this dissertation will be concerned. In short, I wish to investigate whether a desire provides reasons *qua* psychological state.

---

<sup>1</sup> Brentano [1874: 88-89].

<sup>2</sup> The distinction is not meant to be exclusive, at least not in the sense that an attitude can provide reasons in one sense or the other, but not both. For example, the perceptual appearance of a pink elephant provides one with both a pro tanto reason to believe that there is a pink elephant and a pro tanto reason to check oneself into a hospital. However, when the perceptual appearance of a pink elephant provides one with a pro tanto reason to believe there is a pink elephant, it is doing so *qua* psychological state; but when it provides a pro tanto reason to check oneself into a hospital, it is doing so *qua* fact. *Qua* psychological state, a perceptual appearance is a truth-value bearer (or, if you prefer, a bearer of veridicality conditions); *qua* fact, it is a truth-maker (or a determiner of veridicality). Since a perceptual appearance may play the role of both a veridicality bearer and a veridicality determiner, it may provide reasons in both senses. But each sense will be relative to its respective role. Thanks to an anonymous commentator for alerting me to the need for this clarification.

## 1.2. Cognitivism about Desire

According to one widely discussed contemporary theory of desire, a desire provides reasons because a desire represents the desired outcome as good or as something that ought to be brought about. Following David Velleman, I will refer to this theory of desire as *cognitivism*.<sup>3</sup> One of the central motivations for accepting cognitivism comes from an attempt to make sense of a certain aspect of our ordinary linguistic practice; namely, the fact that the specification of a desire is often taken to be a satisfactory answer to the questions: “Why did you  $\varphi$ ?”, “Why are you  $\varphi$ -ing?”, and “Why are you going to  $\varphi$ ?”<sup>4</sup> For example, saying that I want a glass of water is typically taken to be a satisfactory answer to the question “why are you going to the kitchen?” If we conceive of being asked “why?” as a request for reasons<sup>5</sup>, then the fact that citing a desire is typically taken to be a satisfactory answer to the question “why?” suggests that desires provide reasons, or so the argument goes. Hence, if cognitivism were true, it would provide us with a straightforward way of making sense of a significant part of our ordinary reason-giving practice.

A second reason cognitivism may be deemed attractive is because it is consistent with a longstanding philosophical conception of human beings as fundamentally rational agents, understood as agents that act for (or for what are believed to be) reasons. Putatively, there are cases in which an agent does act for reasons, such as when an agent is motivated to perform an action by the belief that she is morally or rationally obligated to perform it. For example, if I adopt the intention not to cheat on my taxes because I believe I have a moral obligation not to cheat on my taxes, my reasons for adopting the intention not to cheat on my taxes seems to correspond with the relevant belief about what constitutes my moral obligations. However, there also seem to be cases in which I am motivated not by a belief that I am morally obligated to perform a certain course of action, but by a desire. For example, I may be motivated to buy vanilla ice cream by a desire to buy vanilla ice cream and yet fail to believe that I am morally (or otherwise) obligated to buy vanilla ice cream. This means that if we wish to preserve the idea that human beings always act for or in the light of reasons, then we cannot only consider those cases in which an agent is motivated by the belief that she is morally or rationally obligated to bring about some outcome. We must also consider those cases in which she is motivated by a

---

<sup>3</sup> Velleman [1992]. See and Cf.: Brunero [2009]; Tenenbaum [2006]. Velleman conceives of cognitivism as a version of the guise of the good theory of desires, the claim that intentional action is “directed at outcomes regarded *sub specie boni*: under the guise of the good”([1992: 3]). Cognitivism is identified with those versions of the guise of the good theory that claim or entail that desires provide reasons in favour of bringing about the desired outcome because desires represent the desired outcome as good. The present notion of cognitivism should not be confused with moral cognitivism, understood as the claim that moral statements are apt for truth or falsity (e.g., Boyd [1988] and Railton [1986]).

<sup>4</sup> For discussion, see: Anscombe [2000: 70ff.].

<sup>5</sup> Here and henceforth, I am using the word “reason” to mean *justificatory reason*. There has been a great deal of discussion in the action theory literature of the difference between justificatory reasons and explanatory reasons. This distinction is typically taken to be similar, if not identical, to the distinction between normative and motivating reasons. I will not attempt to add my own analysis of justificatory reasons, or of the difference between justificatory and explanatory reasons, here.

desire. Thus, if we held that desiring to bring about some outcome provided one with a reason to bring about that outcome, we would be allowed to preserve the idea that human beings always act for reasons, even in those cases in which their actions are based on a desire rather than on a belief. However, if we deny that desires provide reasons, then there seems to be a large subset of our everyday actions that are performed for no reason. This, in turn, seems to cast doubt on the conception of human beings as fundamentally rational agents.

The preceding pair of motivations for adopting cognitivism is meant to be suggestive rather than exhaustive. I do not mean to imply that cognitivism is the only or even the most effective way of achieving the desiderata alluded to above. Rather, I wish to illustrate that the sort of reasons for which one may find cognitivism attractive range from the quotidian to the highly theoretical. Given the range of possible motivations for thinking that desires provide reasons, the question of whether or not they do warrants close, patient investigation. This is what I attempt to do in this dissertation. The present investigation will be for the most part restricted to those desires that entail a disposition to act. This is because I am primarily interested in the kind of desire that may potentially feature in the aetiology of an intentional action.<sup>6</sup> One implication of this restriction is that I will not be concerned with what Tony Milligan refers to as a “helpless desire”, understood as a desire that is directed at an outcome “we cannot do anything to bring about.”<sup>7</sup> Examples of helpless desires include the desire to fly like a bird, to alter the past, or that the sun shines at some remote date in the future.<sup>8</sup>

I will refer to anyone who denies that a desire represents its object as good as a *non-cognitivist*. While it is in principle possible for one to be a non-cognitivist in the present sense and yet affirm that desires provide reasons (i.e., if one thought that desires provide reasons by some means besides representing their object as good), I am not aware of any theorists who both embrace non-cognitivism and also affirm that desires provide reasons. Moreover, it is difficult to imagine what a plausible version of such a position would look like. Nevertheless, I will not be making the assumption that the non-cognitivist is committed to denying that desires provide reasons. The point of contention between the cognitivist and the non-cognitivist is whether or not desires represent their object as good. However, if the non-cognitivist is able to successfully show that a desire does not represent its object as good, this would block the most widely discussed contemporary strategy for arguing that desires provide reasons. Hence, it is only as a potential objection to the claim that desires provide reasons that non-cognitivism is presently of interest to us.

---

<sup>6</sup> Strawson [1994] suggests that it is best to keep the notion of a desire distinct from the notion of a disposition to act. However, the considerations he cites in support of this recommendation only have force if we are interested in providing a much more general and promiscuous account of desire than that which interests us at present.

<sup>7</sup> Milligan [2007: 310].

<sup>8</sup> I borrow these examples from Milligan [2007: 309-310].



### 1.3. The Guise of the Good Theory

The cognitivist conception of desire is typically regarded as a version of the guise of the good theory (henceforth, GG theory). Sergio Tenenbaum defines GG theory as “the view that desire (or perhaps intention, or intentional action) always aims at the good.”<sup>9</sup> According to the cognitivist, a desire aims at the good in virtue of the fact that a desire represents its object as good. However, it may be argued that representing something as good is not the only way in which a desire may be said to aim at the good. If this is right, then it should be possible to buy into GG theory without buying into cognitivism. Nevertheless, most contemporary GG theorists are also cognitivists. Consequently, I will be treating cognitivism and GG theory as more or less interchangeable.

In its Aristotelian formulation, GG theory is widely taken to include a claim about desires. In this regard, the following passage is typically taken by action theorists to be representative of Aristotle’s view:

The object of *desire* and the object of *wish* is either good or the apparent good. (Eudemian Ethics [1235b26-27], Italics mine)<sup>10</sup>

The claim that the good is the object of desire was later codified by the Scholastics in the slogan: *quid appetitur, appetitur sub specie boni* (whatever is desired is desired under the guise of the good). I will refer to the version of GG theory that includes the claim that desire aims at the good as the *Desire Thesis*.

Some contemporary GG theorists have given up on defending the Desire Thesis in favour of simply saying that intention (or intentional action) aims at the good. For example, Joseph Raz [2010] identifies GG theory with the following three propositions:

- (1) Intentional actions are actions performed for reasons, as those are seen by the agents.
- (2) Specifying the intention that makes an action intentional identifies central features of the reason(s) for which the action is performed.

---

<sup>9</sup> Tenenbaum [2010: 3].

<sup>10</sup> It has become all but standard for contemporary GG theorists to cite the following passage from Aristotle’s *De Anima* as the *locus classicus* of GG theory: “It is the object of desire which produces movement, but this is either the good or the apparent good” (*De Anima*, III.10 [433a28-29]). For example, Matthew Boyle and Douglas Lavin cite this passage at the very beginning of their recent paper, “Goodness and Desire”. However, I have opted to set this passage aside since it identifies the “object of desire”, as opposed to desires themselves, as that which moves an agent, and this subtlety of phrasing introduces a number of vexed issues which continue to be hotly debated among Aristotle scholars. For readers with interests in this subject matter, the Greek word translated as “desire” is *orekton*, and the Greek word translated as “wish” is *boulêton*. According to John Cooper [1999: 241-242] and Jessica Moss [2010: 65], Aristotle uses *orekton* (*orexis*) as a generic designation for any desire, while *boulêton* (*boulêsis*) is used to refer to distinctively rational motivations. Since I lack the necessary expertise to evaluate the exegetical claims of Cooper and Moss, I do not wish to either endorse or impugn their interpretation. I present their reading, not to settle the question of what Aristotle is actually committed to saying, but rather as an example of what Aristotle is taken to be committed to saying; in this case, by a pair of widely respected ancient scholars. See and cf. *Nicomachean Ethics* [1094a1-2].

- (3) Reasons for action are such reasons by being facts that establish that the action has some value.<sup>11</sup>

Raz then goes on to emphasise that he “remains noncommittal about the relations of intentions and desires, and will therefore not discuss desires”.<sup>12</sup> This shift away from discussing desires—traditionally taken to be one of the central concerns of GG theory—has come about largely in response to criticisms of the Desire Thesis, such as those advanced by David Velleman.<sup>13</sup> Since the present investigation is concerned with the rational significance of desire, I will not consider those versions of GG theory that do not include the Desire Thesis.

Since beliefs and judgements are typically taken to be the paradigm examples of psychological states that provide reasons, some theorists have sought to establish that desires provide reasons by arguing that desires are a type of belief or judgement.<sup>14</sup> Specifically, the desire to bring about **P** is equated with the belief or judgement that it is good to bringing about **P**. Since the belief or judgement that it is good to bringing about **P** is just the kind of thing that could provide one with a reason to bring about **P**, this view allows us to make sense of the claim that desires provide reasons.

More recently, a few GG theorists have equated a desire with a special kind of perceptual appearance of the good.<sup>15</sup> Since a perceptual appearance is also widely regarded as an example of an attitude that provides reasons, the equation of a desire to a type of perceptual appearance also potentially allows us to make sense of the claim that desires provide reasons. However, the identification of a desire with a perceptual appearance rather than a belief is significant because it imposes an important restriction on what we can plausibly take to be entailed by the claim that desires provide reasons. Some theorists—most notably Peter Railton and Kieran Setiya—have taken the claim that desires provide reasons to entail that desires are reason-responsive.<sup>16</sup> However, if we were to adopt the view that a desire is a type of perceptual appearance, then there is no reason to assume that the entailment holds. There are two senses in which an attitude may be reason-responsive—a descriptive sense and a normative sense. From a purely descriptive perspective, there are indeed circumstances in which a desire may be plausibly described as reason-responsive. For example, suppose that I desire to take a swig from what appears to be a refreshing bottle of apple juice, but (upon closer inspection) is revealed to be a bottle of vinegar. Upon discovering that it is vinegar rather than apple juice in the bottle, I

---

<sup>11</sup> Raz [2010: 111].

<sup>12</sup> Raz [2010: 112].

<sup>13</sup> Velleman’s criticisms of GG theory are directed at the Desire Thesis. For example, in the introduction of his widely discussed paper, “The Guise of the Good”, he describes the rationale behind GG theory as follows: “The reason is that he acts intentionally only when he acts out of a *desire* for some anticipated outcome, and in *desiring* that outcome he must regard it as having some value” (Velleman [1992: 3], *Italics mine*).

<sup>14</sup> For a defence of this view, see Price [1998].

<sup>15</sup> For example: Stampe [1987], Oddie [2005], and Tenenbaum [2007].

<sup>16</sup> Railton [2006].

may lose my desire to take a swig. Moreover, perceptual appearances may be said to be reason-responsive in a similarly limited sense. For example, a shadowy object in the corner of my bedroom may appear to be a person, until I remember that it is actually a lamp, at which point it quickly ceases to “look” like a person.

However, pace the preceding observations, it is also true that desires (like perceptual appearances) are notoriously recalcitrant in the face of countervailing reasons. I may continue to desire to have a cigarette even after I learn that it is bad for my health, and no amount of reminders from my doctor about the dangers of large amounts of deep fried or fatty foods may suffice to silence my desire to partake in them.<sup>17</sup> Hence, from a purely descriptive point of view, desires often fail to be reason-responsive. Thus, the claim that desires are reason-responsive is, from a purely descriptive perspective, not always true. Moreover, there is no categorical difference in this respect between a desire or perceptual appearance, on the one hand, and a belief, on the other, since there may be cases in which a belief is not reason-responsive (such as when a belief is held in an irrationally dogmatic way).

I believe that the more philosophically interesting question is whether desires are reason-responsive in a normative sense. It is with regards to this question that the claim that a desire is a type of perceptual appearance and the claim that a desire is a type of belief pull in opposing directions. Normatively speaking, beliefs are reason-responsive in the sense that we would hold an agent liable to rational criticism if she failed to revise her beliefs in the face of what she acknowledges to be conclusive contravening evidence. The same is not true of perceptual appearances. For example, suppose a partially submerged stick perceptually appears bent to me. Even after I have been presented with what I acknowledge to be conclusive contravening evidence, I am not liable to rational criticism if I continue to have the perceptual appearance of it being bent. Hence, there is a normative sense in which perceptual appearances are not reason-responsive. Consequently, if we take a desire to be a type of perceptual appearance, we are left without any motivation for thinking that desires are, from a normative point of view, reason-responsive. Moreover, this appears consistent with how we ordinarily conceive of and talk about desires. For example, I may decide to give up smoking because I take myself to have conclusive reasons for thinking it is bad for my health, and yet continue to have the desire for a cigarette all the same. The mere fact that my desire to have a cigarette persists does not make me liable to rational criticism. This suggests that desires are not reason-responsive in the normatively significant sense that beliefs or judgements are; an agent is not ipso-facto deemed irrational because his desires fail to line up with what he believes (or takes there to be conclusive evidence for holding) to be good.<sup>18</sup>

The preceding observations suggest that Setiya mischaracterises the GG theorists when he describes her as being committed to the following principle:

---

<sup>17</sup> For a psychologist’s perspective on the claim that desires need not be reason-responsive, see: Zajonc [1980].

<sup>18</sup> Another way that this point may be put would be to say that a desire is *not* an example of what Scanlon [1998] calls a “judgement-sensitive attitude”, understood as the “class of attitudes for which reasons...can sensibly be asked for or offered”[1998: 4-5]. Cf. Brandt [1969].

When someone wants to *X*, or wants it to be the case that **P**, they want it for a reason, and reasons for desire must be respects in which the object of desire is seen as good.<sup>19</sup>

Here, Setiya portrays the GG theorist as being committed to the claim that an agent always has a reason for her desires. However, according to those who equate desires to perceptual appearances, it is not only false, as a psychological claim, that desires are always had for and responsive to reasons, but even more significantly, that we do not require that desires be had for or be responsive to reasons. On this view, it would be no more appropriate to require a reason for a desire than it would be to require a reason for a perceptual appearance. In light of these considerations, I will not be assuming that those who are committed to saying that desires provide reasons are also committed to saying that there may be reasons for a desire. While the claims are consistent with each other, the former does not entail the latter.

#### 1.4. Overview

The remaining chapters of this dissertation form a set of inter-supporting essays in which I criticise and offer an alternative to the cognitivist conception of desire. In chapter 2, I introduce an alternative interpretation of the direction of fit metaphor to that offered by Elizabeth Anscombe. I begin by arguing that the Anscombe interpretation is ill-suited for the purpose for which it is typically employed by GG theorists—namely, to make sense of the idea that desire aims at the good in a sense that is roughly analogous to how belief aims at the truth. I then advance my own interpretation of the direction of fit metaphor, which I refer to as the *two-content interpretation*. Specifically, I argue that those attitudes that display a belief-like direction of fit do so in virtue of having what I call *indicative content*, and that those attitudes that display a desire-like direction of fit do so in virtue of having what I call *imperative content*. I conclude by offering a diagnostic test for determining the correctness conditions of an attitude with imperative content. I demonstrate that when this test is applied to desire, we arrive at the conclusion that desires do not aim at the good. The arguments of chapter 2 are meant to go some distance towards answering the question: *How are we to conceive of the direction of fit of desire?*

In chapter 3, I attempt to further situate the two-content interpretation in the context of the contemporary discussion of Anscombe-inspired forms of GG theory. I begin with an analysis of the two most widely-discussed strategies for unpacking Anscombean GG theory—namely, the thesis that a desire is a type of belief and the thesis that a desire is a type of perception. I argue that, contrary to expectation, the two dominant interpretations of direction of fit—those associated with Anscombe and Michael Smith, respectively—are actually consistent with the equation of a desire with a belief or perception. I take this to be symptomatic of the fact that both Anscombe’s and Smith’s interpretations fail to get at the heart of the matter, as far as the direction of fit analysis is concerned. In this regard, both of the dominant interpretations of direction of fit stand in contrast to the two-content interpretation. The

---

<sup>19</sup> Setiya [2010: 86].

arguments of chapter 3 attempt to answer the question: *How is the analysis of direction of fit on offer in this dissertation importantly different from and preferable to the dominant interpretations?*

In chapter 4, I argue that there are two distinct kinds of rational significance that an attitude may display. The first, which I refer to as *commitment-incursion*, is displayed by belief, and the second, which I refer to as *rational support*, is displayed by perception. I consider and respond to an objection to my claim that there are two kinds of rational significance. In response to these objections, I argue that the distinctive rational significance displayed by belief is tied to the fact that a belief is an attitude that may be had for reasons while a perceptual appearance is not. Getting clear on the rational significance of belief is important for assessing the claim that a desire has the same rational significance as the belief that the desired outcome is good. In this respect, chapter 4 provides a framework for my assessment of David Velleman's criticisms of cognitivism in chapter 5. The arguments in chapter 4 attempt to go some distance towards answering the question: *How are we to understand the difference between the rational significance of belief and rational significance of perceptual appearances?*

In chapter 5, I argue that David Velleman's widely-discussed objections to cognitivism fail to establish their conclusion. I begin by summarising two sets of objections found in Velleman's paper, "The Guise of the Good": the first directed at the version of cognitivism inspired by Donald Davidson, and the second directed at the version of cognitivism inspired by Elizabeth Anscombe. I argue that both sets of objections fail because they saddle the cognitivist with commitments she need not have. In the process of so doing, I argue that the cognitivist is most plausibly and charitably seen as committed to the claim that a desire has the same kind of rational significance displayed by a perceptual appearance. The arguments in chapter 5 attempt to answer the question: *What is the most plausible way of understanding the cognitivist version of the claim that desires provide reasons?*

In chapter 6, I argue that a desire has no more rational significance than a random selection procedure. I begin by presenting a set of considerations that suggest that, in terms of its rational import, a desire has more in common with a coin-flip or guess than with a perceptual appearance. At first pass, this may seem like an unattractive conclusion. However, I attempt to show that this conclusion is made palatable by the fact that the kind of rational significance displayed by perceptual appearances is, strictly speaking, unnecessary in the practical sphere. As such, the fact that a desire is of little or no rational significance should be no cause for alarm. The arguments in chapter 6 attempt to go some distance towards answering the question: *Are desires rationally significant?*

## 2

## DIRECTION OF FIT AND THE AIM OF DESIRE

---

### 2.1. Introduction

The Guise of the Good Theory of Desires (GG theory) is the thesis that desires aim at the good.<sup>20</sup> According to the version of GG theory that derives its primary inspiration from the work of Elizabeth Anscombe, the claim that a desire aims at the good is taken to entail the following set of claims:

- I. A desire aims at the good because a desire represents the desired outcome as good.
- II. A desire is correct only if the desired outcome is good because a desire aims at the good.
- III. A desire has world-to-mind fit because the world is incorrect or subject to revision when there is a mismatch between the world and what is desired, and a belief has mind-to-world fit because the belief is incorrect or subject to revision when there is a mismatch between what is believed and the world.
- IV. The aim of a desire explains, is explained by, or is otherwise linked to the fact that it has world-to-mind fit, and the aim of a belief explains, is explained by, or is otherwise linked to the fact that it has mind-to-world fit.

I will refer to the conjunction of the four preceding claims as Anscombean GG theory. Much of the contemporary interest in Anscombean GG theory stems from the fact that it provides us with a framework for making sense of how a desire may provide a reason for bringing about the desired outcome.<sup>21</sup> According to (I), a desire aims at the good in virtue of the fact that it

---

<sup>20</sup> Philosophers who have held a view recognisably similar to GG theory include Plato, Aristotle, Aquinas, Kant, Anscombe and Davidson, along with numerous contemporary thinkers.

<sup>21</sup> See section 1.1 of the introduction of this dissertation for a clarification of what I mean by “providing reasons” in the present context.

represents the desired action as good.<sup>22</sup> In cases in which a desire represents matters as they truly are, the desire puts the desiring agent in touch with a consideration that speaks in favour of bringing about the desired outcome—namely, the fact that the outcome in question is good. Hence, a desire may provide the desiring agent with a reason to bring about the desired outcome.

In this chapter, I will be offering a critique of Anscombean GG theory. My aim will be both critical and constructive. On the critical side of things, I argue that (III) is an unsatisfactory way of understanding the direction of fit metaphor. Moreover, I argue that there is a tension between (I), (III) and (IV). On the constructive side of things, I advance an alternative interpretation of the direction of fit metaphor to that offered by Anscombe, which I refer to as the *two-content interpretation*. One advantage that the two-content interpretation has over the Anscombe interpretation is that the former provides us with the resources necessary to hold that two attitudes with the same direction of fit may have different correctness conditions.

## 2.2. The Anscombe Interpretation

In §32 of *Intention*, Anscombe employs an example comparing a shopping list and a detective's record to illustrate two different ways in which our words (written or spoken) may relate to the world. She writes:

Let us consider a man going round a town with a shopping list in hand. Now it is clear that the relation of this list to things he actually buys is one and the same whether his wife gave him the list or it is his own list; and that there is a different relation when a list is made by a detective following him about. If he made the list itself, it was an expression of an intention; if his wife gave it to him, it has the role of an order. What then is the identical relation to what happens, in the order and the intention, which is not shared by the record? It is precisely this: if the list and the things that the man actually buys do not agree, and if this and this alone constitutes a mistake, then the mistake is not in the list but in the man's performance...; whereas if the detective's record and what the man actually buys do not agree, then the mistake is in the record.<sup>23</sup>

Let us refer to the example discussed here as the *shopping list example*. Anscombe notes that, in the case of the shopping list, the task is to get the world (the items purchased) to match our words (the shopping list), while in the case of the detective's record, the task is to get the word (the detective's record) to match the world (the items purchased). In the terminology of speech act theorists, the former is an example of *world-to-word* fit, and the latter of *word-to-world* fit.<sup>24</sup>

---

<sup>22</sup> There are two different accounts of why desires entail representations of the good currently found in the literature: (i) because desires involve beliefs or judgements of the good (e.g., Price [1989]; Byrne and Hájek [1997]), or (ii) because desires involve perceptions of the good (e.g., Stampe [1987]; Tenenbaum [2007]).

<sup>23</sup> Anscombe [2000: 56].

<sup>24</sup> The expression “direction of fit” seems to have been first used John Searle, who employed it to describe the two different ways in which words may relate to the world, as highlighted by Anscombe (See: Searle [1985]).

Significantly, Anscombe sees an intention (i.e., a certain type of attitude) and an order (i.e., a certain type of speech act) as standing in the same logical relation to the man's actions. By her lights, it does not matter if we see the shopping list as an expression of an intention (on the part of the man) or as an order (on the part of his wife); in either case, the shopping list displays the same direction of fit. Thus, Anscombe takes the observation that our words may relate to the world in two very different ways as also applicable to attitudes. This idea has been taken up by several theorists in moral psychology and the philosophy of mind, where a distinction is often drawn between attitudes that display *mind-to-world* fit and attitudes that display *world-to-mind* fit. Huw Price summarises the central intuition behind the direction of fit analysis of attitudes as follows:

Beliefs have a 'mind-to-world' direction of fit; they aim at fitting the world, at being true. Desires, on the other hand have a 'world to mind' direction of fit; we aim to change the world to fit desires, and not vice versa.<sup>25</sup>

The standard way of unpacking the direction of fit metaphor, which builds on Anscombe's shopping list example, is in terms of whether an attitude or the world is deemed to be incorrect or subject to revision when there is a lack of correspondence between the two. According to this approach, what it means for a belief to have mind-to-world fit is that the belief is deemed to be incorrect or subject to revision when there is a mismatch between the world and what is believed. By contrast, what it means for a desire to have world-to-mind fit is that the world is deemed to be incorrect or subject to revision when there is a mismatch between that which is desired and the world. Call this the *Anscombe interpretation*.

Unfortunately, the Anscombe interpretation seems to conflate the claim that an attitude has a certain direction of fit with the claim that an attitude has certain correctness conditions. There is widespread agreement that cognitive attitudes like hypothesising that **P**, supposing that **P**, and pretending that **P** have the same direction of fit as believing that **P**. All four attitudes represent things as being a certain way. However, while believing that **P** and hypothesising that **P** are plausibly thought of as incorrect or subject to revision if it is not true that **P**, the same cannot be said of the attitudes of supposing and pretending that **P**. Although supposing that **P** and pretending that **P** represent things as being a certain way, neither is necessarily incorrect or subject to revision if **P** is false. Consequently, if we assume, in keeping with the Anscombe interpretation, that an attitude has mind-to-world fit only if it is incorrect or subject to revision when there is a mismatch between the attitude and the world, then we seem forced to deny that supposing and pretending have mind-to-world fit. In order to avoid this difficulty, there may be a temptation to weaken the notion of "incorrectness" implicated by the Anscombe interpretation so that it can plausibly be applied to the attitudes of supposing and pretending. However, there does not seem to be any plausible sense in which pretending is incorrect if things turn out to be different from how they are pretended to be. On the contrary, it would be more plausible to say that pretending aims to represent things in a way that they are not.

---

<sup>25</sup> Price [1989: 120].



The preceding observation suggests a possible diagnosis of where the Anscombe interpretation goes wrong. To say that a particular attitude has mind-to-world fit is not yet to say anything about when that attitude is incorrect or subject to revision. Whether or not an attitude is incorrect or subject to revision seems tied to its aim, and not to its direction of fit. Hence, a belief is correct only if the proposition believed is true because belief aims at truth.<sup>26</sup> Since the attitude of pretending, by contrast, does not aim at the truth, it is not correct only if the proposition pretended is true. In order to avoid this difficulty, we need to make sense of the direction of fit metaphor in a way that preserves the distinction between saying that an attitude has a certain aim (and consequently, certain correctness conditions) and the claim that an attitude has a particular direction of fit. I submit that the Anscombe interpretation is unsatisfactory because it falls short in precisely this respect.

Putting aside the above difficulty with the Anscombe interpretation, there is an additional problem with Anscombean GG theory that I wish to highlight. When the claim that a desire represents the desired action as good is combined with the Anscombe interpretation, it seems to follow that a desire has mind-to-world fit. Recall, according to Anscombean GG theory, a desire is correct only if the desired action is good. This suggests that it is the desire, rather than the world, that is incorrect or subject to revision if the desired action is not good. For example, suppose I desire to kick a puppy. Suppose further that it would not be good to kick a puppy. Under such circumstances, we certainly would not wish to say that it is the world—that is, the fact that it would not be good to kick a puppy—that is at fault and should be revised to suit my desire. Rather, it is my desire to kick a puppy that should be revised, given up, or be left unsatisfied. If this is right, then the Anscombe interpretation entails that the desire to kick a puppy has mind-to-world fit.

From the point of view of the Anscombean GG theory, the above conclusion is problematic since it links the aim of a desire to the fact that it has mind-to-world fit. Recall, according to (I), a desire aims at the good in virtue of the fact that it represents the desired outcome as good. Moreover, we just observed that a desire has mind-to-world fit in virtue of the fact that it represents the desired outcome as good. This means that a desire both aims at the good and has the mind-to-world in virtue of the fact that it represents the desired outcome as good. Hence, the fact that a desire aims at the good is connected to the fact that it has mind-to-world fit—to wit, it has both properties in virtue of the fact that it represents the desired outcome as good. The end result is that, contra (IV), the aim of a desire is linked to the fact that it has mind-to-world fit, rather than to the fact that it has world-to-mind fit. This is a very surprising and unhappy result. One would have expected that the feature in virtue of which a desire aims at the good be identified with or closely linked to the feature in virtue of which it has a different direction of fit to belief. This is precisely the view to which (IV) commits the advocate of Anscombean GG theory. But if the present objection is right, the feature in virtue of which a desire aims at the good turns out to be exactly the same feature in virtue of which a desire has the same direction of fit as belief; namely, the fact that a desire represents the desired outcome as good.

---

<sup>26</sup> For a discussion of some of the different ways we can understand the claim that belief aims at truth, see Engel [2004].

### 2.3. The Independence Thesis

The preceding analysis calls attention to two shortcomings in Anscombean GG theory. First, the Anscombe interpretation of the direction of fit metaphor is unable to accommodate the intuition that attitudes such as supposing and pretending have the same direction of fit as belief. Second, when the Anscombe interpretation is combined with (II), it does not only have the unhappy consequence that a desire has mind-to-world fit, but also that it has mind-to-world fit in virtue of the very same property in virtue of which it aims at the good. The end result is that the claim that desires aim at the good cannot satisfy the requirements of (IV), according to which the aim of a desire is linked to the fact that it has world-to-mind fit. In light of these difficulties with Anscombean GG theory—and with the Anscombe interpretation of the direction of fit metaphor, in particular—I wish to propose an alternative interpretation of the notion of direction of fit, one that (i) preserves the intuition that the aim of a desire is something that explains, is explained by, or otherwise linked to the fact that it has world-to-mind fit, and (ii) is able to accommodate the intuition that attitudes like supposing and pretending have the same direction of fit as belief. To this end, I claim that attitudes with mind-to-world fit have a different type of content (in a sense soon to be specified) to attitudes with world-to-mind fit. I will refer to this interpretation of the direction-of-fit metaphor as the *two-content interpretation*.

Before we embark on a detailed discussion of the two-content interpretation, it will be helpful to address a few terminological issues. Typically, an assertion is defined as a speech act in which a proposition is represented as being true.<sup>27</sup> It is widely held that an assertion may share the same propositional content as an attitude, such as belief, hope or doubt. For example, the assertion that Goldbach's conjecture is true has the same propositional content as the belief that Goldbach's conjecture is true. Moreover, an agent may use the assertion that Goldbach's conjecture is true to express her belief that Goldbach's conjecture is true. I will refer to an assertion that expresses a belief of the agent making the assertion as a *sincere assertion*. There are cases in which an agent makes a putative assertion, but in which the putative assertion does not express one of the speaker's beliefs. I will refer to such putative assertions as *insincere assertions*. Significantly, to say that a putative assertion is insincere, in the present sense, is not to say that it is somehow infelicitous. Whether an insincere assertion is infelicitous will depend on other factors, such as the context of utterance. For example, if an insincere assertion is made in the context of a theatrical performance, the fact that it does not express a belief of the speaker does not immediately render the assertion infelicitous. However, in a context in which an agent is asked to give her honest opinion on a matter, it would be infelicitous for her to make an insincere assertion.

An assertion may be codified in a sentence (written, spoken, or signed) by which the asserted proposition is conveyed. I will refer to particular tokens of such written, spoken, or signed sentences as *utterances*. Two or more different utterances may be used to assert the same proposition. For example, the utterances "It is raining today" and "Es regnet heute" may both be used to assert the proposition 'it is raining today'. By necessity, I will be employing written sentences to express the various propositions discussed in this paper. However, it is important to keep in mind that these written sentences will be merely standing proxy for the propositions

---

<sup>27</sup> If one is an anti-realist about propositions, one may replace talk about propositions with talk about that-clause complements.

they are typically used to convey. In order to avoid confusion on this score, I will use double quotation marks to indicate when I'm talking about the utterance "It is raining today" and single quotation marks to indicate that I am referring to the proposition 'it is raining today'.

In addition to assertoric utterances, there are also non-assertoric utterances, such as questions and commands. Assertions, questions, and commands differ in their illocutionary force.<sup>28</sup> However, it is sometimes assumed that utterances with different illocutionary force may have the same propositional content. For example, consider the following three utterances:

- (A): "The office door is shut."
- (B): "Is the office door shut?"
- (C): "Shut the office door!"

It is sometimes held that (A), (B), and (C) share the same propositional content; namely, the proposition: 'the office door is shut'.<sup>29</sup> The difference between the three utterances has to do with the illocutionary force with which this single proposition is expressed, with the proposition being asserted in (A), questioned in (B), and commanded in (C). Since, according to this view, it is possible to vary the force of an utterance while keeping the content fixed, it follows that content and force are independent features of an utterance. I will refer to this view as the *independence thesis*.<sup>30</sup>

## 2.4. Motivating Illocutionary Content

There are two features of the notion of content presupposed by the independence thesis that makes it less than ideal for our present purposes. First, the independence thesis relies on a notion of content that leaves the logical properties of an utterance underspecified. This is because a complete characterisation of the logical properties of an utterance—including the specification of the various deductive inferences for which a particular utterance may be employed—is at least partly determined by the sentence's illocutionary force. Recall, according

---

<sup>28</sup> See and Cf. Searle and Vanderveken [1985].

<sup>29</sup> While this is not a view I endorse, it is nevertheless one that I wish to accommodate. However, it is worth noting that, among philosophers who specialise in the semantics of questions, the dominant view is that the content of a question is identical to or to be understood in terms of its answerhood conditions, where "answerhood conditions" is taken to refer to the set of all possible answers, true answers, partial answers, etc. According to this view, which we may refer to as *Hamblin question-semantics*, a question picks out a set of propositions, rather than a single proposition (see: Hamblin [1973] and Karttunen [1977]. For overview of the relevant literature, see: Groenendijk and Stokhof [1997], Higginbotham [1996], and Ginzburg [1996]). Thus, many philosophers would reject the claim that (A) and (B) share the same propositional content. Insofar as such philosophers are committed to saying that (A) and (B) have different types of content, I would consider them allies since, as will soon become clear, I wish to make a similar claim. However, in order to make my account appealing to the widest possible audience, I will attempt to set things up in a way that is neutral between the assumption that (A) and (B) have the same propositional content and Hamblin question-semantics. Similarly, there is a growing number of theorists who reject the claim that (A) and (C) share the same propositional content (e.g., Chellas [1971], Vranas [2008; 2010], Warnock [1976], and Stalley [1972]). I also view such theorists as allies, but will endeavour to remain neutral on this issue, here.

<sup>30</sup> For a defence of the independence thesis, see Stenius [1967].

to the independence thesis, (A), (B), and (C) all have the same content: the proposition 'the office door is shut'. If we assume that this content is sufficient to determine the logical character of all three utterances, then it would follow that (A), (B), and (C) should be logically interchangeable, despite their contrasting illocutionary force. But this does not seem to be the case. For example, the following is a valid deductive inference:

- (A1): "The office door is shut."
- (A2): "If the office door is shut, then Professor Smith is away."
- (A3): "Therefore, Professor Smith is away."

However, neither of the following appears to be a valid deductive inference:

- (B1): "Is the office door shut?"
- (B2): "If the office door is shut, then Professor Smith is away."
- (B3): "Therefore, Professor Smith is away."

and

- (C1): "Shut the office door!"
- (C2): "If the office door is shut, then Professor Smith is away."
- (C3): "Therefore, Professor Smith is away."

The problem with (B1)-(B3) and (C1)-(C3) is that in both cases, the initial premise fails to satisfy the antecedent of the conditional specified in the second premise ((B2) and (C2), respectively). This leaves us without any basis for inferring the consequent of the conditional, as specified in (B3) and (C3), respectively. Moreover, the failure of (B1) and (C1) to satisfy the antecedent of (B2) and (C2) is directly due to their illocutionary force. (B1) and (C1) both fail to satisfy the antecedent of (B2) and (C2) precisely because they fail to depict the proposition 'the office door is shut' as true. In fact, (B1) and (C1) are both perfectly consistent with the falsity of the proposition 'the office door is shut'. The same, of course, cannot be said of (A1), which is clearly inconsistent with the falsity of the proposition 'the office door is shut'. Thus, although according to the independence thesis, (A1), (B1), and (C1) share the same content, (A1) is logically inconsistent with the negation of the proposition 'the office door is shut', while (B1) and (C1) are both logically consistent with the negation of the proposition 'the office door is shut'. Given that the only difference between (A1), (B1), and (C1) has to do with their illocutionary force, it follows that illocutionary force makes a logical difference. The end result is that if we aspire to have a complete characterisation of the logical properties of an utterance, then we must take into consideration its illocutionary force.

The notion of content that interests me in the present investigation is one that corresponds with a complete characterisation of the logical properties of an utterance—to wit, one that includes illocutionary force. As such, I will not be presupposing the notion of content implicated by the independence thesis. To avoid confusion, and in order to emphasise that the notion of content currently at play is one that includes illocutionary force, I will employ the expression "illocutionary content" as an umbrella term for the notion of content that corresponds with a logically complete characterisation of (A), (B), and (C). There are three

distinct types of illocutionary content that are relevant to the present discussion: indicative content (e.g., 'the office door is shut'), interrogative content (e.g., 'is the office door shut?'), and imperative content (e.g., 'shut the office door!'). According to the present view, just as a sincere assertion expresses the attitude of believing, so, too, does a sincere question express the attitude of wondering and a sincere request expresses an attitude of wanting. Moreover, just as an assertoric utterance may be used to convey indicative content, an interrogative utterance may be used to convey interrogative content and an imperative utterance may be used to convey imperative content. Hence, the illocutionary content of an utterance or attitude corresponds with that which distinguishes a case of believing or asserting from a case of wondering or questioning, and distinguishes both from a case of wanting or requesting.

Having such a notion of content is generally important because we assess an agent's rational standing in light of the illocutionary content of their utterances and attitudes, and not in terms of the propositional content of their utterances and attitudes. For example, if an agent believes or sincerely asserts that the office door is shut, we take them to have or express an attitude that is inconsistent with being agnostic about whether or not the office door is shut. Thus, if an agent were to adopt both an attitude of belief and an attitude of agnosticism towards the office door being shut (assuming that such a combination of doxastic attitudes is possible), we would deem them guilty of irrationality. However, consider the case of an agent who wonders if the office door is shut. Such an agent has an attitude that may be expressed by sincerely asking if the office door is shut. Unlike the agent who believes or sincerely asserts that the office door is shut, the agent who wonders or sincerely asks if the office door is shut has or expresses an attitude that is consistent with being agnostic about the office door being shut. Such an agent would not be liable to rational criticism for simultaneously having both attitudes. However, according to the independence thesis, wondering if the office door is shut, sincerely asking if the office door is shut, believing that the office door is shut, and sincerely asserting that the office door is shut all have the same propositional content: namely, the proposition 'the office door is shut'. Thus, we could not hope to evaluate an agent's rational standing by simply considering the propositional content of their utterances or attitudes. We must look, instead, to the illocutionary content of an agent's utterances and attitudes.

The notion of illocutionary content is also of specific importance to the task of providing an adequate characterisation of the differing directions of fit of beliefs and desires. If we conceive of a command as expressing a desire that someone bring about some outcome, then we may view a desire as akin to a self-issued command that one bring about some outcome. On this view, a desire has the same illocutionary content as an imperative utterance: namely, imperative content. By contrast, a belief, as we have already observed, has the same illocutionary content as an assertoric utterance: namely, indicative content. In order for an attempt to make sense of the direction of fit metaphor to be adequate, it must not only account for the fact that beliefs and desires have different directions of fit, but it must also account for the fact that a belief has the same direction of fit as the attitudes of supposing and pretending.

With respect to the immediately preceding point, I believe that the two-content interpretation is superior to the Anscombe interpretation. According to the two-content interpretation, beliefs have the same direction of fit as the attitudes of supposing and pretending because all three have indicative content, while desires have the same direction of fit

as wishes and hopes because all three have imperative content. Insofar as indicative content is truth-evaluable, the present view entails that beliefs, supposings, and pretendings all have truth-evaluable illocutionary content. However, what sets the attitude of belief apart from the attitudes of supposing and pretending is that a belief is correct only if its illocutionary content is true. This is what it means to say that belief aims at truth. On the present account, since it is possible for a supposing or pretending to represent matters in a way that they are not, it is also possible for their illocutionary content to be false. However, it does not follow from the fact that the illocutionary content of a supposing or pretending is false that either attitude is incorrect or subject to revision. This is what it means to say that supposing and pretending do *not* aim at truth. Hence, the present framework is able to preserve the distinction between saying that belief has a certain direction of fit and saying that belief has a certain aim.

## 2.5. Desires, Wishes, and Wants

The truism that there are other attitudes—such as hypothesising or pretending—that share the same direction of fit as belief is often overlooked or ignored. However, most theorists, if pressed, would agree that it is a truism. Likewise, it is relatively uncontroversial that other attitudes—such as wishing or wanting—share the same direction of fit as desire. However, it is not obvious that there is a sharp distinction between our quotidian conception of a desire and our quotidian conception of a want or wish. In fact, there are contexts in which the terms are treated as synonymous. For example, saying that one wants to visit Spain, that one wishes to visit Spain, and that one desires to visit Spain may all be taken to express the same attitude.

Nevertheless, it does not follow from the fact that a distinction is not already part of our pre-theoretical conception that there is no distinction to be made or worth making. In order to facilitate greater conceptual clarity and ease of philosophical analysis, I propose the following theoretical operationalization of the words “want”, “desire” and “wish”. I will use the word “want” (and its various cognates) as a generic term for all attitudes with world-to-mind fit. Attitudes with world-to-mind fit are often referred to as conative attitudes. However, this label may be misleading since the word conative literally refers to motivational or volitional attitudes, and some attitudes with world-to-mind fit are not properly regarded as motivational or volitional. For example, if I want it stop raining or I want the West Indies team to win the cricket Test Match, I have an attitude with world-to-mind fit. However, these attitudes may not motivate me to act or change my behaviour in any way. Thus, the set of attitudes picked out by the word “wanting” is strictly speaking broader than the set of attitudes picked out by the label “conative attitude”. For this reason, I will use the word “want” as opposed to the expression “conative attitude” as my generic term for all attitudes with world-to-mind fit.

I will use the label “desire” to refer to cases of wanting that are directed at outcomes an agent believes she can bring about directly via the deliberate exercise of her agency.<sup>31</sup> I will

---

<sup>31</sup> In this respect, my usage of the word “desire” loosely corresponds with that of Velleman [1992: 17]. However, while Velleman uses the word “desire” to refer to cases of wanting that are directed at outcomes an agent believes to be attainable (where “attainable” is defined as “a possible future outcome”), I use the word “desire” to pick out cases of wanting that are directed at outcomes an agent believes she can bring about directly via the deliberate exercise of her agency. On my usage of the term, an agent who wants it to stop raining (but who believes that stopping the rain

refer to any outcome an agent may bring about directly via the deliberate exercise of her agency as *doable*. Hence, a desire is a case of wanting in which the outcome wanted is believed to be doable. According to the present stipulation, one may desire to have a glass of wine (assuming that one believes that having a glass of wine is an outcome one may bring about directly via the deliberate exercise of one's agency), but one may not desire that it stop raining or that the West Indies cricket team win the Test Match (assuming that one believes that stopping the rain and making the West Indies cricket team win are not outcomes one may bring about directly via the deliberate exercise of one's agency). By contrast, I will be using the label "wish" to refer to a case of wanting that is directed at an outcome that an agent believes she cannot bring about directly via the deliberate exercise of her agency. According to the present stipulation, an agent may wish that it stop raining or that the West Indies cricket team win the Test Match, but she may not wish to have a glass of wine (assuming that she believes that having a glass of wine is an outcome she may bring about directly via the deliberate exercise of her agency).<sup>32</sup>

With the preceding terminological distinction in place, we may now see that the two-content interpretation allows us to consistently hold that desires and wishes have the same direction of fit, and yet deny that they share the same correctness conditions. To this end, I hold that desires and wishes both have imperative content. As such, it follows from the two-content interpretation that they have the same direction of fit. However, I maintain that a desire is correct only if the outcome desired is one the agent can bring about directly via the exercise of her own agency. In short, a desire is correct only if the outcome desired is doable. For example, suppose that it is not doable for me to walk across the River Thames. To wit, my walking across the Thames is not an outcome I can bring about directly via the deliberate exercise of my agency. If I were to desire to walk across the Thames (a desire I could have only if I also had the false belief that walking across the Thames is doable), then my desire would be incorrect. By contrast, a wish may be correct even if the outcome wished for is not doable. For example, if I wished that I could walk across the Thames (a case of wanting that counts as a wish because I had the true belief that walking across the Thames is not doable), then my wish may be nevertheless correct. Hence, my desire and my wish have different correctness conditions even though they have the same direction of fit. The end result is that the two-content interpretation allows us to preserve the distinction between saying that two attitudes have the same direction of fit and saying that they have the same correctness conditions.

---

is not an outcome she can bring about) does not have a desire. She has a wish. By contrast, according to Velleman's taxonomy, an agent who wants it to stop raining still counts as having a desire, so long as she believes that the rain stopping is a possible future outcome. Hence, although my way of distinguishing between desires and wishes is similar to Velleman's, the two taxonomies carve up the space of attitudes with world-to-mind fit differently.

<sup>32</sup> It bears repeating that the present terminological stipulation, while largely inspired by our quotidian usage of the words "desire" and "wish", is not meant to exactly mirror said usage. Rather, it aims to introduce a level of theoretical precision not already found in our ordinary linguistic practice in order to better facilitate philosophical analysis.

## 2.6. Presuppositions and Illocutionary Implications

At the end of the previous section, I suggest that what distinguishes a desire from a wish is that the former is correct only if the outcome wanted is doable. This introduces a puzzle. How can we consistently affirm that a desire is correct only if the desired action is doable and yet deny that a desire represents its object as doable? After all, if one were to hold that a desire represents its object as doable, then there is an obvious and intuitive sense in which a desire gets things wrong if the desired action is not doable. Simply put, the desire represents matters incorrectly. However, if we hold that a desire does not represent its object as doable, then we seem to be left with no basis for saying that a desire gets things wrong if the desired outcome is not doable. If, on the other hand, we grant that a desire does represent the desired outcome as doable, then we seem forced to conclude that desires have indicative content after all. My strategy for solving this puzzle is to argue that a desire presupposes (in a sense to be specified momentarily) that its object is doable, and may therefore be described as correct only if its object is doable. This will allow me to say that a desire is correct only if its object is doable, even though a desire does not represent its object as doable.

In non-technical terms, a presupposition is defined as information that is assumed or taken for granted by an individual or utterance.<sup>33</sup> For example, suppose I were to make the following sincere assertion:

(E1): “Usain Bolt has won the 100 meters once again.”

In so doing, I may plausibly be said to presuppose that the following propositions are true:

(E2) ‘There is a (salient and identifiable) Usain Bolt.’

(E3) ‘There is a (salient and identifiable) 100 meters.’

(E4) ‘Usain Bolt has won the 100 meters at least once before.’

The technical notion of a presupposition that I wish to introduce here aims to mirror the results of the non-technical notion. Thus, I wish to preserve the intuition that (E1) presupposes (E2), (E3), and (E4). However, the technical notion derives much of its theoretical motivation from Frege’s observation “[t]hat the name ‘Kepler’ denotes something is just as much a presupposition for the assertion ‘Kepler died in misery’ as for the contrary assertion.”<sup>34</sup> Within Frege’s framework, a presupposition is conceived of as a definedness condition, a requirement

---

<sup>33</sup> Describing a presupposition as an “assumption” is potentially controversial, if by “assumption” we mean the assumption of an agent. Theories of presupposition may be divided into *pragmatic* and *semantic*. According to the standard pragmatic theory, presuppositions are tied to the attitudes of a particular agent. For example, according to Stalnaker [1974], an utterance presupposes a proposition because the speaker took the proposition for granted when she made the utterance. By contrast, according to the standard semantic theory, presuppositions are tied to sentences, rather than to the attitudes of an agent. For example, according to one interpretation of Strawson [1950], an utterance presupposes a proposition because the truth of the proposition is a necessary condition for the utterance to denote, be meaningful, or have a truth-value. The present use of the word “assume” is not an attempt to take a stand on this debate but is rather meant to convey the general significance of the quotidian usage of the word.

<sup>34</sup> Frege [1892: 168].



for an expression to have a truth-value. Strawson takes up this idea by defining a presupposition along the following lines:

**Definition 2.1 (Strawsonian Presupposition):**

Sentence X presupposes sentence Y *IFF* Y is true whenever X is true or false.<sup>35</sup>

Significantly, Strawson is primarily concerned with when one sentence presupposes another. As such, he is interested in offering a semantic theory of presuppositions. I do not wish to impugn either the need for or the prospects of having a successful semantic theory of presuppositions. However, at present, I am interested in utterances and attitudes, not sentences. Consequently, I am not interested in offering a semantic theory of presuppositions. Nonetheless, I wish to incorporate the central intuition behind the Frege-Strawson approach into my own by claiming that if the truth of a certain proposition is presupposed (rather than asserted) by an utterance, then it should be implied (in some sense) by both the utterance and its negation.

The Frege-Strawson intuition that I wish to incorporate into my utterance-oriented account may be summarised as follows. On the one hand, saying that an utterance, X, presupposes some proposition, Y, means that an agent who utters X can be expected to believe Y. This suggests that if Y is presupposed by X, then Y should be implied (in some sense) by X. On the other hand, insofar as a presupposed proposition is not asserted by an utterance, then accepting the proposition(s) asserted by an utterance should not be a prerequisite for accepting a presupposed proposition. This means that whatever property of an utterance in virtue of which it implies a presupposed proposition should be equally effective in giving rise to the relevant implication even if one were to reject all the propositions that the utterance asserts. Thus, the implication should also hold if one were to make an utterance that was like the original utterance in every respect except that all of the propositions asserted in the original utterance were now denied. Otherwise one would be justified in concluding that the proposition was a member of the set of propositions asserted by the original utterance, in which case it would fail to qualify as a presupposed proposition. The end result is that Y may be said to be presupposed by X if and only if X implies Y and  $\neg X$  implies Y. I will refer to this technical conception of a presupposition as an *agential presupposition*.<sup>36</sup>

---

<sup>35</sup> Strawson [1952]. For a discussion of Strawson's proposals, see: Sellers [1954]; Strawson [1954]; and Nerlich [1967].

<sup>36</sup> Survival under negation is widely regarded as the most important diagnostic test of a presupposition. A presupposition that passes the negation-embedding test is said to be "projected" under negation-embedding. However, projection under negation-embedding is typically seen as falling short of a necessary condition for something to count as a presupposition because there are circumstances in which presuppositions fail to project. Indeed, attempting to explain why presuppositions sometimes fail to project is one of the central problems in presupposition theory. The notion of presupposition presently on offer departs from the standard picture since it takes passing the negation-embedding test to be both necessary and sufficient for a proposition to be presupposed by an utterance or attitude. This is largely due to the fact that the current technical account of a presupposition does not exploit the notion of "projection" that exercises so much of presupposition theory. I believe I am allowed to get away with this because I am presently concerned with the presuppositions of an agent, as indicated by her utterances and attitudes, rather than with the presuppositions of a sentence. Consequently, the problem of projection failure simply fails to arise in the present context. It is also worth noting that, because agential presuppositions are

It is important to get clear on the notion of implication at play in the current context. Specifically, classic semantic entailment seems ill suited for our present purposes. Firstly, if we model the present notion of an “implication” on classic semantic entailment, then  $X$  implies  $Y$  just in case it is impossible for  $X$  to be true and  $Y$  to be false. However, since I will be considering utterances and attitudes that are not truth-evaluable (or, to be more precise, utterances and attitudes whose illocutionary content is not truth-evaluable), the classic notion of semantic entailment will not do for present purposes. Secondly, according to classic semantic entailment, if  $X$  implies  $Y$  and  $\neg X$  implies  $Y$ , then it follows that  $Y$  is a tautology. This is a very unhappy result since it would follow from our definition of an agential presupposition that all presuppositions are tautologies. In order to avoid the above difficulties, I wish to appeal to a notion of implication that is both broader and weaker than classic semantic entailment. I will refer to this notion of implication as an illocutionary implication (*i*-implication), which I define as follows:

**Definition 2.2 (Illocutionary Implication):**

$X$  *i*-implies  $Y$  IFF an utterance of  $X$  indicates that the speaker believes  $Y$ .<sup>37</sup>

Illocutionary implication is broader than classic semantic entailment since it has application not only to utterances with truth-conditions (i.e., ones with indicative content), but also to utterances with doability-conditions (i.e., ones with imperative content). It is also weaker than the classic semantic entailment since it does not track what propositions are entailed by an utterance, but what an utterance indicates that a speaker believes.<sup>38</sup> With the broader notion of an *i*-implication in hand, we may now define an agential presupposition as follows:

**Definition 2.3 (Agential Presupposition):**

$X$  presupposes  $Y$  IFF  $X$  *i*-implies  $Y$  and  $\neg X$  *i*-implies  $Y$ .

Definition 2.3 provides us with a diagnostic test for determining if a particular proposition is agentially presupposed by an utterance. On the current view,  $X$  presupposes  $Y$  just in case  $X$  *i*-implies  $Y$  and there is some  $Z$  such that  $Z$  is formed by embedding  $X$  under a negation operator, and  $Z$  *i*-implies  $Y$ . Hence, we may determine if a proposition that is *i*-implied by an utterance is agentially presupposed by that utterance by first embedding the illocutionary content of the utterance under a negation operator, and then checking to see if the *i*-implication still holds.

---

concerned with utterances and attitudes, it bears a number of similarities to Stalnaker’s pragmatic conception of a presupposition (Stalnaker [1972; 1973; 1974; 1998]). See and Cf. Bas van Fraassen [1968]. For an excellent overview of presupposition literature, see Soames [1989].

<sup>37</sup> See and Cf. Beaver [2001: 15].

<sup>38</sup> The notion of an utterance indicating what the speaker means is here deliberately left vague in order to remain as ecumenical as possible. The reader is therefore free to unpack the notion in terms of her preferred theoretical account. I believe it would be particularly fruitful to consider the present account of a presupposition in light of Stalnaker’s notion of a pragmatic presupposition. However, I will not attempt to do that here since it would take us too far afield from our present concerns. For a discussion of Stalnaker’s pragmatic conception, see: Stalnaker [1972, 1973, 1974, 1998]. Like Stalnaker’s notion of a pragmatic presupposition, my notion of an agential presupposition is not inconsistent with the existence of a Frege-Strawson-style semantic presupposition.

Call this the *negation-embedding test*. For example, if we embed (E1) under the negation operator we arrive at (E1\*):

(E1\*) “Usain Bolt has not won the 100 meters once again.”

According to the negation-embedding test, if (E2)-(E4) are *i*-implied by both (E1) and (E1\*), then (E2)-(E4) are agentially presupposed by (E1). An examination of (E1) and (E1\*) reveals that both *i*-imply (E2)-(E4). That is to say, an utterance of (E1), as well as an utterance of (E1\*), indicate that the speaker believes the propositions (E2), (E3), and (E4). Consequently, (E2), (E3), and (E4) all pass the negation-embedding test. It follows that (E2)-(E4) are agential presuppositions of (E1). Contrast this with the proposition (E5):

(E5) ‘Usain Bolt did not lose the 100 meters.’

It is clear that (E1\*) does not *i*-imply (E5), since an utterance of (E1\*) would not indicate that the speaker believes (E5). Consequently, (E5) fails the negation-embedding test; it is not an agential presupposition of (E1).

## 2.7. Desires, the Doable, and the Good

Let us now return to the question of how a desire may be correct only if the desired action is doable even though a desire does not represent its object as doable. The first step towards solving this puzzle is to register that imperative utterances also have presuppositions. For example, consider the following imperative utterance:

(F1): “Shut the office door!”

(F1) has the following putative presuppositions:

(F2): ‘There is an (salient and identifiable) office door.’

(F3): ‘The office door is open.’

It should be obvious that an utterance of (F1) indicates that the speaker believes (F2) and (F3). We may therefore conclude that (F1) *i*-implies (F2) and (F3). However, it is not immediately clear that (F1) is a candidate for the negation-embedding test. After all, I have been insisting that imperative content is not truth-evaluable. How then are we to characterise the negation of a non-truth-evaluable item? In response to this problem, I propose that we conceive of the negation of, for instance, a command to  $\phi$  in terms of a command to refrain from  $\phi$ -ing. Hence, we may form the negation of (F1) by embedding it under a “do not” operator, in order to form (F1\*):

(F1\*): “Do not shut the office door!”

We may now ask if (F1\*) *i*-implies (F2) and (F3). That is to say, would an utterance of (F1\*) indicate that the speaker believes (F2) and (F3)? Since it would be pointless to command someone not to shut the office door if there were no (salient and identifiable) office door and if the office door were not open, an utterance of (F1\*) does indicate that the speaker believes (F2) and (F3). Consequently, (F2) and (F3) both pass the negation-embedding test for imperative utterances, and are therefore agential presuppositions of (F1).

With the just introduced theoretical apparatus for determining the agential presuppositions of an imperative utterance now in hand, we may ask if the command to shut the office door agentially presupposes that shutting the office door is doable. In response to this question, I submit that the set of propositions that must be true in order for shutting the office door to be doable is always a subset of the set of propositions agentially presupposed by the command to shut the office door. Or, to put the same point more generally and in terms of a conditional, if a certain proposition must be true in order for an imperative utterance to be doable, then that imperative utterance agentially presupposes that proposition. Let Y stand for a proposition, X stand for an imperative utterance, and  $\neg X$  stand for the imperative utterance formed by embedding X under the “do not” operator. The claim that an imperative utterance always agentially presupposes a proposition that must be true in order for the imperative to be doable may be demonstrated by the following argument:

- (1) If Y must be true for X to be doable, then X *i*-implies Y. (Premise defended below)
- (2) If Y must be true for X to be doable, then  $\neg X$  *i*-implies Y. (Premise defended below)
- (3) If Y must be true for X to be doable, then X *i*-implies Y and  $\neg X$  *i*-implies Y. (From (1) and (2))
- (4) If X *i*-implies Y and  $\neg X$  *i*-implies Y, then X agentially presupposes Y. (From Definition 2.3)
- (5) If Y must be true for X to be doable, then X agentially presupposes Y. (From (3) and (4))

Premise (1) rests on the intuition that it would be pointless to command someone to perform an action they could not perform. As such, a sincere imperative utterance would indicate that the speaker believes that the imperative is doable. Since A *i*-implies B just in case an utterance of A would indicate that the speaker believed B, a sincere imperative utterance always *i*-implies that the imperative is doable. Premise (2) rests on the intuition that it would be pointless to command someone to refrain from performing an action that they could not perform. For example, it would be pointless to issue the command, “Do not shut the office door!” if it was not possible to shut the office door. If this intuition is right, then  $\neg X$  always *i*-implies those propositions that must be true in order for X to be doable. The other premises of the argument follow from (1), (2), and the definition of an agential presupposition, as specified in (4). We may therefore conclude that the set of propositions that must be true for an imperative utterance to

be doable is always as subset of the set of propositions that are agentially presupposed by that imperative utterance.<sup>39</sup>

Let us conclude the present section by considering the question of whether a desire agentially presupposes that the desired action is good. If the GG theorist is going to claim that, for example, the command to shut the office door agentially presupposes that shutting the office door is good, then she will need to show that (F4), or some comparable proposition, passes the negation-embedding test for (F1):

(F4): It is good to shut the office door.

However, in order to pass the negation-embedding test for (F1), (F4) must be *i*-implied by both (F1) and (F1\*).

(F1): "Shut the office door!"

(F1\*): "Do not shut the office door!"

But this is clearly not the case. Even if a sincere utterance of (F1) indicates that the speaker believes (F4)—a claim that strikes me as dubious—a sincere utterance of (F1\*) certainly does not indicate that the speaker believes (F4). Therefore, it follows from the definition of an *i*-implication that (F1\*) does not *i*-imply (F4). (F4) therefore fails the negation-embedding test for (F1).

The preceding results fits hand in glove with my characterisation of a desire as an instance of wanting in which the outcome wanted is believed to be doable. According to the presupposition analysis, a desire is correct only if the outcome desired is doable because the imperative that constitutes its illocutionary content indicates that the person with the desire believes that it is doable. This comports with my stipulation that a desire only counts as such if the outcome desired is believed to be doable. It immediately follows from my definition of a desire that desiring to bring about **P** always implicates the belief that **P** is doable, and the fact that an agent has a desire may safely be taken to indicate that the agent believes **P** to be doable. It also immediately follows from my definition of a desire that the illocutionary content of the desire to bring about **P** *i*-implies that **P** is doable. Given my characterisation of the aim of an attitude as a condition whose satisfaction is necessary for the attitude to be correct, it follows that the aim of a desire (i.e., the doable) is intimately connected both to what makes a desire a

---

<sup>39</sup> It is worth noting that the claim that a desire (or more accurately, its illocutionary content) agentially presupposes that the outcome desired is doable is not the counterpart to (nor is it offered as a substitute for) the GG theorist's claim that desire aims at the good. According to the GG theorist, desire aims at the good in roughly the sense that belief aims at the truth. The GG theorist therefore relies on an analogy between desire and belief or between the practical and the theoretical. By contrast, I am committed to no such analogy. On the contrary, one of the goals of this dissertation will be to argue that the practical and theoretical spheres are disanalogous in ways that problematize the heavy reliance on analogies of which the GG theorist is guilty. Moreover, this allows me to sidestep the worry that the doable does not (and could not) play the same role for desire that truth plays for belief. Even if this were true, this would not count, in the least, against my claim that a desire agentially presupposes that the desired outcome is doable. My claim that a desire agentially presupposes that the desired outcome is doable is based on the fact that the doable passes the negation embedding test, not on an alleged analogy between desire and belief. Thus, whether or not such an analogy is apt has no bearing on the soundness of my argument.

desire (i.e., the fact that a want only counts as a desire if the outcome wanted is believed to be doable) and to that in virtue of which a desire has world-to-mind fit (i.e., the fact that it has imperative content). However, it bears repeating that there remains a gap between saying that an attitude has imperative content (which is equivalent to saying that it has doability conditions) and saying that it is correct only if it is doable (which is equivalent to saying that it aims at the doable), just as there is a gap between saying that an attitude has indicative content (which is equivalent to saying that it has truth conditions) and saying that it is correct only if it is true (which is equivalent to saying that it aims at the truth). Thus, according to the view presently on offer, the distinction between the aim and direction of fit of an attitude is preserved.

## 2.8. Conclusion

In this chapter, I have set out to accomplish two things. First, to show that the two-content interpretation is superior to the Anscombe interpretation because it allows us to preserve the distinction between saying that an attitude has a certain direction of fit and saying that it has certain correctness conditions. According to the two-content interpretation, desires and beliefs have different directions of fit because they have different kinds of illocutionary content: indicative content in the case of the former, and imperative content in the case of the latter. However, a belief is correct only if it is true, not because it has indicative content, but because it aims at the truth. Consequently, the two-content interpretation preserves the distinction between saying that, for example, a belief has mind-to-world fit and saying that a belief aims at the truth. Second, I have attempted to provide both a theoretical framework for making sense of how an attitude that lacks indicative content could nevertheless have correctness-conditions. Third, I have offered a diagnostic test for determining the correctness conditions of an attitude with imperative content—namely, the negation-embedding test for agential presuppositions. Moreover, we have observed that when the negation-embedding test is applied to both the doable and the good, only the former passes the test.

## 3

## DESIRE AS BELIEF OR PERCEPTION

---

### 3.1. Introduction

Anscombe posits that “truth is the object of judgement, and good the object of ‘wanting’”.<sup>40</sup> In this chapter, I will consider the two most common strategies for unpacking Anscombe’s thesis. They are the *Desire-as-Belief Thesis*, the claim that the desire to bring about **P** is an instance of believing that it is good to bring about **P**, and the *Desire-as-Perception Thesis*, the claim that the desire to bring about **P** is an instance of perceiving that it is good to bring about **P**. What both of these approaches have in common is that they posit that desiring to bring about **P** entails representing it as good to bring about **P**. One putative objection to the Desire-as-Belief and Desire-as-Perception Theses is that a desire cannot be a kind of belief or perception because a desire has a different direction of fit to a belief and perceptual appearance. However, in this chapter, I argue that the two most widely discussed interpretations of the direction-of-fit metaphor—that of Anscombe and Michael Smith, respectively—are actually consistent with the Desire-as-Belief or Desire-as-Perception Theses. I claim that this is symptomatic of the fact that these interpretations fail to get at the key intuition that the direction of fit metaphor is meant to capture—namely, that the attitude of desire displays a different logical structure to that of belief. The two-content interpretation (introduced in chapter 2) is presented as an attempt to take this intuition seriously. Unlike the Anscombe and Smith interpretations, it entails that both the Desire-as-Belief and Desire-as-Perception Theses are false. If successful, the arguments in this chapter will establish that the two-content interpretation is superior to the those offered by Anscombe and Smith, not only because it is immune to the problems (highlighted in this dissertation) that plague these competing interpretations, but also because it comes closer to capturing the central intuition motivating the direction of fit analysis. Moreover, since the two-content interpretation entails the falsity of the Desire-as-Belief and Desire-as-Perception Thesis, the arguments in this chapter also attempt to establish that both are at odds with the direction of fit analysis when it is correctly understood.

---

<sup>40</sup> Anscombe [2000: 76].

### 3.2. The (Broadly) Anscombean View

I wish to begin, in the present section, by clarifying what I refer to as the *Broadly Anscombean View*: the thesis that the good is the object of desire in a sense roughly analogous to how truth is the object of belief. I suggest that while it is a view that Anscombe herself does not explicitly defend (a fact that is often overlooked), it nevertheless derives its chief inspiration from Anscombe.

When Anscombe claims that “truth is the object of judgement, and good the object of ‘wanting’”, her use of the word ‘want’ (as opposed to the word ‘desire’) is not accidental. She actually has something significantly different from our ordinary conception of a desire in mind:

‘Wanting’ may of course be applied to the prick of desire at the thought or sight of an object, even though a man then does nothing towards getting the object. . . . The wanting that interests us, however, is neither wishing nor hoping nor feeling nor *desire*, and cannot be said to exist in a man who does nothing towards getting what he wants. (*Italics mine*)<sup>41</sup>

Anscombe's claim that wanting, as she understands the term, does not apply to an agent who does nothing towards getting what he wants, suggests that her notion of wanting may have more in common with an intention than with a desire, as the term is ordinarily understood.<sup>42</sup> According to the pre-philosophical conception of a desire, one may desire to bring about **P** without intending to bring about **P** or doing anything to bring it about that **P**. For example, I may desire to purchase a flat screen television and yet fail to act on my desire because I believe that purchasing a flat screen television is outside my budget.<sup>43</sup> Thus, our pre-philosophical conception of a desire does not require that the desiring agent take steps to bring about the desired outcome in order to count as desiring it.

A significant feature of our pre-philosophical conception of a desire is that it is possible to have a pair of desires that one recognises to be mutually exclusive—such that acting on one desire precludes acting on the other—without being liable to rational criticism. For example, I may *desire* to leave work early (because I have a headache) and I may also *desire* to work late (because I hope to secure a promotion), where these represent mutually exclusive desires (i.e., acting on the first precludes the possibility of acting on the second, and vice versa). Of course, I am liable to rational criticism if I intend to both leave work early and leave work late, while recognising that these are mutually exclusive intentions. However, merely having inconsistent desires does not make one liable to rational criticism. By contrast, if we buy into the

---

<sup>41</sup> Anscombe [2000: 67-68].

<sup>42</sup> As will soon become clear, Anscombe's notion of wanting also differs from my usage of the term in chapter 2, section 2.5.

<sup>43</sup> In at least this respect, the pre-philosophical conception of a desire seems closer to Davidson's notion of a pro-attitude than to Anscombe's notion of wanting, since the latter requires that the agent takes steps towards getting what she wants while the former does not. By Davidson's lights, one may have two conflicting pro-attitudes and yet only act upon one of them. This suggests that one may have a pro-attitude towards some X even though one has taken no steps towards bringing about X. Moreover, on the Davidsonian account, the pro-attitude one fails to act on may still count as something one desired to do. However, by Anscombe's lights, the pro-attitude one fails to act on, in cases in which one has two conflicting pro-attitudes, does not count as an instance of wanting.



Anscombean view that wanting already involves an intention to bring about the desired outcome, then my wanting to both leave work early and leave work late already makes me liable to rational criticism. Thus, Anscombe's notion of wanting has very different implications for when an agent is liable to rational criticism than does our pre-philosophical conception of a desire. Moreover, on the Anscombean view, it does not make sense to say I *wanted* to leave work early (because I have a headache) and that I also *wanted* to work late (because I hoped to secure a promotion), if I ultimately decide to work late rather than leave work early. Given that I never acted on my desire to leave work early, it follows that (by Anscombe's lights) I never really wanted to leave work early.<sup>44</sup>

The preceding considerations suggest that Anscombe's notion of wanting is closer to our ordinary conception of an intention than our ordinary conception of a desire. But if this is right, one may very well wonder why Anscombe did not just say that the good is the object of intention, instead of saying that the good is the object of wanting. This question seems particularly perplexing since it is not clear what attitude Anscombe could be referring to other than an intention. One suggestion is that Anscombe actually has in mind a subset of intentions; namely, intentions that are formed in response to a desire. On the present suggestion, Anscombe uses the word "wanting" to describe a hybrid state, one that entails the presence of both a desire to bring about **P** and an intention to bring about **P**. Arguably, only some of our intentions are formed in response to a desire. For example, suppose I adopt the intention to bring about **P** not because I desire to bring about **P**, but because I believe it is my duty to bring about **P**. In such a case, it would hardly seem appropriate to say that I wanted to bring about **P**. If this is right, then by restricting herself to cases in which an agent wants to bring about **P**, Anscombe *ipso facto* restricts herself to intentions that are formed in response to a desire. One motivation for such a restriction is that it allows her to zero in on desires that an agent acts upon. This would be an attractive move if one thought that the desires an agent acts upon are importantly different from the desires that an agent has not yet acted upon or fails to act upon. For example, one may think that an agent is rationally or morally criticisable for the former in ways in which one is not rationally or morally criticisable for the latter, and that the former therefore deserve special consideration.

My aim, here, is not to argue that Anscombe's conception of wanting ought to be rejected because it fails to correspond with our pre-philosophical conception, nor is it to insist that the pre-philosophical conception of a desire is somehow preferable to Anscombe's notion of wanting. As we just noted, there may be good theoretical motivations for buying into Anscombe's notion of wanting. Rather, it is to register that there are important differences between the two conceptions. However, among those who advocate a guise of the good theory of desires, this difference often seems to go unrecognised. For example, Stampe [1987], Tenenbaum [2003], and Setiya [2007] all take some version of Anscombe's claim that "good is the object of wanting" to apply to desires that fall short of being intended.<sup>45</sup> This is suggested

---

<sup>44</sup> Furthermore, the conception of desire that I am interested in allows for the possibility that I may continue to want to bring about **P** even after I have adopted some incompatible intention. For example, we can imagine that my pounding headache may continue to generate the desire to leave work early, even after I have decided to remain at my desk.

<sup>45</sup> See, Thompson [2008: 104, fn. 12] for discussion.

by the fact that such thinkers typically assume that Anscombe's claim applies to the kind of desires at play in cases of weakness of the will.

On the standard description of weakness of the will—e.g., that found in the discussions of Donald Davidson—the weak-willed agent has two conflicting or mutually exclusive desires, only one of which corresponds with her all-things-considered judgement. Moreover, the weak-willed agent ultimately acts on the desire that is at odds with her all things considered judgement. For example, consider the case in which an agent, Chen, who has the desire to have a slice of chocolate cake and a desire to stick to his diet. However, only one of his desires—namely, his desire to stick to his diet—corresponds with what he judges to be best, all things considered. On the standard conception, should Chen decide to have the slice of chocolate cake anyway, he would be guilty of weakness of the will. But if we assume that the kind of desire at play in cases of weakness of the will entails taking steps to bring about the desired outcome, then it follows that Chen never really had conflicting desires. On the contrary, the desire that corresponds with his all things considered judgement—viz., the desire to stick to his diet—would not count as a desire at all. This suggests that, unlike Anscombe's notion of wanting, the conception of desires typically assumed to be in play in discussions of weakness of the will is one that does not require that the agent take steps to bring about the desired outcome. Moreover, given that most GG theorists are interested in the conception of desire that they take to be involved in cases of weakness of the will—i.e., cases in which an agent finds herself with conflicting desires—it follows that they are not interested in the conception of wanting that requires that the agent take steps towards getting what she wants—i.e., the conception of wanting that we find in Anscombe.

While I stand behind the reading of Anscombe just adumbrated—one that sees Anscombe's notion of wanting as different from the notion of desire typically presupposed in discussions of weakness of the will—this dissertation is primarily concerned with the conception of desire that features in the standard discussions of weakness of will, and which roughly corresponds with our pre-philosophical conception. I wish to explore the possibility of extending Anscombe's analysis of wanting to the conception of desire that features in discussions of weakness of the will and that roughly corresponds with our pre-philosophical conception—i.e., a conception of desire that does not require that the agent take steps towards getting what she desires. However, I will flag the fact that I do not wish to attribute such a position to Anscombe herself by referring to the thesis that the good is the object of desire (in the non-intention-involving sense) as the *Broadly Anscombean View*. While not a view that Anscombe herself advocates (at least in print), the Broadly Anscombean View draws its chief inspiration from Anscombe's account of wanting.

### 3.3. The Desire-as-Belief Thesis

The most widely discussed way of unpacking the Broadly Anscombean view is in terms of the *Desire-as-Belief Thesis*. John Collins defines the Desire-as-Belief Thesis as “the thesis that desire

is a particular kind of belief—that to desire A is simply to believe that A would be good.”<sup>46</sup> We may distinguish between weak and strong versions of the Desire-as-Belief Thesis; the claim that *some* desires are beliefs and the claim that *all* desires are beliefs, respectively. But this distinction, while significant in its own right, will have little bearing on the present discussion. This is because we are only concerned with the Desire-as-Belief Thesis as a means of unpacking the Broadly Anscombean View; namely, the claim that the good is the object of desire. When we unpack the Broadly Anscombean View in terms of the Desire-as-Belief Thesis, we arrive at the claim that desiring to bring about **P** is an instance of believing that it is good to bring about **P**. Now, according to the weak version of the Desire-as-Belief Thesis, only some desires to bring about **P** are instances of believing that it is good to bring about **P**. However, on those occasions in which the desire to bring about **P** is not an instance of believing that it is good to bring about **P**, the desire in question does not have the good as its object. Therefore, the central claim of the Broadly Anscombean View does not apply to such desires. Hence, insofar as we are interested in the Desire-as-Belief Thesis as a means of unpacking the Broadly Anscombean View, we are only concerned with those cases in which a desire may be said to be a type of belief. Consequently, the distinction between the weak and strong versions of the Desire-as-Belief Thesis is superfluous in the context of the present discussion.<sup>47</sup>

In his paper, “Defending Desire-as-Belief”, Huw Price highlights two motivations for the desire-as-belief thesis; a rejection of the Humean theory of desire and a defence of Anti-Emotivism in moral discourse. He writes:

In modern terminology the Humean view is thus that action is a joint product of an agent's beliefs and desires; and that these are distinct kinds of mental states, desires being distinguished from beliefs in virtue of their motivational role. . . . In recent years, however, several philosophers have questioned the Humean orthodoxy. They suggest that certain beliefs might be intrinsically motivational—in effect, in other words, that some or all desires might themselves be beliefs. . . . An attractive feature of the suggestion that (some) desires might be beliefs has been its evident potential to undermine emotivist accounts of moral discourse (and so to enable ethical statements to be brought within the scope of a truth-conditional general semantics).<sup>48</sup>

According to Price, the Humean view amounts to the thesis that desires are intrinsically motivational while beliefs are not. Moreover, Price construes emotivist accounts of moral discourse as entailing the denial of the claim that ethical statements display a truth-conditional semantics. I will refer to the rejection of the Humean view as *Anti-Humeanism*, and the rejection of emotivist accounts as *Anti-Emotivism*.

---

<sup>46</sup> Collins [1988] impugns the Desire-as-Belief Thesis on the grounds that it is incompatible with non-quantitative decision theory, and David Lewis [1988] argues that the desire-as-belief thesis is incompatible with formal theories of Bayesian decision and belief revision. For replies to the objections of Collins and Lewis, along with a defence of the Desire-as-Belief Thesis, see Price [1989] and Humberstone [1987].

<sup>47</sup> Cf. Lewis [1988: 325].

<sup>48</sup> Price [1989: 119].

Although, in the passage just cited, Price associates the Desire-as-Belief Thesis with Anti-Humeanism and Anti-Emotivism, it is worth emphasising that the denial of the Desire-as-Belief Thesis is consistent with both. It is common ground between the Humean and the Anti-Humean that desires are intrinsically motivating.<sup>49</sup> However, the Humean denies, and the Anti-Humean affirms, that some beliefs are also intrinsically motivating. But one may consistently subscribe to the claim that some beliefs are intrinsically motivating and yet deny that desires are beliefs. For example, suppose one held that being intrinsically motivational is a necessary but insufficient condition for an attitude to be a desire.<sup>50</sup> Then one may also consistently hold that desires are distinct attitudes from beliefs (thereby denying the Desire-as-Belief Thesis) and that some beliefs are intrinsically motivational (thereby affirming Anti-Humeanism). Moreover, if we identify the content of moral discourse with the content of an intrinsically motivating belief (rather than with the content of a desire), then we may consistently hold that moral discourse displays a truth-conditional semantics and that the content of desire does not. (One only needs to add the further assumption that the content of an intrinsically motivating belief displays a truth-conditional semantics.) Hence, if we reject the Desire-as-Belief Thesis, it does not follow from this rejection that either Anti-Humeanism or Anti-Emotivism is false.

The preceding observations draw attention to an important feature of Price's account. Price associates the claim that "certain beliefs might be intrinsically motivational" with the claim that "some or all desires might themselves be beliefs", either equating the two claims or taking the second to be entailed by the first. However, the equivalence or entailment only holds if we assume that being motivational is a sufficient condition for an attitude to be a desire. In brief, Price assumes that only desires are intrinsically motivational. However, it seems like both the Humean and the Anti-Humean alike are well within their rights in rejecting this assumption. For example, both the Humean and the Anti-Humean may hold that certain emotions, like love and hatred, are intrinsically motivational. However, it is not immediately clear, nor is it a fundamental assumption of either Humeanism or Anti-Humeanism, that love and hate are desires. I do not wish to either endorse or impugn the claims that some emotions are intrinsically motivational, or that emotions are not desires. I merely wish to stress that neither claim is inconsistent with either Humeanism or Anti-Humeanism.<sup>51</sup> The end result is that both the Humean and Anti-Humean may reject the claim that only desires are intrinsically motivating. If this suggestion is right, then it is best to see Humeanism as being committed, not to the thesis that desires are the only attitudes that are intrinsically motivational, but rather to the claim that beliefs are not among those attitudes that are. The Anti-Humean, by contrast, wants to insist that at least some beliefs (perhaps, beliefs about the good or about what one should do) are included among the set of attitudes that are intrinsically motivational. But once

---

<sup>49</sup> Strictly speaking, it would be more accurate to say that *some* desires are motivating, since there may be some desires, such as the desire that it rain tomorrow, that are not intrinsically motivating. However, since we are only concerned with the desire to bring about **P**, we will not be concerned with such desires.

<sup>50</sup> Recall, the present usage of the word "desire" is meant to exclude "helpless desires", understood as desires that one is unable to bring about directly via the exercise of one's own agency.

<sup>51</sup> Hume himself typically speaks about the "passions", a label that includes, but is by no means limited to, desires. See A. J. Kenny [1963], esp. chapters 4 and 5.

it is acknowledge that the set of intrinsically motivational attitudes is not limited to desires, then (as far as Anti-Humeanism is concerned) there is no need to insist that desires are beliefs. In light of the preceding considerations, it should not be assumed that the rejection of the Desire-as-Belief Thesis is necessarily motivated by Humeanism or that the acceptance of the Desire-as-Belief Thesis is necessarily motivated by Anti-Humeanism.<sup>52</sup>

### 3.4. The Desire-as-Perception Thesis

Another common strategy for unpacking the Broadly Anscombean View is by appealing to the thesis that desires are a type of perception. According to this view, which we may refer to as the *Desire-as-Perception Thesis*, to desire to bring about **P** is to perceive that it is good to bring about **P**. One putative advantage of the Desire-as-Perception Thesis over the Desire-as-Belief Thesis is that it is able to preserve the intuition that one may have desires that one knows to be inconsistent without being guilty of irrationality. This is because, unlike beliefs, one may knowingly have inconsistent perceptual appearances without being liable to rational criticism. For example, an agent may visually perceive that a partially submerged stick, **X**, is bent and tactilely perceive that **X** is straight (where the agent in question recognises that these represent inconsistent perceptions) without being liable to rational criticism. We may explain this fact by registering that perceiving that **P** does not entail that one is rationally committed to the truth of **P**.<sup>53</sup> Hence, if desires are a type of perception, then it does not follow from the fact that one desires to bring about **P** that one is rationally committed to the goodness of bringing about **P**.

The above point should not be taken to suggest that the advocate of the Desire-as-Belief Thesis is left without any resources for explaining why having conflicting desires does not make one guilty of irrationality. For example, the defender of the Desire-as-Belief Thesis may claim that a desire is an instance of believing that it is good to bring about **P** from a certain perspective or relative to a certain set of considerations. Since, one may consistently believe that it is good to bring about **P** relative a perspective or set of considerations, **A**, and believe that it is not good to bring about **P** relative to another perspective or set of considerations, **B**, the defender of the Desire-as-Belief Thesis may claim that conflicting desires are not actually inconsistent (in a rationally significant sense) since they are always relative to a certain perspective or set of considerations. This argument may be used to explain why having conflicting desires does not make one liable to rational criticism. However, this line of reply saves one intuition (the intuition that one may have conflicting desires without being guilty of irrationality) by sacrificing another (the intuition that one may have genuinely inconsistent desires). The Desire-as-Perception Thesis, by contrast, is able to preserve both intuitions. Which of the two explanations one finds most satisfying will partly depend on whether or not one thinks the intuition that there may be genuinely inconsistent desires is worth preserving.

We may include, under the umbrella of the Desire-as-Perception Thesis, the account defended by Sergio Tenenbaum, who argues that desires are appearances of the good from a

---

<sup>52</sup> Price is not alone in associating the Desire-as-Belief Thesis with Humeanism. David Lewis [1988, 1996] does so as well. Given the pedigree of support for this association, it is important to emphasise that it is not inevitable.

<sup>53</sup> For an elaboration of this claim, see chapter 4, section 4.5, of this dissertation.

certain perspective.<sup>54</sup> However, it is important to register that the claim that desires are appearances (or perceptions) of the good from a certain perspective differs from the claim that desiring to bring about **P** entails the belief that it is good to bring about **P** from a certain perspective, since the latter is a claim about the content of the desire, while the former is a claim about the mode of presentation that a desire embodies. For example, if we say that my perception that a partially submerged stick is bent represents a certain perspective, we are not saying that the content of my visual experience is that ‘the stick is bent from a visual perspective’. The relevant content is simply that ‘the stick is bent’. And if the agent felt that the stick was straight with her hands, then the content of her tactile experience would be that ‘the stick is straight’, not that ‘the stick is straight from a tactile perspective’. In other words, it is not part of the content of the perceptual appearance that what is perceived is true from a certain perspective. Rather, the thing to say is that the perceptual appearance itself *instantiates* a particular perspective.

The end result is that when the advocate of the Desire-as-Perception Thesis says that desires are perspectival, she is saying something, not about the content of the desire, but about the mode under which that content is entertained. She is, in effect, registering that there are alternative perceptual modes under which that very same content may be entertained. Each perceptual mode is on par with the other in the sense that none of them involve rational commitments. By contrast, when the advocate of the Desire-as-Belief Thesis says that a desire is perspectival, she is saying something about the content of the desire. Moreover, the advocate of the Desire-as-Belief Thesis remains committed to the claim that desires involve rational commitments. Hence, the crucial difference between the Desire-as-Belief and Desire-as-Perception Theses is that the former (but not the latter) entails that desires involve rational commitments.

### 3.5. The Anscombe Interpretation

How are we to make sense of the claim that a desire is a type of belief or perception in light of the direction of fit analysis? I believe that, properly understood, the direction of fit analysis entails the falsity of the Desire-as-Belief and Desire-as-Perception Theses. However, upon close inspection, the two most widely discussed interpretations of the direction of fit metaphor—the interpretations due to Elizabeth Anscombe and Michael Smith—turn out to be consistent with either the Desire-as-Belief or Desire-as-Perception Thesis. I take this to be symptomatic of the fact that neither interpretation succeeds in getting to the heart of the matter as far as the direction of fit analysis is concerned.

---

<sup>54</sup> Strictly speaking, Tenenbaum's [1999] claim that desires are appearances of the good does not entail that desires are a type of perception. For example, Tenenbaum often draws an analogy between desires and non-perceptual appearances, such as when an invalid inference appears valid because the agent is in the grip of a fallacious thought process. For example, it may non-perceptually appear that the next coin-flip will land head-side-up to an agent who is in the grip of the gambler's fallacy and who has witnessed the coin land tail-side-up ten times in a row. It would therefore be more accurate to say that Tenenbaum is committed to a Desire-as-Appearance Thesis than a Desire-as-Perception Thesis. Nevertheless, the points I raise against the Desire-as-Perception Thesis apply equally to the Desire-as-Appearance Thesis.

As we noted in chapter 2, section 2.2, the most common way of unpacking the direction-of-fit metaphor—what I refer to as the *Anscombe interpretation*—is in terms of whether the attitude or the world is deemed to be at fault or subject to revision when there is a lack of correspondence between the two. On this view, what it means for a belief to have a mind-to-world fit is that the belief is deemed to be at fault (and therefore subject to revision) when there is a mismatch between the world and what is believed. By contrast, what it means for a desire to have a world-to-mind fit is that the world is deemed to be at fault (and therefore subject to revision) when there is a mismatch between what is desired and the world.

Contra initial appearances, the Anscombe interpretation of the direction of fit metaphor is perfectly consistent with both the Desire-as-Belief and Desire-as-Perception Theses. There can be little doubt that the Anscombe interpretation entails that the desire to bring about **P** is not identical to the belief that **P** (or the belief that one has, is, or will bring about **P**). For example, if I believe that I am having wine with dinner when I am in fact not having wine with dinner, then (according to the Anscombe interpretation) my belief is at fault; it is my belief that should be changed to suit the world, rather than the other way around. By contrast, if I desire that I have wine with dinner when I am in fact not having wine with dinner, then (according to the Anscombe interpretation) the problem lies with the world; it is the world that should be changed to suit my desire, rather than the other way around. The end result is that my desire to have wine with dinner cannot be identical with my belief that I am having wine with dinner since the two states have different success and revision conditions.

However, what remains unestablished is that the Anscombe interpretation entails that the desire to bring about **P** is not an instance of believing that it is good to bring about **P**. Moreover, upon closer inspection, this turns out to be the salient question when assessing the consistency of the Anscombe interpretation and the Desire-as-Belief and Desire-as-Perception Theses. The first step to seeing that this is so is to register that the Desire-as-Belief Thesis does not entail that the desire to bring about **P** is an instance of believing that **P** (or believing that I have brought about **P**, am bringing about **P**, or will bring about **P**). For example, according to the advocate of the Desire-as-Belief Thesis, my desire to have a glass of wine with dinner is not an instance of believing that I am having (or will have) a glass of wine with dinner. Rather, it is an instance of believing that my having a glass of wine with dinner is good. The salient question, then, is not whether my desire to have wine with dinner is at fault or subject to revision when I do not have wine with dinner. It is whether my desire to have wine with dinner is at fault or subject to revision when my having a glass of wine with dinner is not good. Since the advocate of the Desire-as-Belief Thesis is committed to saying that the desire to have wine with dinner is an instance of believing that having wine with dinner is good, she is also committed to saying that the desire to have wine with dinner is at fault or subject to revision when it is not good to have wine with dinner. But this is certainly not a mark against the Desire-as-Belief Thesis. On the contrary, this is just the sort of result we would expect if the Broadly Anscombean View were true. After all, when we say that a desire aims at the good, we mean to say that a desire is somehow defective, gets things wrong, or is subject to revision, when the desired outcome is not good. What we do not want to say is that the world is at fault or subject to revision when an agent desires something that is not good. The end result is that, insofar as we are committed to the claim that desire aims at the good, the Anscombe

interpretation forces us to also hold that the desire to bring about **P** has the same direction of fit as the belief that it is good to bring about **P**. Consequently, the Anscombe interpretation is consistent with the claim that the desire to bring about **P** is an instance of believing that it is good to bring about **P**. The same line of argument, *mutatis mutandis*, may be used to argue that the Anscombe interpretation is consistent with the Desire-as-Perception Thesis.

The Anscombe interpretation lacks the resources to establish that the desire to bring about **P** is not an instance of believing or perceiving that it is good to bring about **P** because of its exclusive focus on when a belief or desire is deemed to be at fault or subject to revision. Insofar as the GG theorist is committed to the claim that the good is the object of desire, she is committed to saying that a desire is faulty or subject to revision when the desired outcome is not good. However, this makes the conditions under which the desire to bring about **P** is faulty or subject to revision the same as the conditions under which the belief that it is good to bring about **P** is faulty or subject to revision. Since differences in when a pair of attitudes are faulty or subject to revision are the only tools available to the Anscombean interpretation for saying that they have different directions of fit, it follows that any two attitudes with the same correctness or revision conditions have the same direction of fit.

### 3.6. The Smith Interpretation

An alternative interpretation of the direction of fit metaphor is that of Michael Smith, who suggests that we unpack the metaphor in terms of a dispositional account of desire and belief:

[A] dispositional conception of desires enables us to cash the metaphor characterising beliefs and desires in terms of their direction of fit . . . . [T]he difference between beliefs and desires in terms of direction of fit comes down to a difference between the counterfactual dependence of a belief and a desire that **P**, on a perception that not **P**: roughly, a belief that **P** is a state that tends to go out of existence in the presence of a perception that not **P**, whereas a desire that **P** is a state that tends to endure, disposing the subject in the state to bring it about that **P**. Thus, we may say, attributions of beliefs and desires require that different kinds of counterfactuals are true of the subject to whom they are attributed. We may say that this is what a difference in their direction of fit is.<sup>55</sup>

Smith points out that, given the above characterisation, it is impossible for the same attitude to have both directions of fit. After all, this would require that the attitude in question have the tendency to both go out of existence in presence of a perception that not **P** and endure in the presence of a perception that not **P**.<sup>56</sup> Consequently, Smith concludes that desires and beliefs must be distinct attitudes. Let us refer to Smith's unpacking of direction of fit metaphor as the *Smith interpretation*.

Unlike the Anscombe interpretation, the Smith interpretation clearly entails that the Desire-as-Belief Thesis is false. However, it is less clear that it entails the falsity of the Desire-as-Perception Thesis. Recall, according to the Smith interpretation, desires and beliefs have

---

<sup>55</sup> Smith [1987: 54].

<sup>56</sup> Smith [1987: 56].



different directions of fit since the latter tends to go out of existence in the presence of a perception that  $\neg P$ , while the former tends to persist in the presence of a perception that  $\neg P$ . However, on this score, a perceptual appearance seems to be more like a desire than like a belief. Consider once again the case in which I visually perceive that a partially submerged stick is bent and tactilely perceive that it is straight. Since my visual experience of the stick being bent tends to persist in the presence of my tactile perception of it being straight, it seems to follow from Smith's analysis that my visual experience has the same direction of fit as a desire rather than a belief. Since the Smith interpretation is perfectly consistent with the claim that desires and perceptual appearances have the same direction of fit, it gives us no reason to doubt that desires are types of perceptions.

The preceding observations suggest not only that the Smith interpretation is consistent with the Desire-as-Perception Thesis, but also that it is of questionable plausibility. In order to constitute a satisfactory interpretation of the direction of fit metaphor, a given proposal must not only preserve the intuition that desires and beliefs have different directions of fit, but also the intuition that believing, imagining and pretending all share the same direction of fit. However, by Smith's lights, imagining and pretending turn out to have a different direction of fit to believing, owing to the fact that imagining that  $P$  and pretending that  $P$  tends to persist in the presence of the perception that  $\neg P$  while believing that  $P$  tends to go out of existence in the presence of the perception that  $\neg P$ . The end result is that the Smith interpretation cannot be generalised in order to provide a comprehensive analysis of the notion of direction of fit.

Additionally, it is not clear that Smith is even able to offer us a satisfactory analysis of the difference between belief and desire. The Smith interpretation takes it to be a brute psychological fact about us that our desires tend to endure in the presence of a perception that  $\neg P$  and that our beliefs tend to go out of existence in the presence of a perception that  $\neg P$ . However, it is entirely conceivable that there may be exceptions to the above rule. For example, there may be a putative belief that does not tend to go out of existence in the presence of the perception that  $\neg P$ , simply because it is dogmatically held. (Consider, for example, the dogmatic beliefs of a religious fanatic.) Moreover, we may wish to regard such a putative belief as a genuine belief, despite its dogmatic nature. However, Smith's interpretation of the direction of fit metaphor forces us to regard such a putative belief as a desire. Thus, Smith's framework seems to define out of existence what many would regard as a real possibility; namely, a belief that is so dogmatically held that it is immune to contravening perceptual evidence. Moreover, if we were to encounter a religious fanatic who held such a "dogmatic belief", we would be left without grounds for criticising such an individual since (by Smith's lights) their dogmatic "belief" would not constitute a belief at all, but a desire. Thus, the Smith interpretation offers the rather surprising prescription for arriving at epistemically faultless beliefs; one merely has to hold such beliefs in such a dogmatic fashion that they become *de facto* desires. These are all rather unattractive consequences of the Smith interpretation, ones that cast doubt on its adequacy as an interpretation of the direction of fit metaphor.

### 3.7. Diagnosing the Direction of Fit Analysis

The take away from sections 3.4 and 3.5 is that the two most influential interpretations of the direction of fit metaphor are consistent with the Desire-as-Belief or Desire-as-Perception Thesis. However, the defender of the Desire-as-Belief or Desire-as-Perception Thesis should not see this as cause for celebration since I have also argued that both the Anscombe and Smith interpretations are inadequate. The first is inadequate because it is unable to preserve the distinction between saying that an attitude has a certain direction of fit and saying that an attitude has certain correctness conditions.<sup>57</sup> The second is inadequate because it is unable to preserve the intuition that beliefs share the same direction of fit as imaginings and pretendings, and accommodate the possibility of dogmatic beliefs. Below, I present the two-content interpretation as an alternative to the Anscombe and Smith interpretations, one that does not suffer from the aforementioned shortcomings. Moreover, I show that the two-content interpretation entails the falsity of both the Desire-as-Belief and Desire-as-Perception Theses. If my claim that the two-content interpretation is superior to the Anscombe and Smith interpretation is sound, then the defender of the Desire-as-Belief or Desire-as-Perception Thesis can take little comfort in the fact that these interpretations are consistent with their account.

In the remainder of this chapter, I will attempt to do three things. First, identify what I take to be the key intuition that the notion of direction of fit is meant to capture; namely, the intuition that there is a difference in the logical structure of belief-like attitudes, on the one hand, and desire-like attitudes, on the other. Second, present the two-content interpretation as a plausible way of making sense of this intuition. Third, show that the two-content interpretation entails the falsity of the Desire-as-Belief and Desire-as-Perception Theses.

The intuition that there is a difference in the logical structure of belief-like and desire-like attitudes is inchoately registered by the observation that the former, but not the latter, is truth-assessable. This comports with our quotidian conception of the relevant attitudes—to wit, we do not ordinarily conceive of attitudes like desiring, wishing, hoping and intending as truth-assessable. There are at least two things we could mean when we say that desires are not ordinarily seen as truth-assessable. We could mean that we ordinarily believe that desires are not truth-assessable. Call this the *strong reading*. On the strong reading, our quotidian conception actually takes a stand on whether or not a desire is truth-assessable; to wit, it denies that it is. On the other hand, we could mean that we do not ordinarily believe that a desire is truth-assessable, but neither do we believe that it is not truth-assessable. Call this the *weak reading*. According to the weak reading, our quotidian conception does not take a stand on whether or not a desire is truth-assessable; it is agnostic on the issue.<sup>58</sup>

Whether the strong or weak reading accurately mirrors our ordinary linguistic practice is a sociological question, one that can hardly be answered from the philosopher's armchair. However, as I understand it, the direction of fit analysis is not meant to serve as a description of our quotidian conception of beliefs, desires, etc. Rather, it is part of a robust theoretical account

---

<sup>57</sup> See chapter 2, section 2.2, for my argument.

<sup>58</sup> The strong reading suggests that any theory of desires that depicts such desires as truth-assessable is revisionary with respect to our quotidian intuitions. The weak reading suggests that any theory of desires that depicts such desires as truth-assessable is an addendum to our quotidian intuitions.

of the relevant attitudes. Consequently, there is no need for us to answer the sociological question of whether the weak or strong reading is right. What interests us is the corresponding theoretical question: should we conceive of intentions and desires as truth-assessable in the very same sense that beliefs and perceptions are? The direction of fit analysis is first and foremost a negative answer to this question.

Consider Anscombe's shopping list example. One of its central points is that, unlike the detective's record, the shopping list is not deemed false because it fails to match the items in the shopping cart. This is because the shopping list does not represent the shopping cart as having certain items. By contrast, the detective's record does represent the shopping cart as having certain items. As such, the detective's record may be described as representing matters falsely if it fails to match the items in the shopping cart. Moreover, Anscombe's point may be applied not only to shopping lists and detective's records, but also to attitudes. If we assume that an attitude is true just in case matters are as the attitude represents them to be, and false just in case matters are not as the attitude represents them to be, then we may conclude that an attitude is truth-assessable just in case it represents matters as being a certain way. On this much, Anscombe and I agree. However, where I believe Anscombe goes wrong is that she assumes that the fact that an attitude represents matters as being a certain way entails that it is incorrect or subject to revision if matters are not that way. What the existence of attitudes like imagining and pretending illustrates is that it is possible to represent matters as being a certain way without purporting to represent matters truly. For example, if I were to pretend that there were magic beans in my shopping cart, my pretence does represent matters as being a certain way—to wit, it represents matters as though there were magic beans in my shopping cart. However, my pretence is not incorrect or subject to revision if there are no magic beans in my shopping cart because it does not purport to represent matters truly. What Anscombe overlooks is that there is a gap between merely representing matters as being a certain way, and representing matters as being a certain way while purporting to represent matters truly.

Once we recognise this gap, it becomes clear that we can only preserve the intuition that supposing, imagining, and pretending have the same direction of fit as belief if we assume that the direction of fit metaphor tracks the contrast between attitudes that represent matters as being a certain way and attitudes that do not, rather than the contrast between attitudes that represent matters as being a certain way while purporting to represent truly, and attitudes that do not. This is because only the former dichotomy keeps all four attitudes on the same side of the direction of fit divide. Moreover, all that follows from saying that an attitude represents matters a certain way is that the attitude is false if things are not that way. What does not follow is that the attitude is incorrect or subject to revision if matters are not as the attitude represents them to be. Of course, there is a natural tendency to think that saying that an attitude is false is equivalent to saying that it is incorrect or subject to revision. However, as natural as it may be, this tendency should be resisted. Saying that an attitude is false is to make a formal claim about its logical structure. Saying that an attitude is incorrect or subject to revision is to make an evaluative claim about an attitude's normative import. In short, these are very two different sorts of claims.

My contention, then, is that the direction of fit analysis is a claim about an attitude's logical structure, rather than a claim about its normative import. The first half of the preceding

claim appears to be common ground between Anscombe and I since she would no doubt agree that beliefs are truth assessable and that desires are not. Where we seem to disagree is with respect the second half of the preceding claim—to wit, my insistence that the notion of direction of fit lacks a certain normative import. The normative import in question is that having to do with the correctness or revision conditions of an attitude. This is the lesson to be learned from the existence of attitudes like imagining and pretending. Saying that an attitude has a certain direction of fit is not yet to say what its correctness or revision conditions are. It is not yet to specify what I have been calling an attitude's normative import. Consequently, if we wish to arrive at an accurate characterisation of the notion of direction of fit, we must restrict ourselves to the logical properties of an attitude. Specifically, we must restrict ourselves to the question of whether or not the attitude is truth-assessable.

Anscombe violates this restriction when she assumes that the direction of fit analysis is meant to capture an attitude's normative import, and this violation explains why her account conflates an attitude's correctness conditions with its direction of fit.<sup>59</sup> Smith also violates this restriction when he assumes that the direction of fit analysis is meant to capture an attitude's dispositional properties, and this violation explains why he is unable to preserve the intuition that imagining and pretending have the same direction of fit as believing. Since two truth-assessable attitudes may involve very different dispositions (e.g., believing that **P** tends to go out of existence in the presence of a perception that  $\neg\mathbf{P}$  while pretending that **P** tends to persist in the presence of a perception that  $\neg\mathbf{P}$ , despite the fact that both are truth-assessable), the attempt to make sense of direction of fit via a dispositional analysis is ill conceived, if not moribund. Hence, when viewed from a sufficient level of abstraction, the Anscombe and Smith interpretations may be seen as suffering from the same flaw. Both take the direction of fit analysis to involve more than a claim about the logical structure of two broad classes of attitudes—namely, those that are truth-assessable and those that are not. If my proposed restriction is right, and the claim that believing and pretending share the same direction of fit is merely meant to capture the idea that they both have truth-assessable content, while the claim that believing and desiring have different directions of fit is merely meant to capture the idea that, unlike the former, the latter does not have truth-assessable content, then an interpretation of the direction of fit metaphor which focuses on the correctness conditions, revision conditions or dispositions implicated by an attitude runs the risk of being little more than a distraction from the central issue.

### 3.8. The Two Content Interpretation

The two-content interpretation represents an attempt to take seriously the idea that the notion of direction of fit is meant to capture a logical property of an attitude. As such, it unpacks the claim that beliefs and desires have different directions of fit by saying that the two attitudes belong to different Boolean domains. This proposal combines the negative claim that a desire is not truth assessable in the way that a belief is, and the positive claim that a desire has its own

---

<sup>59</sup> For discussion, see chapter 2, section 2.2.

two-valued logical structure, one that implicates doability conditions. In the present section, I will briefly examine both claims, beginning with the negative claim.

According to the negative claim, attitudes with a desire-like direction of fit (e.g., desires, wishes, intentions) are not truth-assessable. It is important to note that saying that desires are not truth-assessable does not entail that they have no truth-assessable features or parts. To assume that it does would be to commit a fallacy of composition. Hence, the claim that a desire is not truth assessable is consistent with the claim that the propositional object of a desire is truth-assessable. The direction of fit analysis is therefore consistent with the claim that desires are propositional attitudes, where a proposition is conceived of as a truth-assessable entity. What the direction of fit analysis does entail is that a desire does not display the same two values displayed by its propositional content. Another way this point may be put is to say that the attitudes of desire is not a member of the same Boolean domain as its propositional content, if it is a member of a Boolean domain at all.

The possibility that an attitude may fail to display the same logical structure as its propositional content introduces a need to talk about an attitude in a way that is independent of its propositional content. This is one of the reasons why the notion of illocutionary content was introduced. Appealing to this notion allows us to buy into something like the following picture. We may say that both believing that **P** and desiring that **P** have the same propositional content. Moreover, like all propositions, **P** is truth-assessable. However, there is a difference between saying that the propositional content of an attitude is truth-assessable and saying that the attitude itself is truth-assessable. The latter is a claim about an attitude's illocutionary content. Thus, in the case of belief, both the attitude's illocutionary content and propositional content are truth assessable. As such, the illocutionary content and propositional content of a belief are members of the same Boolean domain. However, in the case of a desire, the illocutionary content is not a member of the same Boolean domain as its propositional content since the illocutionary content of a desire is not truth-assessable. This means that if desires have a two-valued logical structure at all, then the values are not truth and falsity.

There is still disagreement among those who buy into the direction of fit analysis with regards to the question of whether desires have a two-valued logical structure, and (if so) what are the two relevant values. These are the issues taken up by the positive claim. The two-content interpretation goes beyond the negative claim by not only positing that desires have a two-valued structure (analogous to but different from truth-values), but also by identifying the values in question with the "doable" and "undoable". It is at this stage that the notion of imperative content takes on particular significance. Simply put, a desire is not truth assessable because it has imperative content. Moreover, the diagnostic test limned in sections 2.6 and 2.7 allows us to determine that attitudes with imperative content have doability conditions. Thus, we can characterise the difference between indicative content and imperative content in terms of a difference in Boolean domains, with indicative content referring to illocutionary content belonging to the domain of truth-assessable items and imperative content referring to illocutionary content belonging to the domain of doability-assessable items.

Significantly, once we have taken on board the idea that what it means for an attitude to have mind-to-world fit is for it to have truth-assessable illocutionary content (i.e., indicative content), and what it means for an attitude to have world-to-mind fit is for it to have doability-

assessable illocutionary content (i.e., imperative content), there can be no question of whether a desire is a type of belief or perceptual appearance. Since a desire has a different kind of illocutionary content to a belief or perceptual appearance, the former cannot be an instance of the latter. Hence, the two-content interpretation entails the falsity of both the Desire-as-Belief and Desire-as-Perception Theses. Moreover, the two-content interpretation provides us with the resources to say that two attitudes have different directions of fit (in virtue of having different kinds of illocutionary content) even if they have the same correctness or revision conditions (as is alleged in the case of desire to bring about **P** and the belief that it is good to bring about **P**) and even if they involve the same dispositions (as in the case of a desire and a dogmatic belief). Hence, the two-content interpretation is immune to the problems plaguing the Anscombe and Smith interpretations.

### 3.9. Conclusion

In this chapter, I have attempted to further situate the account of desire on offer in this dissertation in the context of ongoing attempts to make sense of the claim that desires aim at the good. Specifically, I have argued that the two-content interpretation better captures the central intuition underlying the direction of fit analysis than the accounts of Anscombe and Smith. Moreover, I have argued that, in contradistinction to the Anscombe and Smith interpretations, the two-content interpretation entails the falsity of the Desire-as-Belief and Desire-as-Perception Theses. This does not only underscore the distinctiveness of the two-content interpretation (as compared to the Anscombe and Smith interpretations), but is also symptomatic of the fact that the Anscombe and Smith interpretations fail to get to the heart of the matter, as far as direction of fit analysis is concerned. I submit that an account of direction of fit, if it is adequate, will preclude the possibility of an attitude with mind-to-world fit being identical to or an instance of an attitude with world-to-mind fit. Of course, the immediately preceding claim does not entail that either the Desire-as-Belief or Desire-as-Perception Thesis is false. What it entails is that if the direction of fit analysis (properly understood) is true, then the Desire-as-Belief and Desire-as-Perception Theses are both false. In short, I am (in this chapter) making a claim about the incompatibility of two sets of claims, rather than a claim about one set of claims being true and another set of claims being false. Nonetheless, those who accept the direction of fit analysis, and who also buy into the two-content interpretation defended in this dissertation, may take the arguments advanced in this chapter as a basis for rejecting the Desire-as-Belief and Desire-as-Perception Theses.

## 4

## TWO KINDS OF RATIONAL SIGNIFICANCE

---

### 4.1. Introduction

In this chapter, I argue that there are two distinct kinds of rational significance that an attitude may display. This will turn out to have important implications for the question of whether or not a desire is rationally significant, since (if the arguments presented below are correct) it will mean that there are two things the cognitivist may mean when she says that desires are rationally significant. Ultimately, I will argue that we should be sceptical about desires being rationally significant in either of the two senses I describe in this chapter. But we will get to that in due course. At present, I simply wish to get clear on the two kinds of rational significance I have in mind, what they amount to, and why we should be interested in them.

In section 4.2, I define and motivate the notion of rational significance with which I am interested. Then, in section 4.3, I describe the two kinds of rational significance that an attitude may display under the heading of what I refer to as the *Dual-Significance Thesis*. In section 4.4, I consider an objection to the Dual-Significance Thesis—one directed at the idea that beliefs may be said to be rationally significant. I devote the rest of the chapter (sections 4.5–4.6) to responding to the objection limned in section 4.4, and (in the process of so doing) attempt to further clarify and motivate the Dual-Significance Thesis. If successful, the arguments in this chapter will (1) establish that there are two kinds of rational significance that an attitude may display, (2) go some distance towards clarifying the rational significance of belief, and (3) provide a framework (to be used in later chapters) for assessing the claim that desires are rationally significant.

### 4.2. Rational Significance

Let us say that an attitude is rationally significant just in case it plays a role (qua psychological state) in generating a rational obligation. In this dissertation, I will be primarily interested in rational obligations that are related to the adoption of a belief or intention. Specifically, I will be concerned with rational obligations to adopt a belief or intention, and rational obligations to refrain from adopting a belief or intention. Hence, I define rational significance as follows:

**Definition 4.1 (Rational Significance):**

X is rationally significant with respect to a belief or intention, Y, *IFF* X plays a role (qua psychological state) in generating a rational obligation to adopt Y or refrain from adopting Y.

Two points of clarification are worth making with respect to the preceding definition. First, the “qua psychological state” clause imposes a restriction similar to that limned in section 1.1 in the introduction of this dissertation. An attitude does not count as rationally significant, in the sense I presently have in mind, when it constitutes a reason *qua fact*. This is not to deny that an attitude may play a role in generating a rational obligation by constituting a reason, in the sense in which the fact that I am having a perceptual appearance of an elephant in the corner of the room constitutes a reason for me to check myself into a hospital. I affirm that, like any other fact, the fact that I have an attitude may, under the right circumstances, constitute a reason. However, the notion of rational significance I currently have in mind is one that only applies to an attitude when it generates a rational obligation in virtue of the fact that it is a psychological state—i.e., a state with the property of aboutness.

Second, I will be presupposing the following definition of a rational obligation:

**Definition 4.2 (Rational Obligation):**

S is rationally obligated to X *IFF* S would be liable to rational criticism if S failed to X.

For example, S is rationally obligated to believe **P** if and only if S would be liable to rational criticism for failing to believe **P**. Similarly, an agent, S, is rationally obligated to refrain from intending to bring about **P** if and only if S would be liable to rational criticism if S intended to bring about **P**. Furthermore, I will say that X is rationally permissible if and only if one is not rationally obligated to refrain from X. Thus understood, the notion of rational permissibility is parasitic on the notion of a rational obligation.

The present conception of rational significance yields something like the following picture: Suppose that (1) an agent is rationally obligated to believe **P** only if the agent has a mentally-grasped justification for believing **P**, and that (2) perceiving **P** provides one with a mentally-grasped justification for believing **P**. Given (1) and (2), it would follow, on the present picture, that perceiving **P** is rationally significant relative to the belief that **P**. Two further points of clarification are worth noting. First, saying that an attitude is rationally significant does not entail that it is sufficient to generate a rational obligation. The connection between rational significance and a rational obligation is contributory rather than decisive.<sup>60</sup> Second, saying that an attitude, X, plays a role in generating a rational obligation does not entail that some other attitude, Y, could not also play this role. For example, the present account allows that both perceiving **P** and remembering **P** may be rationally significant because they provide an agent with a mentally-grasped justification for believing **P**.

Definitions 4.1 and 4.2 are meant to be stipulative. As such, they are not meant to settle any pre-existing philosophical debates. That would constitute a misuse of a stipulation.

---

<sup>60</sup> For a discussion of this terminology, see Dancy [2004].



Stipulations, properly understood and employed, are meant to circumscribe a particular area of concern. The present stipulation is motivated by a concern with minimising the rational criticism for which one is liable. Motivations for having this concern range from the social (e.g., we may wish to gain the approval or avoid censure of other members of the community of reasoners to which we belong) to the epistemic (e.g., we believe that minimising the rational criticism to which we are liable happens to be the most effective way of arriving at knowledge or maximising true beliefs) and the practical (e.g., we believe that minimising the rational criticism to which we are liable happens to be the most effective way of arriving at appropriate or effective intentions). Whatever our motivation, we often find ourselves concerned with what we are rationally obligated to believe, intend, etc. It is this sort of concern that the present stipulation is meant to circumscribe. However, I am no more committed to saying that the present conception of rational significance is the only legitimate one than I am to saying that my particular concerns are the only ones worth having. Nevertheless, in order to further motivate my operationalization of the terms, it may be worthwhile to briefly sketch what I believe to be at stake with regards to the present conception of rational significance.

I begin with a few clarificatory remarks about what I mean by the expression “liability to rational criticism”. Being free from liability to rational criticism does not amount to freedom from error, tout court. For example, consider an agent, Tanya, who has her brain removed while asleep and placed into a vat of the kind epistemologists dread. Let us assume that the envatment process is carried out seamlessly, so that Tanya fails to realise that the world she now perceives is not the actual world, but rather a virtual reality made to look and feel like the actual world. Furthermore, suppose that Tanya goes about forming her beliefs in the most responsible manner possible. She is careful not to adopt any inconsistent beliefs, observes the appropriate standards of evidence, etc. Nonetheless, insofar as Tanya believes that the ‘tree’ she now sees outside her window is the same sort of thing as the actual trees she frequently saw prior to her envatment, she has a false belief. A fortiori, Tanya does not know that there is a tree outside her window. As such, she may be described as falling short of certain epistemic ideals. However, Tanya is not liable to rational criticism, at least not of the kind with which I am presently concerned. Insofar as Tanya went about forming her belief responsibly (e.g., by avoiding inconsistent beliefs, observing the appropriate standards of evidence, etc.) she still counts as having fulfilled all her rational obligations, in my sense of the term.

It bears repeating that the preceding analysis of rational obligations is not offered as an argument to the effect that we should conceive of rational obligations along the lines just described. Rather, it is merely an attempt to get clear on what I mean—and the sorts of considerations I will be concerned with—when I talk about rational obligations. The reason I am interested in this conception of a rational obligation is because I am concerned with the questions “What shall I believe?” and “What shall I do?” when posed from the perspective of the deliberating agent. This is a perspective that we have all occupied. For example, suppose that I am trying to decide whether I should take the local train, which is now pulling into the station, but which will make multiple stops before getting to my destination, or whether I should wait for an express train, which would get me to my destination without extra stops, but is currently running late. There is, of course, an objective fact of the matter as to which train will get me to my destination first, and if this is my sole concern, then this fact will determine

which train I should take. However, when I am engaged in deliberation, I am not privy to this fact. What I am privy to is my past experience, my knowledge of past local and express train frequency, etc. Within this context, making a decision that is not liable to rational criticism means weighing the reasons that are within my ken. It does not involve making a decision that simply gets things right, in a sense that ignores the epistemic limitations with which actual reasoners are typically saddled. It is in this sense that we are often in a situation analogous to that of Tanya. Not that we are typically mistaken in the manner Tanya is (or at least so we hope). But in the sense that, when we find ourselves in the situation of deliberating about what to believe or intend, we can only be expected to consider and be held responsible for what is within our ken, and what is in our ken may (wittingly or unwittingly) fall short of the fact of the matter. Moreover, even when we are saddled with the aforementioned limitations, we still find ourselves confronted with the questions “What shall I believe?” and “What shall I do?” This is the sort of predicament to which the present notions of rational significance and rational obligation speak.

### 4.3. The Dual-Significance Thesis

It is typically assumed that beliefs are rationally significant because they provide us with reasons to adopt other beliefs. However, John Broome calls attention to a potential problem with this assumption. According to Broome, believing **P** does not provide one with a reason to believe the logical consequences of **P**.<sup>61</sup> His argument begins with the observation that **P** logically entails **P**. If we assume that believing **P** provides one with a reason to believe the logical consequences of **P**, then it follows that the belief that **P** is a reason to believe **P**. This generates the rather unattractive conclusion that every belief is a reason for itself.<sup>62</sup> If we assume that one is rationally obligated to believe **P** only if one has a reason to believe **P**, then it seems to follow that believing **P** cannot generate a rational obligation to believe the logical consequences of **P**. The point of the preceding appropriation of Broome’s argument is not to deny that an agent who believes **P** is typically rationally obligated to believe the logical

---

<sup>61</sup> Broome [2000].

<sup>62</sup> Broome [2000: 85; 2002: 92-95]. For discussion, see: Brunero [2007: 426-427]. While I share Broome’s general misgivings about the claim that all beliefs are self-justifying, it should be noted that versions of this position have been embraced by a number of prominent philosophers, including Quine [1951], Sklar [1975], Chisholm [1980], and Harman [1986]. Some versions of epistemic conservatism, as this family of views is typically called, are more defensible than others, and a detailed examination of its many permutations will not be possible here. (For a helpful discussion of the various kinds of epistemic conservatism, see: Christensen [1994] and Vahid [2004]. See also: Fumerton [2007].) However, I believe this omission is tolerable since my main contention in this chapter—namely, that there are two kinds of rational significance that an attitude may display—is perfectly consistent with most versions of epistemic conservatism. While I am committed to saying that a belief displays a different kind of rational significance to perception, this does not prevent me from saying that a belief also displays the same kind of rational significance as perception. After all, it is possible that belief may display two different kinds of rational significance, only one of which it shares with perception. Hence, while I will be following Broome’s lead by assuming that belief’s cannot be self-justifying (and, therefore, cannot be said to provide reasons), this assumption need not be taken as a premise in the main argument of this chapter.

consequences of **P**. It is to cast doubt on the assumption that believing **P** plays a role in generating said obligation.

Even if one finds the preceding argument persuasive with respect to the conclusion that beliefs fail to provide reasons for other beliefs, one need not conclude that belief plays no role whatsoever in generating a rational obligation. This is because there are at least two distinct kinds of rational significance that an attitude may display: one having to do with the incursion of rational commitments and another having to do with providing rational support. Belief displays the first kind of rational significance. When I adopt the belief that **P**, I incur a rational commitment to the logical consequences of **P**. Incurring such a commitment means that, so long as I continue to believe **P**, I am rationally committed to everything that logically follows from **P**. This, of course, includes being rationally committed to **P** itself. However, saying that believing **P** entails being rationally committed to **P** is different from saying that the belief that **P** gives one a reason to believe **P**. While the latter amounts to the doctrine that all beliefs are self-justifying (i.e., believing **P** gives one a reason to believe **P**), the former is a truism about the beliefs of rational agents (i.e., believing **P** entails that one is rationally committed to **P**).<sup>63</sup>

Perceptual appearances display the second kind of rational significance. Perceiving **P** does not rationally commit me to **P**. For example, suppose that I perceive **P** but fail to adopt the belief that **P** because I deem my perceptual appearance unreliable. In such a case, it would be rationally permissible to adopt the belief that  $\neg$ **P**. Hence, perceiving **P** does not rationally commit one to **P**. Nevertheless, perceiving **P** favours the adoption of the belief that **P**. It does this by putting one in touch with the fact that **P**, the very fact that makes one's belief that **P** true, or by providing one with evidence for **P**. Thus, while perceiving **P** does not rationally commit one to **P**, it provides one with rational support for an attitude that rationally commits one to **P**.

Incurring rational commitments and providing rational support are species of the same genus, the genus of rational significance. Both kinds of rational significance are *pro tanto* in the sense that they are either *contributory* or *defeasible*. To say that a type of rational significance is contributory means that it is insufficient for generating a rational obligation on its own. A conjunction of contributory factors may be sufficient to generate a rational obligation. What makes each individual factor qualify as contributory is that it could not generate a rational obligation in isolation from other factors.<sup>64</sup> To say that a type of rational significance is defeasible means that it may be defeated, undermined or overturned. For example, the perceptual appearance of it raining may, in the absence of defeaters, be sufficient to generate a rational obligation to believe that it is raining. However, insofar as my perceptual appearance may be defeated (for example, if I were to learn that I imbibed a powerful hallucinogen that typically made it look like it was raining), it still counts as *pro tanto* because of its defeasibility.

---

<sup>63</sup> Another difference between saying that believing **P** gives one a reason to believe the logical consequences of **P** and saying that believing **P** rationally commits one to the logical consequences of **P** is that the latter, but not the former, entails that one is liable to rational criticism if one adopts the belief that  $\neg$ **P**. I am here, of course, assuming that the reason in question is *pro tanto* rather than *decisive*. One may have a *pro tanto* reason for believing **P** and yet rationally believe that  $\neg$ **P** if the weight of one's reasons favour believing  $\neg$ **P**. However, if one believes that **P**, and is therefore committed to **P**, it is not rationally permissible to adopt the belief that  $\neg$ **P**, so long as one continues to believe **P**.

<sup>64</sup> See and Cf. Dancy [2004: 17-29].

If the rational significance of an attitude (or set of attitudes) is sufficient to generate a rational obligation, I will say that it is *decisive*. However, we may distinguish between saying that the rational significance displayed by an attitude is always sufficient to generate a rational obligation, and saying that, on a particular occasion, it is sufficient to generate a rational obligation. The former is a claim about a certain type of rational significance, and the latter is a claim about a certain token of rational significance. I will refer to an attitude that displays a kind of rational significance that is always sufficient to generate a rational obligation as *type-decisive*. Saying that an attitude is type-decisive entails that the rational significance it displays is not pro tanto. By contrast, I will refer to an attitude that displays a kind of rational significance that is, on a particular occasion, sufficient to generate a rational obligation as *token-decisive*.<sup>65</sup> Saying that an attitude is token-decisive is consistent with saying that the rational significance it displays is pro tanto. Hence, when I say that commitment incursion and rational support are pro tanto, I am only denying that they are type-decisive. I have not ruled out the possibility that they are token-decisive.

While there are basic cases in which a single attitude, with a single kind of rational significance, is sufficient to generate a rational obligation (i.e., be token-decisive), there are also more complex cases in which several different attitudes (displaying both kinds of rational significance) are required to generate a rational obligation. For example, we need several different attitudes, and both kinds of rational significance, in order to explain how believing the premises of a modus ponens inference with empirical content makes one rationally obligated to believe the conclusion. For example, consider the following inference:

- (A1): The streets are wet.
- (A2): If the streets are wet, then it rained last night.
- (A3): It rained last night.

Believing (A1) and (A2) incurs a commitment to (A3). This means that so long as one believes (A1) and (A2), one is only rationally permitted to adopt an attitude of belief towards (A3), insofar as one adopts any attitude towards (A3) at all. There are three possible doxastic attitudes one may adopt towards (A3). One may believe (A3), one may disbelieve (A3), or one may withhold belief and disbelief with respect to (A3).<sup>66</sup> Saying that believing (A1) and (A2) incurs a rational commitment to (A3) means that of these three options, belief is the only attitude one is rationally permitted to adopt towards (A3), so long as one remains committed to (A1) and (A2).

---

<sup>65</sup> I am sceptical about the existence of type-decisive attitudes. I only use the notion of a type-decisive attitude as a foil, against which we may better understand what it means to say that an attitude is token-decisive.

<sup>66</sup> Significantly, I take withholding belief and disbelief towards **P** to be its own distinct attitude towards **P**. As such, it should not be confused with cases in which I fail to take any attitude towards **P** because I am still contemplating whether **P** or  $\neg$ **P** or because **P** simply has not occurred to me. For simplicity, I will henceforth refer to the attitude of withholding belief and disbelief with respect to **P** as simply withholding **P**.

Another way this point may be put is to say that one would be liable to rational criticism if one were to either disbelieve (A3) or withhold (A3), so long as one believes (A1) and (A2).

However, saying that believing **P** incurs a rational commitment to the logical consequences of **P** does not explain how believing (A1) and (A2) may make one rationally obligated to believe (A3). This is because there are always two ways to discharge the commitment incurred by a belief. One may discharge the commitment incurred by believing (A1) and (A2) by either adopting the belief that (A3) or by giving up one's commitment to (A1) or (A2). Moreover, the rational significance displayed by believing (A1) and (A2) is indifferent to these two ways of discharging one's rational commitment. As such, the kind of rational significance displayed by belief—i.e., rational commitment-incursion—leaves one's rational obligations underspecified.

It is at this point that the kind of rational significance displayed by perception becomes salient. Suppose that one's basis for believing (A1) is one's currently perceiving the street is wet. (Let us also assume, for the sake of argument, that the conditional depicted in (A2) is true and that one's belief in (A2) is well founded.) One way to make sense of the claim that perceiving the street is wet provides rational support for the belief that the street is wet is to say that one's perceptual appearance privileges one way of discharging one's rational commitment over another. To wit, visually perceiving the street is wet favours remaining committed to (A1), and adopting the belief that (A3) over giving up one's commitment to (A1) and thereby foregoing the need to believe (A3). Or, to put the same point more simply, perceiving **P** favours believing **P** over disbelieving **P** and withholding **P**.

The end result of the preceding analysis is that both kinds of rational significance collaborate to explain how believing the premises of a modus ponens inference with empirical content generates a rational obligation to believe the conclusion. If the premises of an argument imply the conclusion, then belief in the premises incurs a commitment to the conclusion. This means that one is rationally obligated to refrain from both believing the premises of the inference and disbelieving or withholding the conclusion. However, there are two ways of discharging this commitment: to either believe the conclusion or give up belief in one or more of the premises. If one has equal amounts of rational support for and against the truth of the premises (a possibility that includes cases in which one has no reasons for and no reasons against the truth of one's premises), then one is rationally obligated to withhold the premise(s). In such cases, believing the premise(s) fails to generate a rational obligation to believe the conclusion. If one has (net) reasons in favour of the falsity of one or more of the premises, then one's (net) reasons favour giving up the relevant premise(s) over believing the conclusion of the inference. In such cases, believing the premises again fails to generate a rational obligation to believe the conclusion. However, if one's (net) reasons speak in favour of the truth of the premise(s), then one's (net) reasons favour believing the conclusion of the inference over giving up the premise(s). In such cases, and only in such cases, believing the premise(s) generates a rational obligation to believe the conclusion. Hence, on the present picture, both kinds of rational significance—i.e., commitment-incursion and rational support—are necessary for explaining how believing (A1) and (A2) generates a rational obligation to believe (A3). I will refer to the account just limned—one that posits two distinct kinds of rational significance that an attitude may display—as the *Dual-Significance Thesis*.

#### 4.4. The Dispensability Objection

It is my contention that there are two distinct kinds of rational significance that an attitude may display. However, it may be argued that, given how I have set things up, it is unclear that believing **P** plays any role in generating the rational obligation to believe the logical consequences of **P** since, in principle, all the normative work may be done by perceiving **P**.<sup>67</sup> After all, by my own admission, perceiving **P** may, in the absence of defeaters, be sufficient to generate a rational obligation to believe **P**. Given that this is so, it is unclear that there is any need to invoke the commitment-incurring power of belief. For example, suppose that I have token-decisive rational support for believing the following propositions:

(A1): The streets are wet.

(A2): If the streets are wet, then it rained last night.

Saying that the rational support is token-decisive means that (on the occasion in question) it is sufficient to generate a rational obligation to believe (A1) and (A2). For example, it means that (on the occasion in question) there are no defeaters in the vicinity that could potentially undermine my perceptual evidence. But suppose that I have not adopted the corresponding beliefs. In short, I have failed to fulfil my rational obligation. Nevertheless, it still seems as though I am rationally obligated to adopt an attitude of belief towards the following proposition:

(A3): It rained last night.

In other words, given that I have token-decisive rational support for (A1) and (A2), it seems as though I would be liable to rational criticism for either disbelieving (A3) or withholding (A3). This seems to be true even in cases in which I have failed to fulfil my rational obligation to believe (A1) and (A2). And therein lays the potential problem for my view. For it seems as though I have incurred a commitment to believe the logical consequences of (A1) and (A2) simply in virtue of having token-decisive rational support for (A1) and (A2). There is therefore no work left for belief to do. When it comes to generating rational obligations, belief is dispensable. I will refer to the present objection to the Dual-Significance Thesis as the *Dispensability Objection*.

The remainder of this chapter will be devoted to responding to the Dispensability Objection. This will not only constitute a defence of the Dual-Significance Thesis, but will also highlight a number of important normative features of belief. A preliminary (but ultimately unsuccessful) response to the Dispensability Objection is to point out that a perceptual appearance can only provide rational support in the absence of defeaters. It may therefore be argued that a perceptual appearance does not incur a rational commitment on its own.

---

<sup>67</sup> I am here restricting myself to cases of belief based on a perceptual appearance. This is solely for the sake of simplicity. I do not believe that perceptual appearance is the only attitude that provides rational support. For example, memory does as well. Moreover, the same argument may be reformulated, *mutatis mutandis*, in terms of beliefs based on memory, testimony, intuition, etc., as the case might be.

However, this argument seems to offer little comfort since the very same point may be made with respect to belief. For example, believing (A1) and (A2) can only generate a rational obligation to believe (A3) if there are no defeaters in the vicinity. Consequently, there is no salient difference between perception and belief in this regard.

I wish to pursue a different line of argument. To this end, I will argue that it is a mistake to assume that one must have a belief in order for the kind of rational significance displayed by belief to generate a rational obligation. Contra this assumption (and in agreement with Broome), I hold that the kind of rational significance displayed by belief is wide in scope in the sense that an agent need not have a belief in order for the rational norms governing belief to have application to the agent.<sup>68</sup> Consider, for example, the norm that says that an agent is irrational if she believes **P** and also disbelieves or withholds **P**. A narrow scope formulation of this norm may be expressed as follows:

**Example 4.1 (Narrow Scope Norm):**

If one believes that **P**, then one is rationally obligated to refrain from disbelieving **P** and withholding **P**.

By contrast, a wide scope formulation of the aforementioned norm is as follows:

**Example 4.2 (Wide Scope Norm):**

One is rationally obligated to refrain from simultaneously believing **P** and disbelieving **P** or withholding **P**.

One difference between the narrow and wide scope formulations of the preceding norm is that the former is inconsistent with the observation that we are sometimes rationally obligated to give up a belief we already have and adopt an attitude of disbelief or withholding in its stead. For example, if I believe **P**, and then encounter token-decisive rational support for disbelieving **P**, I am rationally obligated to give up my belief in **P** and adopt an attitude of disbelief towards **P**. However, according to the narrow scope norm, if I believe that **P**, I am rationally obligated to refrain from disbelieving and withholding **P**. This rules out the possibility of revising the belief I already have, and adopting an attitude of disbelief or withholding towards **P** in its stead. As

---

<sup>68</sup> The question of whether or not rational norms (or what Broome calls “rational requirements”) are narrow or wide in scope remains a hotly contested issue. Broome typically frames the debate in terms of whether or not we have reasons to satisfy the requirements of rationality. However, at present, I am not interested in answering this question. I appeal to the narrow/wide scope distinction only to get a better handle on the difference between the kinds of rational significance displayed by belief and perception. Significantly, while I affirm that there are wide scope rational norms, I do not deny that there are also narrow scope rational norms. On the contrary, my own suspicion is that the norms governing perception (and other attitudes that provide rational support) are narrow in scope. Thus, I remain open to the possibility that while some rational requirements are wide in scope, others are narrow in scope. This puts me at odds with Broome, who not only affirms the existence of wide scope rational requirements, but also denies the existence of narrow scope rational requirements. I will not attempt to engage in a detailed examination of how my proposal fits into the narrow/wide scope debate here, since doing so will take us too far afield from the primary concern of this dissertation. For arguments to the effect that the requirements of rationality are narrow in scope, see Schroeder [2004, 2009] Kolodny [2005]. For contemporary defenses of the claim that the requirements of rationality are wide in scope, see Broome [1999, 2007], Dancy [2000], Wallace [2001], Brunero [2010] and Way [2010].

such, I take the narrow scope norm to be implausible. By contrast, the wide scope formulation is perfectly consistent with the observation that we are sometimes rationally obligated to revise a belief we already have in favour of disbelief or withholding. The only rational obligation specified by the wide scope formulation is that I ensure that I do not both believe **P** and disbelieve or withhold **P**. Fulfilling this norm allows that I may either abandon my belief that **P** or continue to believe that **P** and refrain from disbelieving and withholding **P**.

Another difference between the narrow and wide scope norms is that the former only applies to an agent who has a belief (and who has therefore fulfilled the antecedent of the conditional), while the latter applies to an agent irrespective of whether or not she has a belief. Hence, if we assume that the relevant norms governing belief are narrow in scope, an agent would have to actually have a belief in order for the belief to play a role in generating the agent's rational obligations. However, if we assume that the relevant norms are wide in scope, then they may generate a rational obligation even if the agent in question does not have a belief. It is this final point that the Dispensability Objection overlooks when it assumes that the fact that the agent in question does not have a belief entails that no normative properties of belief play a role in generating her rational obligation. Contra this assumption, I maintain that it is the fact that the norms governing belief are wide in scope that explains why the normative properties of belief can play a role in generating a rational obligation for an agent who does not have a belief.

However, if the rational norms governing belief are wide in scope (in the sense that I have suggested), this seems to introduce a puzzle. If a certain norm applies to an agent irrespective of whether or not they have a belief, then why should we conceive of that norm as a reflection of the rational significance of belief? The short answer to this question is that the relevant norm applies to belief and fails to apply to other attitudes that share the same propositional content as belief. The slightly longer answer to this question may be put as follows. Consider the norm that says that one is rationally obligated to refrain from believing **P** and also disbelieving or withholding **P**. (In the next section, I will offer an explanation of why this norm applies to belief.) It is important to register that a comparable norm fails to apply to perceiving **P**. It is false that one is rationally obligated to refrain from perceiving **P** and also disbelieving and withholding **P**. Since perceiving **P** and believing **P** both share the same propositional content, it means that the aforementioned norm does not simply track the logical relations in which one proposition stands to another. If it did, then it would apply equally to every attitude that shared the same propositional content. This means that the norm must be conceived of as applying to the attitude itself, rather than to its propositional content. It is this idea that I aim to capture when I say that the relevant norm reflects a rational significance of belief.

#### **4.5. Belief as a Reason-Receiving Attitude**

In this section, I will attempt to fill in some of the details omitted in my initial reply to the Dispensability Objection. My aim is to show that the kind of rational significance displayed by belief does something that the rational significance displayed by perception cannot do. Hence, contra the Dispensability Objection, there is still normative work left for belief to do.



I begin with the observation that a belief is the sort of attitude that one may have a justification for having, while a perceptual appearance is not. There is a clear and intuitive sense in which it is appropriate to talk about an agent's justification for believing **P** and in which it would be inappropriate to talk about her justification for perceiving **P**. Suppose I were to ask why you believe it is raining outside. If you responded by saying that you saw that it was raining outside, it would be inappropriate for me to then ask for your justification for seeing that it is raining outside.<sup>69</sup> Perceptual appearances are simply not the sort of attitudes one may have a justification for having. In short, beliefs are *reason-receiving attitudes* while perceptual appearances are not. I will refer to the present disanalogy between a belief and a perceptual appearance as the *justification disanalogy*.

The thesis that beliefs are reason-receiving attitudes has important implications for the normative nature of belief. Let us say that an agent has sufficient justification for taking **P** to be true just in case the total evidence an agent has available makes **P** more probable than  $\neg\mathbf{P}$ . I take the present notion of sufficient justification to be a minimalist requirement on well-formed belief. Presumably, we will eventually wish to invoke a higher standard for epistemically well-formed belief. However, the minimalist requirement will do for our present purposes since I am only concerned with articulating a necessary condition for reasonable belief. Given that believing **P** entails taking **P** to be true, I hold that one is only rationally permitted to believe **P** if one has sufficient justification for taking **P** to be true. Saying that one is rationally permitted to believe that **P** only if one has sufficient justification for taking **P** to be true means that one is liable to rational criticism if one lacks sufficient justification for taking **P** to be true. Thus, we arrive at the first implication of the claim that beliefs are reason-receiving for the normative nature of belief:

**(B1):** The fact that a belief is a reason-receiving attitude is sufficient to explain why believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be true.

The fact that a belief is a reason-receiving attitude also explains why having inconsistent beliefs makes one liable to rational criticism. If believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be true, then believing  $\neg\mathbf{P}$  makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be false. Thus, believing both **P** and  $\neg\mathbf{P}$  makes one liable to rational criticism if one does not simultaneously have sufficient justification for taking **P** to be true and false. However, it is not possible to simultaneously have sufficient justification for taking **P** to be true and false. Given the minimalist conception of sufficient justification currently at play, having sufficient justification for taking **P** to be true precludes the possibility of having sufficient justification for taking **P** to be false, and vice versa. The end result is that if one believes both **P** and  $\neg\mathbf{P}$ , one must be liable to some kind of rational criticism (whether for believing **P** without sufficient justification, believing  $\neg\mathbf{P}$  without sufficient

---

<sup>69</sup> On this score, it is important to distinguish between asking for someone's justification for thinking that their perceptual appearance is reliable, and their justification for having a certain perceptual appearance. While the former constitutes a legitimate request for reasons, the latter does not.

justification, or both). Thus, we arrive at my second hypothesis about the fact that a belief is a reason-receiving attitude:

**(B2):** The fact that a belief is a reason-receiving attitude is sufficient to explain why believing both **P** and  $\neg\mathbf{P}$  makes one liable to rational criticism.

The above point may be expanded to include not only inconsistent beliefs, but also inconsistent doxastic attitudes in general. On the present analysis, believing **P** and disbelieving **P** makes one liable to rational criticism because believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be true while disbelieving **P** (which is equivalent to believing  $\neg\mathbf{P}$ ) makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be false. This means that there is no way to simultaneously have sufficient justification for both. It therefore follows that one must be liable to rational criticism. A similar argument may be put forward to show that simultaneously believing **P** and withholding **P** makes one liable to rational criticism. This is because believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be true, while withholding **P** makes one liable to rational criticism if one has sufficient justification for taking **P** to be either true or false. Since withholding **P** amounts to an acknowledgement that one lacks sufficient justification for taking **P** to be true, and since believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be true, it follows that one is liable to rational criticism for both believing **P** and withholding **P**. This brings us to my third hypothesis about the fact that a belief is a reason-receiving attitude:

**(B3):** The fact that a belief is a reason-receiving attitude is sufficient to explain why simultaneously believing **P** and disbelieving **P** or simultaneously believing **P** and withholding **P** makes one liable to rational criticism.

It is worth noting that (B3) corresponds with the wide scope norm described in Example 4.2. We can now see that this norm is tied to the fact that a belief is a reason-receiving attitude.

Next, we turn our attention to the logical consequences of a belief. If believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be true, and **P** is true only if **Q** is true, then believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **Q** to be true. This explains why one is liable to rational criticism if one either disbelieves or withholds the logical consequences of one's belief. Given that believing **P** makes one liable to rational criticism if one lacks sufficient justification for taking **Q** to be true (where **Q** must be true in order for **P** to be true), it follows that one would be liable to rational criticism if one were either to disbelieve or withhold **Q** (which we have already established by our argument for (B3)). Thus, we arrive at my fourth hypothesis about the fact that a belief is a reason-receiving attitude:

**(B4):** The fact that a belief is a reason-receiving attitude is sufficient to explain why one is liable to rational criticism for disbelieving or withholding the logical consequences of one's belief.

I believe that (B1)-(B4) displays four important respects in which believing **P** differs from perceiving **P**. Two points are worth noting on this score. First, (B1)-(B4) suggests that the fact that a belief is a reason-receiving attitude explains a number of important normative features of belief. Specifically, the fact that a belief is a reason-receiving attitude explains why one is liable to rational criticism for having inconsistent beliefs, inconsistent doxastic attitudes, and for disbelieving or withholding the logical consequences of one's beliefs. Since a belief is a reason-receiving attitude independently of whether or not it is justified, and (by extension) independently of whether or not there is a defeater present, it follows from (B1)-(B4) that a belief displays these normative features whether or not they are justified and whether or not there is a defeater present. Second, since perceiving **P** is not a reason-receiving attitude, it follows that we cannot substitute perception for belief in (B1)-(B4). This means that even if perception displayed the various normative properties that constitute the explananda of (B1)-(B4), the explanans would be different.

Of course, it does not follow from the fact that perceiving **P** is not a reason-receiving attitude that perceiving **P** does not display the same normative properties that believing **P** displays in virtue of being a reason-receiving attitude. After all, perceiving **P** may display these very normative properties in virtue of some feature other than being a reason-receiving attitude. However, I believe that there are independent grounds for thinking that it does not. Let us consider each of the aforementioned normative properties in turn:

First, it is simply not plausible that perceiving **P** makes one liable to rational criticism if one lacks sufficient justification for taking **P** to be true. After all, that would entail that one would be liable to rational criticism whenever there was a defeater for one's perceptual evidence. Granted, one would be liable to rational criticism if one adopted a belief based on defeated perceptual evidence. But one is not liable to rational criticism for simply having defeated perceptual evidence.

Second, one is not liable to rational criticism for having a perceptual appearance that is inconsistent with what one believes. For example, if the two lines of a Müller-Lyer illusion perceptually appeared to be different in length even though (having measured them with a ruler) one believed them to be equal in length, one would not be liable to rational criticism for having a perceptual appearance that was inconsistent with what one believed. A similar line of argument may be given to show that one is not liable to rational criticism for having a perceptual appearance that is inconsistent with what one disbelieves or withholds.

Third, one is not liable to rational criticism for disbelieving or withholding the logical consequences of one's perceptual appearance. For example, suppose that line A perceptually appears to be longer than line B. Suppose further that I know that line C is equal in length to line B. It would logically follow from the truth of my perceptual appearance that line A is longer than line C. However, if I know that lines A and B are part of a Müller-Lyer illusion, I may disbelieve that line A is longer than line C. Moreover, I would not be liable to rational criticism for disbelieving that line A is longer than line C, even though it is logically entailed by my perceiving that line A is longer than line B. In light of the preceding considerations, we may not only conclude that perceiving **P** is not a reason-receiving attitude, but also that perceiving **P**

lacks the normative properties that believing **P** has in virtue of being a reason-receiving attitude.

#### 4.6. Implications of the Dual-Significance Thesis

Three implications of the analysis of belief and desire offered in the previous section are worth noting. The first is that a belief may be said to incur a rational commitment irrespective of the presence or absence of a defeater. This seems to give rise to the following question: If the presence of a defeater is sufficient to prevent believing **P** from generating a rational obligation to believe the logical consequences of **P**, then does it not follow that it is also sufficient to prevent believing **P** from incurring a rational commitment to the logical consequences of **P**? The answer is no. This is because incurring a rational commitment to the logical consequences of **P** does not entail that one is rationally obligated to believe the logical consequences of **P**. Recall, saying that believing **P** incurs a rational commitment to **Q** entails that one is rationally obligated to refrain from simultaneously believing **P** and disbelieving or withholding **Q**. This is consistent with being rationally obligated to give up one's belief in both **P** and **Q**. Hence, saying that an agent has incurred a rational commitment to the logical consequences of **P** does not entail that she is rationally obligated to believe the logical consequences of **P**.

A second noteworthy implication of the analysis of belief and perception offered in the previous section is that an agent may be guilty of multiple instances of irrationality relative to a single inference. Consider, for example, the modus ponens inference introduced in section 4.3 (reproduced below):

- (A1): The streets are wet.
- (A2): If the streets are wet, then it rained last night.
- (A3): It rained last night.

According to the view currently on offer, an agent, Aasif, would be liable to rational criticism for believing (A1) and (A2), and disbelieving (A3), even if his belief in (A1) were unjustified. In such a case, Aasif would be guilty of two separate instances of irrationality: the irrationality of believing something he had insufficient justification for believing, and the irrationality of having inconsistent doxastic attitudes. The same point holds, *mutatis mutandis*, if Aasif believes (A1) and (A2), based on insufficient evidence, and withholds (A3).

The preceding argument highlights an important difference between saying that believing **P** incurs a rational commitment to the logical consequences of **P** irrespective of the presence or absence of a defeater (i.e., the view I accept), and saying that believing **P** incurs a rational commitment to the logical consequences of **P** only if there are no defeaters present (i.e., the view I reject). According to the view I reject, Aasif is guilty of a single instance of irrationality if he believes (A1) and (A2), and disbelieves or withholds (A3). The fact that Aasif's belief in (A1) is unjustified means that he is rationally permitted to adopt a doxastic attitude that is inconsistent with believing (A3), such as disbelieving or withholding (A3). Consequently, Aasif is not liable to rational criticism for failing to disbelieve or withhold the

logical consequences of his unjustified beliefs. The only irrationality Aasif is guilty of is the irrationality of having a belief based on insufficient (i.e., defeated) perceptual evidence.

However, according to the view I accept, Aasif is guilty of two separate instances of irrationality. This feature of the view I accept may generate the following worry: if Aasif is guilty of two instances of irrationality, then why think that he has a rational obligation to give up his belief in (A1) rather than adopt an attitude of belief towards (A3)? After all, aren't both rational obligations on a par? Again, the answer is no. If Aasif is guilty of two instances of irrationality, then he is rationally obligated to extricate himself from both. However, the only way to do this would be to give up his belief in (A1). If Aasif were to simply adopt an attitude of belief towards (A3), he would still be left with the irrationality of believing (A1) based on insufficient evidence. However, if Aasif were to give up his belief in (A1), then he would both forego the need to believe (A3), and he would no longer be liable to rational criticism for having a belief based on insufficient evidence. This means that the only way for Aasif to avoid rational criticism altogether would be to give up his belief in (A1). Hence, the account presently on offer is able to preserve the intuition that Aasif is rationally obligated to give up his belief that (A1) rather than adopt the belief that (A3).

This brings us to a third important implication of the analysis of belief offered in the previous section: The fact that believing **P** incurs a rational commitment to the logical consequences of **P** is a necessary but insufficient condition for believing **P** to give rise to a rational obligation to believe the logical consequences of **P**. There is therefore a sense in which the commitment incurred by believing **P** is *pro tanto vis-à-vis* an agent's rational obligations. However, it is not *pro tanto* in the sense that believing **P** sometimes fails to incur a rational commitment to the logical consequences of **P**. Rather, it is *pro tanto* in the sense that incurring a commitment to the logical consequences of **P** does not always generate a rational obligation to believe the logical consequences of **P**. This comports with my insistence (in section 4.3) that the kind of rational significance displayed by belief is not type-decisive.

Finally, it may be argued that I have still failed to address one of the key intuitions motivating the Dispensibility Objection: namely, the intuition that we can determine what an agent's rational obligations are by simply looking at their (perceptual) evidence. For example, consider the case of an agent who perceives that it is raining outside. Suppose further that the rational support provided by the agent's perceptual appearance of it raining outside is token-decisive. Given this assumption, we do not need to inquire into whether or not the agent believes that it is raining outside in order to know what her rational obligations are. Insofar as she has token-decisive rational support of the proposition 'it is raining outside', we are already in a position to say that she has a rational obligation to refrain from disbelieving and withholding the proposition 'it is raining outside' along with any proposition that logically follows from the proposition 'it is raining outside'. It may therefore be argued that talk of "rational commitment incursion" is unnecessary when it comes to determining our rational obligations.<sup>70</sup>

Where I believe the above argument goes wrong is that it conflates an epistemic claim with an explanatory one. The Dual-Significance Thesis does not claim that, in cases in which a

---

<sup>70</sup> Thanks to an anonymous referee for alerting me to this particular objection.

belief is rationally significant, the belief plays a role in our *knowing* what our rational obligations are. Rather, it claims that the belief plays a role in our *having* a rational obligation. The present distinction may be illustrated by the following example. Suppose that I am able to tell that a certain physical reaction has taken place by observing tracks in a bubble chamber. It would be absurd to conclude from the fact that I did not have to observe the atomic particles causing the tracks in order to know that the physical reaction took place that the atomic particles did not play a role in generating the physical reaction. In short, there is a difference between saying that an atomic particle is necessary for a certain physical reaction to take place, and saying that observing the atomic particle is necessary for knowing that a certain physical reaction has taken place. Analogously, I may grant that observing the rational support an agent has available is enough for us to know what her rational obligations are. However, this does not settle the question of whether or not belief plays a role in an agent having the rational obligations that she has. Moreover, it is this question that the arguments presented in this chapter have sought to address by casting belief as an attitude governed by wide scope norms. What the preceding objection overlooks, and what the last two sections attempt to show, is that the only reason someone's rational support may play the epistemic role of revealing an agent's rational obligations is because there is a need for rational support, and the only reason there is a need for rational support (as a type of rational significance) is because of the existence of reason-receiving attitudes, like belief. If there were no reason-receiving attitudes like belief, there would be nothing for which to have reasons, no work for perceptual appearances to do, and therefore no rational obligations for an agent's rational support to reveal.

#### 4.7. Conclusion

If the arguments presented in this chapter are right, we are left with the following picture: In order for *S*'s belief that **P** to generate a rational obligation to believe **Q** (where **Q** is a logical consequence of **P**), the following two conditions must be satisfied: (1) *S*'s belief that **P** must incur a rational commitment to **Q**, and (2) *S* must have sufficient justification for believing **P**. Condition (1) is satisfied by the belief itself, owing to the fact that it is a reason-receiving attitude. Borrowing the notion of a wide scope requirement from Broome, I have argued that the rational commitment incurred by a belief is best conceived of in terms of the following wide scope norm: one is rationally obligated to refrain from simultaneously believing **P** and disbelieving or withholding the logical consequences of **P**. Saying that it is the belief that incurs the rational commitment is a way to register that we are making a claim about a certain attitude, and not merely a claim about the logical relations that obtain between certain propositions. Unlike condition (1), I hold that condition (2) cannot be satisfied by the belief itself. On this point Broome and I also agree. I believe that this is why the type of rational significance displayed by perception is necessary. However, even if one holds that beliefs are self-justifying (a la epistemic conservatism), one may consistently hold that a perceptual appearance may also fulfil condition (2). Hence, we may conclude that the kind of rational significance displayed by perceptual appearances satisfies condition (2). The end result is that there are two distinct kinds of rational significance that conspire to explain how believing **P** generates a rational obligation to believe the logical consequences of **P**.

## 5

## COGNITIVISM AND THE AIM OF DESIRE

---

### 5.1. Introduction

According to the cognitivist, the desire to bring about **P** provides rational support for bringing about **P** because the desire to bring about **P** represents bringing about **P** as good. The standard strategy for resisting the claim that desires provide rational support is to argue that the desire to bring about **P** does not represent bringing about **P** as good. In this chapter, I present Velleman's objections to cognitivism as an example of the standard strategy for arguing against the claim that desires provide rational support. I argue that Velleman's objections ultimately fail. I then attempt to fill in the lacuna left by the failure of Velleman's objections by raising a worry of my own. However, unlike Velleman, who attempts to meet the cognitivist head on, I attempt to show that the cognitivist claims remain unmotivated by offering an alternative picture of practical rationality. The present chapter concludes (in section 6) by sketching this alternative picture: one that removes the need to say that desires provide rational support. In the next and final chapter, I attempt to fill out the picture further, using a number of putative objections to my proposal as a foil for highlighting the more distinctive features of my view.

### 5.2. The Objection to Simple Cognitivism

Velleman discusses two types of cognitivism: *simple* and *sophisticated*. According to simple cognitivism, desires provide rational support because the object of a desire is an evaluative proposition. Velleman attributes this approach to Donald Davidson.<sup>71</sup> On this view, the object of, for example, the desire to have a glass of water is an evaluative proposition, such as 'having a glass of water is good'. Insofar as an evaluation of having a glass of water as good speaks in favour of having a glass of water, simple cognitivism entails that the desire to have a glass of water provides one with a reason to have a glass of water.

Velleman rejects simple cognitivism because it makes the possession of an evaluative concept a necessary condition for having a desire, and therefore precludes the possibility that infants and animals—who lack the relevant evaluative concepts—have desires. He puts the point as follows:

---

<sup>71</sup> See: Davidson [1978: 86].

If the cognitivist seriously means to characterize desire as an attitude toward an evaluative proposition, then he implies that the capacity to desire requires the possession of evaluative concepts. Yet a young child can want things long before it has acquired the concept of their being worth wanting or desirable.<sup>72</sup>

Insofar as the simple cognitivist is interested in offering a general theory of desire, Velleman's objection seems apt. However, it is not clear that simple cognitivism is most charitably conceived of in this way. Recall, simple cognitivism aims to explain how desires may provide rational support. As such, the simple cognitivist need not see herself as offering a general theory of desire. Instead, she may see herself as offering a theory of a subset of desires—namely, those that provide rational support. Thus, the simple cognitivist may consistently hold that the desires of infants and animals fail to provide them with rational support because they lack the relevant evaluative concept and also insist that the desires of an agent equipped with the relevant concepts may provide her with rational support. In fact, since there is agreement on both sides of the cognitivism debate that infants and animals are not the kinds of agents of whom talk of rational support is appropriate, this is precisely the sort of result one would expect.

Velleman anticipates something along the lines of the preceding response to his objection:

Of course, the young child may not be susceptible to rational guidance, either, but this point hardly counts in Davidson's favour. When Davidson characterises belief-desire motivation as equivalent to rational guidance, he leaves no room for agents who are moved by desires without being guided by reasons. The fact that children, who pursue desired ends, can nevertheless be too young for rational guidance is therefore a point against Davidson, on a par with my point that they can be too young for the concept of the desirable.<sup>73</sup>

The above passage suggests that Velleman takes Davidson to be committed to the claim that all desires constitute value judgements.<sup>74</sup> As such, the observation that we do not ordinarily expect infants to possess mentally-grasped justification is as much an objection to Davidson as Velleman's original criticism.<sup>75</sup> I will not attempt to settle the exegetical question of whether or not Velleman has accurately characterised Davidson's view here. Rather, I am interested in the more philosophically substantive question of whether or not the simple cognitivist must be saddled with such a view. I believe that the answer is no. To see why this is so, the following analogy from belief may be helpful. There are things we take to be true of the beliefs of rational agents that we do not take to be true of the beliefs of animal and infants. For example, the fact that a rational agent believes **P** entails that she is rationally committed to the logical

---

<sup>72</sup> Velleman [1999: 7].

<sup>73</sup> Velleman [1999: 22, note 12].

<sup>74</sup> A similar view is defended by Price [1989] and Humberstone [1987].

<sup>75</sup> See, for example: Davidson [1978: 102].



consequences of **P**. As we noted in chapter 4, this means that a rational agent is liable to rational criticism if she believes **P** and **Q**, and also believes that **P** entails  $\neg\mathbf{Q}$ . However, the same cannot be said of non-rational agents since part of what we mean when we say that an agent is non-rational is that the agent is not subject to rational appraisal. As such, if a non-rational agent believes **P**, it does not follow that she is rationally committed to the logical consequences of **P**. Consequently, there is something we take to be true of the beliefs of rational agents that we do not take to be true of the beliefs of non-rational agents.

The preceding disparity between the beliefs of rational and non-rational agents stands in need of explanation. One explanation of this disparity is that there is a difference in kind between the beliefs of rational agents and the beliefs of non-rational agents. On this view, our ordinary concept of a belief picks out two different metaphysical kinds—the beliefs of rational agents and the beliefs of non-rational agents—which are to be distinguished based on their different properties—e.g., the fact that the former generates rational commitments while the latter does not. Admittedly, there are theorists—most notably, Davidson—who insist that only the beliefs of rational agents are deserving of the title. Such theorists seem happy to take a revisionist approach to our ordinary concept of a belief by substituting a much more narrowly circumscribed theoretical concept in its stead. However, once one grants that there is a difference in kind between the beliefs of rational agents and those of non-rational agents, the question of whether or not the latter still deserves the label “belief” turns out to be largely terminological. It basically amounts to the question of whether or not we are willing to have the term belief applied to two different metaphysical kinds (thereby ensuring that the technical usage of the word “belief” is consistent with the quotidian usage) or only one (thereby ensuring a technical usage that is more restrictive than the quotidian).

I do not wish to either endorse or impugn the claim that there is a difference in kind between the beliefs of rational agents and the beliefs of non-rational agents, here. There may be other possible explanations of why there are things we take to be true of the beliefs of rational agents that we do not take to be true of the beliefs of non-rational agents, explanations that do not require that we say that the ordinary concept of belief picks out two distinct metaphysical kinds. For our present purposes, it is sufficient to note that the explanation which posits two distinct metaphysical kinds is included in the logical space of possibilities, is one that many theorists have taken seriously as a candidate explanation, and is one that remains a live option in contemporary philosophical debates. Moreover, an analogous strategy may be adopted by the simple cognitivist vis-à-vis desire. To this end, the simple cognitivist may hold that our ordinary concept of a desire picks out two distinct metaphysical kinds: the desires of agents equipped with some set of salient evaluative concepts and the desires of agents that lack said evaluative concepts. Moreover, the simple cognitivist may see herself as only offering an account of the former. On this view, it is no criticism of simple cognitivism to show that it fails to apply to the desires of agents that lack the relevant evaluative concepts since the theory was never intended to provide an account of the desires of such agents.

At first pass, the above argumentative strategy may seem ad hoc. However, depending on how we fill in the details of the view, the simple cognitivist may turn out to have a principled reason for the preceding restriction. For example, suppose that the simple cognitivist was committed to the view that only agents equipped with the relevant evaluative concepts are

liable to rational criticism for their actions and intentions. Suppose further that only agents that are liable to rational criticism for their intentions or actions are correctly regarded as rational, while agents that are not liable to rational criticism for their intentions or actions (e.g., animals and infants) are non-rational. Assuming that the simple cognitivist had reasons for holding the preceding views—both of which strike me as plausible—she seems to have a principled basis for distinguishing between the desires of rational agents and the desires of non-rational agents. Moreover, since simple cognitivism is being appealed to in order to explain how desires may provide rational support, and given that rational support is only relevant in the case of rational agents, then the desires of rational agents are the only desires that are relevant to the simple cognitivist's aims.<sup>76</sup>

Since I do not hold that desires provide rational support, it goes without saying that I do not find the preceding considerations to speak decisively in favour of the idea that desires do provide rational support. However, what I take the preceding considerations to show is that Velleman has not established that simple cognitivism is mistaken or moribund. There are still a number of philosophical/rhetorical moves open to the simple cognitivist that Velleman has failed to forestall. As such, there is still work for Velleman to do if his objections to simple cognitivism are to succeed.

### 5.3. The Objection to Sophisticated Cognitivism

Having impugned simple cognitivism (at least to his own satisfaction), Velleman turns his attention to sophisticated cognitivism. The sophisticated cognitivist holds that a desire provides rational support, not because its object is an evaluative proposition, but because of its direction of fit. Dennis Stampe espouses a version of sophisticated cognitivism when he writes:

While the belief and the desire that **P** have the same propositional content and represent the same state of affairs, there is a difference in the way it is represented in the two states of mind. In belief it is represented as obtaining, whereas in desire, it is represented as a state of affairs the obtaining of which would be good.<sup>77</sup>

By Stampe's lights, the desire to have a glass of water has the same propositional object as the belief that one will have a glass of water. Hence, the object of the desire to have a glass of water is not an evaluative proposition, such as 'having a glass of water is good'. What distinguishes

---

<sup>76</sup> The suggestion that the desires of an agent equipped with the relevant concepts may play a justificatory role that the desires of an infant cannot is not as strange or novel as it may initially seem. It is widely held that the perceptual appearances of an agent equipped with the appropriate concepts may provide her with justification that could not be had by someone who lacked the concepts in question. For example, an agent who is equipped with the appropriate concepts may come to justifiably believe that there is a fire truck nearby after hearing a blaring siren even though an infant or animal (that lacked the appropriate concepts) could not come to have the same justified belief under similar circumstances. If there were a theory of perception that aimed to explain how the perceptual appearances of an agent equipped with the appropriate concepts could provide her with justification for adopting certain beliefs, it would be no objection to such an account to argue that it failed to apply to agents who lacked the ability to conceive of reasons as such. A similar argument may be advanced on behalf of simple cognitivism.

<sup>77</sup> Stampe [1987: 355].

the desire to have a glass of water from the belief that I will have a glass of water is how the relevant proposition is represented by each attitude, with the latter representing the proposition 'I am having a glass of water' as true, and the former representing the proposition 'I am having a glass of water' as to be made true. The present characterisation of the difference between a desire and a belief is typically summarised by saying that the attitudes have different *directions of fit*. More generally, the belief that **P** represents **P** as true, and the desire to bring about **P** represents **P** as to be made true. Assuming that the fact that **P** is to be made true is a consideration that speaks in favour of making **P** true, then sophisticated cognitivism entails that desires may put us in touch with a reason and that desires therefore provide rational support. Since, according to sophisticated cognitivism, a desire aims at the good in virtue of its direction of fit, rather than in virtue of its propositional content, it does not require that the desiring agent possess a particular evaluative concept. Hence, it sidesteps the standard objection to simple cognitivism. Insofar as all desires share the same direction of fit, the sophisticated cognitivist offers a univocal account of the desires of rational and non-rational agents. Velleman takes this to be an attractive feature of sophisticated cognitivism, one that makes it more plausible than simple cognitivism.

However, Velleman maintains that while more plausible than simple cognitivism, sophisticated cognitivism has problems of its own. If the sophisticated cognitivist means to say that the good is the constitutive aim of desire, then she must be committed to saying that a desire aims at depicting the way the agent wants the world to be with the aim of getting it right that things should be that way. Velleman sums up this idea by saying that a desire must track the *facienda*. Only then is the sophisticated cognitivist entitled to say that the good determines a desire's correctness. However, Velleman insists that desire does not track the *facienda*. At least two separate lines of argument in support of this conclusion may be extracted from Velleman's discussion.

### 5.3.1 *The Constitutive Aim Argument*

The first line of argument takes issue with the sophisticated cognitivist's attempt to ground a desire's capacity to provide rational support in its direction of fit. Specifically, it attempts to establish that the direction of fit of an attitude is poorly suited for the purpose of determining its constitutive aim by depicting the latter as too fine-grained a concept for the former to capture. The label "constitutive aim" is typically applied to a feature of an attitude that distinguishes it from all other attitudes (including those with the same direction of fit).<sup>78</sup> The following analogy is illustrative: to say that having mammary glands is constitutive of being a mammal is to say that having mammary glands is both necessary and sufficient for a species to be a mammal. Having mammary glands distinguishes mammals from all other animal species. Being warm-blooded, by contrast, is not constitutive of being a mammal since other animals (such as birds) are also warm-blooded. Analogously, to say that truth is the constitutive aim of belief is to say that having the truth-aim is both necessary and sufficient for an attitude to be a

---

<sup>78</sup> While I am reluctant to attribute this view to Velleman himself, audience members at graduate philosophy conferences at Princeton University and the University of Waterloo, where earlier versions of this chapter were presented, have either found this view attractive or see Velleman as committed to such a view. Either way, I take the present objection to be one worth responding to, independently of whether or not Velleman actually endorses it.

belief. According to this view, aiming at truth distinguishes belief from all other attitudes, including those that share its direction of fit, such as imagining and fantasising.

Given the preceding conception of a constitutive aim, it quickly becomes clear that the good is not the constitutive aim of desire. Let us assume, for the sake of argument, that a desire aims at the good in virtue of its direction of fit. This is certainly not unique to desire. Ostensibly, a wish or hope seems to aim at the good no less than a desire does. For example, if I wish to bring about **P** or hope to bring about **P**, it seems as though my wish or hope would be deemed incorrect or defective (if it were not good to bring about **P**) to the very same extent as my desire to bring about **P**. There is no salient difference among the three attitudes in this regard. Thus, if one were to hold that a desire aims at the good because a desire is deemed incorrect or defective if the desired outcome is not good, then one should be equally committed to saying that wishes and hopes aim at the good. But this suggests that, in reference to our earlier analogy, the goodness aim is more like being warm-blooded than like possessing mammary glands; it is something a desire shares with other attitudes, rather than something that distinguishes a desire from other attitudes. As such, the good could not be the constitutive aim of desire.

### 5.3.2 *The Perverse Desire Argument*

The second line of argument attempts to exploit the concept of a perverse desire to show that desires do not track the *facienda*. Velleman defines a perverse desire as a desire for some outcome under the very description in which one recognises that outcome to be bad, pointless, or unworthy of being desired. Velleman does not elaborate on what it means to say that an agent desires an outcome under the very description in which she recognises it to be bad, but he seems to be borrowing the idea from Anscombe, who notes that the same act or set of bodily movements may be conceived of under different descriptions. For example, consider the case of an agent, Smith, who is being paid \$50 to pump water to a village using a hand crank. Suppose Smith comes to learn that there is poison in the water, and that the act of pumping water would result in the poisoning the village. There are two different descriptions under which Smith may conceive of the act of pumping water.

- (A): Pumping water in order to make \$50 dollars.
- (B): Pumping water in order to poison the village.

Moreover, it is possible that Smith may regard the act of pumping water as good under one description, (A), and as bad under the other, (B). Moreover, it is possible that Smith may desire to pump water under description (A), but fail to desire to pump water under description (B)—e.g., if Smith wanted to make \$50 dollars but did not want to poison the village. However, such a case would fail to constitute a perverse desire in Velleman's sense. Insofar as Smith believes pumping water to be good under description (A), he does not count as having a perverse desire if he desires to pump water under description (A). However, we can imagine an alternative case, one in which Smith despises the villagers and therefore desires to poison them, despite the fact that he believes it would not be good to do so. In this alternative scenario, Smith desires to pump water under description (B), the very description under which he believes pumping

water to be bad. It is in this second scenario that Smith may be said to have a perverse desire, in Velleman's sense of the term.

According to Velleman, if a desire were really analogous to a belief about what is good or worth desiring then it should be just as incoherent to desire an outcome under the very description in which one recognises it to be bad or unworthy of being desired as it would be to believe that a proposition were true under the very description in which one recognises it to be false. Velleman insists that the incoherence of believing something under the very description in which one recognises it to be false is illustrated by the impossibility of believing the following proposition:

(C): I am five inches taller than I am.

One could not believe (C) because doing so would require that one take something to be true under the very description in which one recognises it to be false.

By contrast, the coherence of desiring something under the very description in which one recognises it to be bad or unworthy of being desired is illustrated by the example of Milton's Satan. Velleman puts the point as follows:

The assumption that desire aims at the good forces the cognitivist to misdescribe examples of perverse desire. Consider, for a particularly vivid example, the figure of Satan in *Paradise Lost*, who responds to his defeat with the cry "Evil be thou my Good." Satan is here resolving to desire and pursue evil, and hence—as he himself puts it—to regard evil as good. But he cannot reasonably be interpreted as adopting new estimates of what is valuable—that is, as resolving to cease judging evil to be evil and to start judging it to be good. If Satan ever loses sight of the evil in what he now desires, if he ever comes to think of what he desires as really good, he will no longer be at all satanic; he'll be just another well-intentioned fool. The ruler of Hell doesn't desire what he wrongly thinks is worthy of approval; he desires what he rightly thinks isn't.<sup>79</sup>

Velleman interprets Satan's cry "Evil be thou my Good" to be an expression of a desire to pursue evil under the very description in which it is recognised as such. If the cognitivist were correct to conceive of desire as aiming at the good in the same sense that belief aims at truth, then Satan's declaration would be analogous to believing something under the very description in which it is recognised to be false. Since this is impossible in the case of the belief, the fact that it is possible in the case of desire indicates that desire does not aim at the good in the same sense that belief aims at the truth. That is to say, desire does not depict the way the agent wants the world to be with the aim of getting it right that the world should be that way; desires do not track the *facienda*.

#### 5.4. Reply to the Constitutive Aim Argument

The constitutive aim argument ultimately fails as an objection to cognitivism because the cognitivist need not be committed to the claim that the good is the constitutive aim of desire. To

---

<sup>79</sup> Velleman [1992: 18].

this end, we may distinguish between two notions of the aim of a psychological attitude: the *constitutive aim* and the *normative aim*. Recall, according to Velleman, the constitutive aim of an attitude does not only specify the correctness-conditions of that attitude, but also distinguishes that attitude from all other attitudes. On this conception, to say that desire aims at the good entails both that (1) the desire to bring about **P** is correct just in case it is good to bring about **P**, and (2) only the desire to bring about **P** is correct just in case it is good to bring about **P**. By contrast, the notion of a normative aim simply specifies the correctness-conditions of an attitude. It does not tell us what distinguishes a given attitude from all other attitudes. To say that the good is the normative aim of desire entails that the desire to bring about **P** is correct just in case it is good to bring about **P**. However, it does not entail that only the desire to bring about **P** is correct just in case it is good to bring about **P**. Hence, the concept of a normative aim is weaker than that of a constitutive aim. While two or more attitudes cannot share the same constitutive aim, they can share the same normative aim.

According to the constitutive aim argument, the desire to bring about **P** cannot aim at the good because there are other attitudes, such as wishing to bring about **P** and hoping to bring about **P**, that are correct just in case it is good to bring about **P**. This argument relies on the implicit assumption that the cognitivist is committed to the claim that the good is the constitutive aim of desire. However, pace Velleman, I will argue that (i) when almost all cognitivists say that desire aims at the good, they are most plausibly understood as having the weaker notion of a normative aim in mind, (ii) the weaker notion of a normative aim is sufficient to play the theoretical role that the cognitivist needs, and (iii) there are independent reasons to prefer the notion of a normative aim to Velleman's conception of a constitutive aim. Consequently, Velleman's assumption that the cognitivist is committed to saying that the good is the constitutive aim of desire is unwarranted.

In support of (i), it is sufficient to note that almost all cognitivists hold that not only desires, but also intentions, aim at the good. By the lights of almost all cognitivists, the intention to bring about **P** is correct just in case it is good to bring about **P**. It follows that, by the lights of almost all cognitivists, the desire to bring about **P** is not the only attitude that is correct when it is good to bring about **P**. This suggests that when almost all cognitivists say that desire aims at the good, they have the notion of a normative aim, rather than the notion of a constitutive aim, in mind.

In support of (ii), it is sufficient to note that none of the claims that Velleman attributes to the cognitivist requires that she be committed to the claim that the good is the constitutive aim of desire. We may, following Velleman, grant that a desire can provide an agent with rational support only if it tracks the *facienda*. But saying that desires track the *facienda* does not require that we say that desires are the only attitudes that track the *facienda*; it does not entail that the desire to bring about **P** is the only attitude that is correct just in case it is good to bring about **P**. It only requires that the desire to bring about **P** be a member of the set of attitudes that are correct just in case it is good to bring about **P**. Hence, the cognitivist need not be committed to the claim that a desire is the only attitude that is correct just in case its object is good. If this is right, then the cognitivist need not be saddled with the position Velleman attributes to her: namely, the claim that the good is the constitutive aim of desire.

Some may object to my suggestion that the notion of a normative aim is sufficient to do all the work that the cognitivist needs on the grounds that the notion of a normative aim is an insufficient basis for individuating attitudes. I believe this charge gets things half right. Given that more than one attitude can have the same normative aim, it certainly follows that the notion of a normative aim cannot serve as a basis for individuating attitudes. That much seems right. But why should we expect the aim of an attitude to play such a role? Firstly, the cognitivist never claimed that her account of the aim of desire should be used as a basis for individuating attitudes. Secondly, there are other features of a desire, apart from its correctness conditions, that may potentially serve as an adequate basis for individuating a desire from other attitudes.<sup>80</sup> In fact, Velleman provides us with what he takes to be a perfectly adequate basis for individuating between desires and other attitudes, one that is completely independent of the fact that it is correct just in case the desired outcome is good. According to Velleman, what distinguishes desire from all other conative attitudes is that it is directed at a possible future outcome. Borrowing this idea from Velleman, the cognitivist may say that what distinguishes a desire from other conative attitudes is not its normative aim—i.e., the fact that it is correct just in case the desired state of affairs is true—but the fact that it is always directed at something believed to be a possible future outcome. Given that, by Velleman's own lights, there are other features of desire that can be used to individuate it from other attitudes, the cognitivist need not be committed to individuating attitudes based on their correctness-conditions.

In support of (iii), it is worth noting that the notion of a normative aim is, from a taxonomic point of view, more useful than the notion of a constitutive aim. This is because the notion of a constitutive aim (which features in the constitutive aim argument) conflates two very different properties of an attitude: that which determines the correctness conditions of an attitude and that which distinguishes a particular attitude from all other attitudes. Recall, according to the constitutive aim argument, truth is the aim of belief because (a) believing **P** is correct just in case **P** is true, and (b) believing **P** is the only attitude towards **P** that is correct just in case **P** is true. However, there are many theorists (who would describe themselves as being committed to the claim that belief aims at the truth) who are committed to saying that not only believing **P**, but also perceiving **P** and remembering **P** are correct just in case **P** is true. If we buy into Velleman's notion of a constitutive aim, we would be forced to classify all such theorists (despite their claims to the contrary) among the ranks of those who deny that belief aims at the truth.

This is a very unattractive consequence of the notion of a constitutive aim that features in the constitutive aim argument. It equates a certain positive claim about memory or

---

<sup>80</sup> I suspect that the assumption that attitudes should be individuated in terms of their correctness-conditions may seem tempting because, as philosophers, we are used to talking about propositions and it is common to individuate propositions in terms of their correctness-conditions. This may prime us to see correctness-conditions as an attractive basis for individuating various items. However, this temptation overlooks the many important differences between a proposition and an attitude. In contradistinction to propositions, attitudes display many non-normative features. After all, attitudes are also the subject matter of a number of different empirical sciences, whose interest in them is often far removed from their normative underpinnings. This should already hint to us that there is no need to make some normative feature of an attitude—namely, its correctness-conditions—the single basis upon which that attitude is to be distinguished from all others. In brief, there are several non-normative features of attitudes that are well positioned to serve as a basis for individuating one attitude from another. In this respect, attitudes differ from abstract items such as propositions.

perception—namely, the claim that remembering **P** or perceiving **P** is correct just in case **P** is true—with a certain negative claim about belief—namely, the claim that beliefs do not aim at truth. Whether or not one buys into the claim that remembering **P** or perceiving **P** is correct just in case **P** is true, it is not difficult to see that there is something awkward about saying that all those who do are ipso facto denying that belief aims at the truth. This is not a problem for the notion of a normative aim, which allows us to consistently affirm that belief aims at the truth and that believing **P**, remembering **P** and perceiving **P** are all correct just in case it is true that **P**. The end result of (i)-(iii) is that the assumption that the cognitivist must be committed to the claim that the good is the constitutive aim of desire turns out to be unwarranted.

### 5.5. The Inconsistency Test

The remainder of this chapter will be devoted to showing that the Perverse Desire Argument fails as an objection to cognitivism. I believe that the problem lies in the fact that the argument assumes that the sophisticated cognitivist is committed to an analogy between desire and belief. This, I will argue, is a mistake. To recap, Velleman depicts the sophisticated cognitivist as being committed to the claim that desires aim at the good in the same sense that a belief aims at the truth. He then proceeds to point out that the sense in which a belief aims at the truth precludes the possibility of perverse belief. Desires, by contrast, do not aim at the good in a sense that precludes the possibility of perverse desires. Consequently, desires do not aim at the good in the same sense that belief aims at truth.

Contra Velleman, I claim that the sophisticated cognitivist is most charitably seen as being committed to an analogy not between desire and belief, but between desire and perception. It is at this point that the theoretical framework introduced in chapter 4 becomes salient. Recall, according to the Dual-Significance Thesis, there are two kinds of rational significance that an attitude may display: *commitment incursion* and *rational support*. The first is the kind displayed by belief, and the second is the kind displayed by perception. The question that now confronts us is this: which of the two kinds of rational significance is the cognitivist most plausibly seen as attributing to desire? The first step to answering this question is to note that it is not plausible that a desire is a commitment-incurring attitude.

One test for determining if an attitude is commitment-incurring exploits the observation that an agent with inconsistent rational commitments is liable to rational criticism. Call this the *inconsistency test*. In order to apply the inconsistency test, we juxtapose the attitude we are inquiring about with an attitudinal contrapositive of the same type, and then ask if an agent with both attitudes would be liable to rational criticism. I define an attitudinal contrapositive as follows:

**Definition 5.1 (Attitudinal Contrapositive):**

An attitude, *X*, is the attitudinal contrapositive of some attitude, *Y*, IFF the illocutionary content of *X* is equivalent to the negation of the illocutionary content of *Y*.

In the case of belief, the notion of an attitudinal contrapositive is relatively straightforward since the illocutionary content of a belief is the same as its propositional content. Thus, the attitudinal



contrapositives of believing **P** include believing  $\neg\mathbf{P}$  and disbelieving **P** (the latter being equivalent to believing  $\neg\mathbf{P}$ ). Other attitudinal contrapositives of believing **P** include imagining  $\neg\mathbf{P}$  and pretending that  $\neg\mathbf{P}$ . However, the attitudes of imagining and pretending are not the same attitude type as believing. Thus, they do not fulfil the requirement that the attitudinal contrapositive be of the same type as the attitude to which the inconsistency test is being applied. Only believing  $\neg\mathbf{P}$  fulfils this requirement. Hence, in order to apply the inconsistency test to belief, we juxtapose the belief that **P** with the belief that  $\neg\mathbf{P}$  (i.e., an attitudinal contrapositive of the same type), and then ask if a person with both attitudes would be liable to rational criticism. Since the answer is yes, we may conclude that belief is a commitment-incurring attitude.

In order to apply the inconsistency test to an attitude with imperative content we will need to employ the “do not” operator I introduced in chapter 2, section 2.7. In the case of the desire to bring about **P** (which has the imperative content: ‘bring about **P**!’), the attitudinal contrapositive is the desire to not bring about **P** (which has the imperative content: ‘do not bring about **P**!’). Having determined the identity of the attitudinal contrapositive of the desire to bring about **P**, we are now in a position to apply the inconsistency test. We merely have to juxtapose the desire to bring about **P** and the desire not to bring about **P**, and then ask if an agent with both attitudes is ipso facto liable to rational criticism. I believe that the most plausible answer to this question is no. For example, consider an agent who wants to quit smoking, but who continues to have the craving for cigarettes. A situation may arise in which such an agent may be described as both desiring to bring it about that she has a cigarette, and desiring to bring it about that she does not have a cigarette. It seems wrong to say that the agent is liable to rational criticism for simply having the inconsistent desires.

If the above application of the inconsistency test seems inconclusive (and some readers may find themselves with less than clear intuitions about what we should say about the preceding case), we may also apply a modified version of the inconsistent test by juxtaposing a desire with an attitude of another type. However, in order for the modified version of the inconsistency test to be effective, the second attitude must display the following two characteristics. Firstly, it must be the attitudinal contrapositive of the attitude we are inquiring about. Since desires have imperative content, this means that (given the definition of an attitudinal contrapositive) the second attitude must also have imperative content. Secondly, it must be pre-determined that the second attitude is commitment-incurring. The rationale for this second requirement stems from the fact that an agent who has a pair of inconsistent attitudes, only one of which is commitment-incurring, is not liable to rational criticism. This means that if the second attitude is not already known to be commitment-incurring, and we find that an agent with both attitudes is not liable to rational criticism, we will have no way of knowing which of the two attitudes is not commitment-incurring. At most, we will know that at least one of the two attitudes is not commitment-incurring, without being sure which one it is, or if both are.

For our second attempt to apply the inconsistency test to desire, I propose to use an intention as the attitudinal contrapositive. We begin by establishing that the intention is indeed a commitment-incurring attitude. To do this, we merely juxtapose the intention to bring about **P** and the intention not to bring about **P**, and then ask if an agent with both attitudes would be

liable to rational criticism. The answer is yes. For example, an agent who both intended to have a cigarette and intended not to have a cigarette would be liable to rational criticism. It follows that an intention is a commitment-incurring attitude. We may now apply the inconsistency test to desire by juxtaposing the desire to bring about **P** with the intention not to bring about **P**. (Notice that since both the desire and intention have imperative content, and the illocutionary content of the intention is equivalent to the negation of the illocutionary content of the desire, the intention is the attitudinal contrapositive of the desire.) For example, we may consider an agent who desires to have a cigarette but who intends not to do so. This is a situation that almost anyone who decides to quit smoking will find themselves in since their desire to have a cigarette is likely to persist even after they have adopted the intention not do so. It is clear that we would not want to say that such an agent is liable to rational criticism. This would have the rather unpalatable consequence that anyone who continues to desire something after adopting an opposing intention would be liable to rational criticism. If my insistence that an agent is not liable to rational criticism because she continues to have a desire after adopting a contrapositive intention is right, then it follows that a desire is not a commitment-incurring attitude.

## 5.6. Reply to the Perverse Desire Argument

Given that a desire is not a commitment-incurring attitude, it follows that a desire does not display the kind of rational significance displayed by a belief. If this is right, then the most plausible remaining strategy for establishing that desires are rationally significant is to argue that desires display the kind of rational significance displayed by perception—namely, by providing rational support for commitment-incurring attitudes. To this end, the cognitivist may hold that desiring to bring about **P** favours intending to bring about **P** over not intending to bring about **P**. In this respect, the desire to bring it about that **P** is once again analogous to perceiving that **P**, since perceiving that **P** provides rational support by favouring believing **P** over not believing **P**. Hence, the cognitivist may claim that desires provide rational support with respect to intentions in a sense analogous to how perceptual appearances provide rational support with respect to beliefs.

Hence, the following account of what it means for a desire to provide rational support seems available to the cognitivist. Consider the following practical analogue to a deductive inference:

(E): I intend to make an omelette.

(F): If I intend to make an omelette, then it is necessary that I intend to break eggs.

(G): I intend to break eggs.

Let us assume that (E) and (G) are reports of an intention, and that one is committed to (E) or (G) just in case one possesses the corresponding intentions: namely, the intention to make an omelette and the intention to break eggs. And let us assume that (F) is something I justifiably believe to be true. Being committed to (E) and (F) incurs a commitment to adopt the intention reported in (G). However, one may discharge this commitment by either adopting the intention

reported in (G) or by giving up one's commitment to (E) or (F). We therefore need some rational basis for determining which way we should discharge this commitment.

The sophisticated cognitivist may insist that it is at this stage that the rational support provided by a desire becomes salient. Suppose that my reason for intending to make an omelette is my desire to make an omelette. We may make sense of the claim that my desire speaks in favour of my intention to make an omelette by observing that the desire is not indifferent with respect to two competing ways of discharging my commitment. To wit, it favours remaining committed to (E), and adopting the intention reported in (G), over giving up my commitment to (E). Hence, the cognitivist may argue that the desire to make an omelette speaks in favour of intending to make an omelette in the very sense that perceiving that the street is wet speaks in favour of believing that the street is wet. Moreover, the cognitivist may argue that we can only make sense of how being committed to (E) and (F) may justify adopting the intention reported in (G) if we assume that desires provide the type of rational support that is the practical analogue to the type of rational support provided by perceptual appearance. Consequently, insofar as we are committed to saying that perceptual appearances provide rational support in the context of theoretical inferences, then, by parity of reasoning, we should also be committed to saying that desires provide rational support in the context of practical inferences.

If we take the cognitivist to be committed to the thesis that the rational support a desire provides is the practical analogue to that provided by a perceptual appearance rather than that provided by a belief,<sup>81</sup> then the perverse desire objection loses its force. Consider an agent who visually perceives the two lines of the Müller-Lyer illusion as being different lengths even though she believes them to be equal in length. The sense in which perceiving that **P** represents **P** as being true is a sense that allows an agent to perceive that **P** even though she fully recognises that **P** is false. Consequently, if the sophisticated cognitivist holds that the desire to bring about **P** represents **P** as an outcome it would be good to bring about, in a sense analogous to how perceiving that **P** represents **P** as being true, then the sophisticated cognitivist may allow for the possibility of perverse desires—i.e., desiring something one fully recognises to be bad. Along these lines, the sophisticated cognitivist may hold that for Satan to desire evil is for evil to appear good to Satan. However, this no more entails that Satan believes or judges that evil is good than experiencing what one knows to be a perceptual illusion entails believing or judging that it is veridical. Since, according to the present view, Satan is neither benighted nor confused about what is actually good, the sophisticated cognitivist is able to preserve the satanic quality of Milton's Satan; to wit, she is able to preserve the intuition that Satan desires evil even though he does not mistakenly believe that it is good.<sup>82</sup>

The preceding cognitivist account allows for a much more flat-footed interpretation of Satan's declaration, "Evil be thou my Good", than Velleman's analysis. Velleman claims that "Satan is here resolving to desire...evil."<sup>83</sup> However, if (as Velleman insists) desire does not aim

---

<sup>81</sup> Sergio Tenenbaum [2007] advocates something very close to this view. See and Cf. Scanlon [1998: 38].

<sup>82</sup> The present reply to the perverse desire objection is consistent with Michael Stocker's [1979] denial of the claim that only the (believed) good attracts.

<sup>83</sup> Velleman [1992: 18].

at the good, then it becomes mysterious why resolving to desire evil amounts to taking evil to be “Good”. If (as Velleman insists) there is absolutely no sense in which evil is good to Satan, then why does not Satan simply declare: “Evil be thou my Evil”? One natural answer to this question is that saying “Evil be thou my Evil” would fail to convey that Satan was resolving to desire and pursue evil. On the contrary, had Satan declared “Evil be thou my Evil”, it would give the misleading impression that Satan did not desire evil after all. If this is right, then it is hard to think of any clearer evidence of a conceptual connection between desire and the good. But if (as Velleman insists) there is no conceptual connection between desire and the good, there would be no reason to think that the declaration “Evil be thou my Good” expresses a resolution to desire evil. This suggests that there is an internal tension in Velleman’s analysis. If Velleman is correct in saying that there is no conceptual connection between desire and the good, then he is not entitled to say—as he in fact does—that the declaration “Evil be thou my Good” expresses a resolution to desire evil. But if Velleman is not entitled to say that the declaration “Evil be thou my Good” expresses a resolution to desire evil, then his would-be counterexample to sophisticated cognitivism never gets off the ground.

## 5.7. Conclusion

In this chapter, I have used the framework for understanding the rational significance of an attitude provided in chapter 4—namely, the Dual-Significance Thesis—to diagnose where Velleman’s objection to sophisticated cognitivism goes wrong. Velleman depicts the sophisticated cognitivist as committed to the claim that desires provide rational support. He also depicts the cognitivist as committed to an analogy between desire and belief. However, these two characterisations are at odds with each other. If the cognitivist were really committed to an analogy between desire and belief, then she would be committed to saying that a desire is a commitment-incurring attitude. But, by Velleman’s own lights, the cognitivist is actually committed to saying that desires provide rational support. This means that the cognitivist is actually committed to an analogy between desire and perception. However, once we register that a desire is analogous to a perception rather than a belief, Velleman’s objection to sophisticated cognitivism lapses.

## 6

## ARE DESIRES RATIONALLY SIGNIFICANT?

---

### 6.1. Introduction

In this chapter, I advance and defend an argument against the claim that desires provide rational support. Unlike the arguments against cognitivism currently represented in the literature (such as the arguments by David Velleman discussed in the previous chapter), which attempt to show that desires do not provide rational support by impugning the claim that desires represent their object as good, the present chapter argues that even if desires do represent their object as good, it does not follow that desires provide rational support. On the contrary, we have good reason to think that even if desires did represent their object as good, desires would fail to provide anything that could plausibly be called rational support. In sections 6.2 and 6.3, I present the core of my argument. In sections 6.4 – 6.7, I respond to a number of objections to my argument, and in the process of so doing elaborate on key features of my analysis of desire. In section 6.8, I conclude by highlighting some of the wider implications of my argument for the nature of practical reasoning.

### 6.2. Desire, Perception, and Coin-flips

There are two ways in which one attitude may favour another that are relevant to the present discussion:

**Definition 6.1 (Conformatively Favour):**

X conformatively favours Y *IFF* X offers a verdict that coincides with the verdict of Y.

**Definition 6.2 (Confirmatively Favour):**

X confirmatively favours Y *IFF* X offers evidence that corroborates the verdict of Y.

The distinction between conformatively favouring and confirmatively favouring may be illustrated by contrasting guessing that **P** and perceiving that **P**. Suppose that I guess that the street is wet. My guess offers a verdict that coincides with the verdict of the belief that the street is wet. As such, my guess that the street is wet conformatively favours the belief that the street is wet. However, the guess that the street is wet fails to lend any evidential support for the belief that the street is wet. As such, my guess that the street is wet does not confirmatively favour believing that the street is wet. Contrast this with perceiving that the street is wet. Like

guessing that the street is wet, perceiving that the street is wet offers a verdict that coincides with the verdict of the belief that the street is wet. As such, perceiving that the street is wet also conformatively favours the belief that the street is wet over the belief that the street is not wet. However, perceiving that the street is wet also offers evidential support for the belief that the street is wet. As such, perceiving that the street is wet confirmatively favours believing that the street is wet.

The fact that some attitude,  $X$ , conformatively favours some other attitude,  $Y$ , is not sufficient for  $X$  to provide rational support for  $Y$ . For example, guessing that  $P$  fails to provide rational support for believing that  $P$ , even though guessing that  $P$  conformatively favours believing that  $P$ . This suggests that we must add a further requirement to the characterisation of the kind of rational significance displayed by perceptual appearances limned in chapter 4. Hence, I propose the following definition of rational support:

**Definition 6.3 (Rational Support):**

$X$  provides  $S$  with rational support for  $P$  IFF  $X$  confirmatively favours the adoption of an attitude that rationally commits  $S$  to  $P$ .

If the cognitivist is going to succeed in showing that desires provide rational support in the same sense that perception does, it is not enough for her to show that the desire to bring about  $P$  conformatively favours intending to bring about  $P$ . She must show that the desire to bring about  $P$  confirmatively favours intending to bring about  $P$ . However, I argue that the cognitivist is at most entitled to say that the desire to bring about  $P$  conformatively favours the intention to bring about  $P$ . This is because the cognitivist has failed to rule out a certain possibility: namely, that when it comes to its rational significance, a desire may be more analogous to a guess than a perceptual experience.

Consider the following example of an agent engaged in a piece of practical reasoning:

**Example 6.1 (Chocolate Intention):**

Lin is at the supermarket attempting to decide between two unfamiliar flavours of ice cream: chocolate and vanilla. Since she has no idea what ice cream ingredients work well together, she has no evidence that either flavour is better than the other. Lin decides to flip a coin. She assigns the value of head-side-up to the outcome: 'purchase chocolate rather than vanilla', and the value of tail-side-up to the outcome: 'purchase vanilla rather than chocolate'. When the coin lands head-side-up, she proceeds to the cash register, pays, and exits the supermarket with chocolate in her grocery bag. When Lin arrives home, her roommate inquires into why she bought chocolate as opposed to some other flavour. Lin responds: "I wasn't familiar with either of the two flavours available, and so I had no reasons in favour of purchasing one flavour over another. So I decided to base my decision on a coin-flip."

I take the following to be true of the Chocolate Intention example:

- (1) The result of Lin's coin-flip (which I am here assuming is no better than a guess) conformatively favours purchasing chocolate rather than vanilla.
- (2) The result of Lin's coin-flip does not confirmatively favour purchasing chocolate rather than vanilla.
- (3) Lin is not liable to rational criticism for basing her decision to purchase chocolate rather than vanilla on a coin-flip.

(1) follows straightforwardly from our definition of the conformatively favouring relation. When Lin assigns the value of head-side-up to the proposition 'purchase chocolate rather than vanilla', and the value of tail-side-up to the proposition 'purchase vanilla rather than chocolate', the fact that the coin lands head-side-up means that the coin-flip offers a verdict that coincides with the intention to purchase chocolate rather than vanilla. (2) follows from the fact that the coin-flip fails to provide Lin with evidence (or anything remotely analogous to evidence) that it would be good for the outcome 'purchase chocolate rather than vanilla' to be brought about. I take (3) to be uncontroversial. Were we in Lin's roommate's shoes, we wouldn't think that there was something rationally problematic about Lin's answer.

Now, consider the following modified version of the Chocolate Intention example:

**Example 6.2 (Chocolate Belief):**

Lin is at the supermarket attempting to decide between two unfamiliar flavours of ice cream: chocolate and vanilla, just as in Example 6.1. However, in addition to adopting the intention to purchase chocolate rather than vanilla, let us suppose that Lin also comes to believe that chocolate is preferable to vanilla. When her roommate inquires into why she believes that chocolate is preferable to vanilla, Lin responds: "Since I wasn't familiar with either of the two flavours available, I didn't have any reason to believe one flavour was preferable to the other. So I decided to base my belief on a coin-flip."

In contradistinction to the first example, Lin's response in the second example is unsatisfying. This suggests an important disanalogy between rationally permissible belief and rationally permissible intention. While it is rationally permissible to adopt an intention to bring about **P** even though one recognises that one has no reason in favour of bringing about **P**, it is not rationally permissible to adopt the belief that **P** if one recognises that one has no reason in favour of the truth of **P**. In section 6.8, I offer an explanation of this disanalogy. However, for now, it is sufficient to note that the disanalogy exists, and to get clear on how it fits into the rhetorical structure of my argument.

Saying that an attitude is rationally significant means, *inter alia*, that it plays a role that cannot be played by an attitude that lacks rational significance. This is easily illustrated by considering the rational significance of perceptual appearances. For example, suppose I am looking out my office window and it perceptually appears to me as if it is raining. *Ex hypothesi*, this perceptual appearance provides rational support for the belief that it raining. By

contrast, consider the case in which I am in a windowless room and I guess that it is raining. Ex hypothesi, my guess that it is raining does not provide rational support for the belief that it is raining. Part of what this means is that the guess that it is raining cannot be substituted for the perceptual appearance of it raining without altering the rational standing of my belief that it is raining. The preceding point generalises. One cannot substitute an attitude, X, which is rationally significant relative to some belief or intention, Y, with another attitude, Z, which is not rationally significant relative to Y, and have the rational standing of Y remain unchanged. Call this the *non-substitutability principle*.

Consider once again the Chocolate Intention example. The situation Lin finds herself in is, in a certain sense, unusual. We do not typically base our practical decisions (i.e., our decisions about which intentions to adopt) on coin-flips. In most real life situations, our practical decisions are based on desires. Of course, since Lin is unfamiliar with the two ice cream flavours (and, as a result, does not have a preference either way), she is not in a position to base her decision on a desire. This is why Lin is forced to resort to a coin-flip. And therein lies the rub. In the absence of a desire, a coin-flip may be deployed to play the same role that the desire would play. Moreover, deploying a coin-flip (in lieu of a desire) preserves the rational standing of the relevant intention. In other words, one may substitute a coin-flip (i.e., something that provides no rational support) for a desire without altering the rational standing of an intention based on said desire. Given the non-substitutability principle, I take this to show that the desire also fails to provide rational support.

Some may object that the Lin example is so unusual that the results of the preceding argument fail to generalise. After all, it has already been acknowledged that the Lin example is an atypical case. Given that we are not typically in the situation in which Lin finds herself, why should we take the Chocolate Intention example as a basis for a general conclusion about desires? I believe that this objection rests on a misunderstanding of the rhetorical structure of my argument. What the Chocolate Intention example illustrates is that if we take the standard case in which an agent makes a decision based on a desire, and we replace the desire with a coin-flip (i.e., something that fails to provide rational support), we in no way alter the rational standing of an intention based on the desire. This, I claim, shows that a desire has no more rational significance than a coin-flip. If we are going to assess whether or not this conclusion generalises, then the question we need to ask is this: Is there some salient respect in which the practical decision with which Lin is confronted is structurally dissimilar to the range of cases in which an agent adopts an intention based on a desire? I will argue that the answer is no.

Let us begin with the most basic case, one in which an agent is confronted with a practical decision involving only one course of action. In such a case, the choice is between action and inaction. The agent may either adopt an intention to act, or adopt an intention not to act. (Of course, there is a third case, one in which an agent simply fails to adopt any intention whatsoever. However, since in such a case there is no intention that stands in need of rational support, it fails to provide us with a test case for determining if desires provide rational support. We may therefore ignore such cases for now.) We can easily imagine a coin-flip being substituted for a desire in such cases without changing the rational standing of the relevant intention. There is therefore no reason to think that the basic case poses any special challenge for my argument. All other (non-basic) cases involve a choice between two or more courses of



action. Moreover, we can easily imagine a coin-flip (or, in the case of practical decisions involving more than two courses of action, some other suitable random selection procedure) being substituted for the desire without changing the rational standing of the intention. Given that the basic cases (i.e., those involving a choice between action and inaction) and non-basic cases (i.e., those involving a choice between two or more actions) exhaust the logical space of possibilities as far as practical decisions are concerned, and given that a coin-flip (or some other random selection procedure) may be substituted for a desire in both basic and non-basic cases without changing the rational standing of the relevant intention, it follows that the conclusion of my argument is generalizable.

### 6.3. The Juxtaposition of Desire and Coin-flip

In the previous section, I argued that the fact that a desire can be replaced by a coin-flip (without altering the rational standing of an intention based on it) suggests that a desire offers no more rational support than a coin-flip. However, the following example seems to cast doubt on this conclusion:

**Example 6.3 (Chocolate Desire):**

Lin is at the supermarket trying to decide between two flavours of ice cream. Suppose that Lin desires to 'purchase chocolate rather than vanilla'. But for no reason whatsoever, she decides to disregard her desire and flip a coin. She assigns the value of head-side-up to the outcome 'purchase chocolate rather than vanilla' and the value of tail-side-up to the outcome 'purchase vanilla rather than chocolate'. She flips the coin and it lands tail-side-up—favouring the intention to 'purchase vanilla rather than chocolate'. Lin then adopts the intention to 'purchase vanilla rather than chocolate' based on her coin-flip.

There seems to be something problematic about Lin's decision to purchase vanilla rather than chocolate based on her coin-flip, given that she desires to purchase chocolate rather than vanilla. In fact, some would go so far as to say that Lin is being irrational. Moreover, her putative irrationality seems tied to the fact that she privileges the coin-flip over her desire. If this is right, then it seems as though a desire does have a rational significance that a coin-flip lacks, and that this fact explains why Lin is irrational for acting on the coin-flip rather than on her desire.

Of course, we can imagine cases in which Lin has independent reasons to disregard her desire or independent reasons to flip a coin. In such cases, deciding to privilege a coin-flip over her desires is unlikely to be rationally problematic. However, the Chocolate Desire example asks us to consider a case in which Lin has no independent reasons to disregard her desire and no independent reasons to flip a coin. This is because if we allow for the possibility that Lin has independent reasons to disregard her desire or independent reasons to flip a coin, then it may turn out that desires do provide reasons, but that the reasons in question are outweighed by the reasons she has for disregarding her desire or the reasons she has for flipping a coin. In order to preclude this possibility, we must restrict our attention to cases in which an agent has no

independent reasons to disregard her desire and no independent reasons to flip a coin. However, it is equally important that the agent in question has no independent reasons for having the desire she has, for it may turn out that she has more reason to purchase chocolate rather than vanilla (in agreement with her desire), but that the reason in question is provided not by the desire, but by whatever consideration provides her with a reason to have the desire. In order to rule out this possibility, we must restrict our attention to cases in which there are no independent considerations that constitute a reason for Lin to desire to purchase chocolate rather than vanilla. Hence, in order for the Chocolate Desire example to successfully show that a desire offers more rational support than a coin-flip, it must include the stipulation that an agent has no independent reasons for either the desire or the coin-flip.

To sum up, the claim of the Chocolate Desire example is this: when we consider a situation in which Lin has no independent reasons for either her desire or for flipping the coin, we find ourselves with the intuition that Lin is rationally required to act on her desire rather than on the coin-flip. This suggests that a desire and a coin-flip are not really on par, rationally speaking.

#### 6.4. The Juxtaposition of Coin-flip and Coin-flip

It may be argued that once we begin to fill out the Chocolate Desire example in such a way that it becomes clear that Lin has no independent reasons for her desire to purchase chocolate rather than vanilla and no independent reasons for her decision to flip a coin, it becomes less intuitively obvious that Lin is being irrational when she decides to act on her coin-flip rather than on her desire. According to this line of thought, we are only inclined to think that Lin is being irrational when she decides to act on the coin-flip rather than on her desire because, under normal circumstances, an agent has reasons for her desires. Since whatever consideration speaks in favour of the desire to bring about **P** is also likely to speak in favour of the intention to bring about **P**, it is likely that whenever an agent desires to bring about **P**, she also has a reason to intend to bring about **P**. Moreover, it is the reasons the agent has for her desire, rather than the desire per se, that account for the intuition that Lin is being irrational when she disregards her desire and acts on her coin-flip. According to the present analysis of the Chocolate Desire example, once we take seriously the idea that Lin has no reason for her desire—i.e., that there is no consideration that speaks in favour of her desire and that may potentially speak in favour of an intention based on her desire—the intuition that Lin is being irrational for disregarding her desire and acting on the coin-flip lapses.

Unfortunately, the above line of reply is not one that is available to me since it is inconsistent with other features of my view. Simply put, it confuses justificatory and explanatory reasons. Recall, by my lights, the only kind of reasons one may have for a desire are the kind that *explain* why one has a desire, rather than the kind that *justify* having a desire. This is because I hold that a desire, like a perceptual appearance, is not a reason-receiving attitude. Consequently, saying that we have justificatory reasons for our desires is on par with saying that we have justificatory reasons for our perceptual appearances. Both claims involve a kind of category mistake. The result is that, on pain of inconsistency with other parts of my view, I cannot argue that the intuition that Lin is irrational when she acts on the coin-flip rather

than on her desire is rooted in the fact that we typically have justificatory reasons for our desires.

Admittedly, there is a similar line of argument that is available to me (albeit one I find less than convincing). Even if we assume that one cannot have reasons for a desire, one may still have independent reasons for bringing about the outcome that one desires. In other words, one may have a justificatory reason for the corresponding intention.<sup>84</sup> Thus, it may be argued that the intuition that Lin is guilty of irrationality for acting on her coin-flip rather than on her desire is rooted in the fact that we normally have independent reasons for bringing about the outcomes we desire. While there may be something right about the preceding reply, I do not think it provides the full or most accurate story. What the preceding reply does seem to get right is that the intuition that Lin is guilty of irrationality when she acts on the coin-flip rather than on her desire weakens once we register that Lin has no independent reasons for bringing about the desired outcome. However, it remains unclear that the intuition fails to persist altogether. Moreover, since the Chocolate Desire example is supposed to already take into consideration the fact that the Lin has no independent reasons for bringing about her desire, the present reply does little more than stubbornly insist that the Chocolate Desire example does not really have the force it purports to have.

Contra the two preceding lines of reply, I wish to not only grant that Lin is being irrational when she decides to act on her coin-flip rather than on her desire, but to also provide additional theoretical motivations for thinking that she is being irrational. I agree that the intuition that Lin is guilty of irrationality for acting on her desire rather than on her coin-flip becomes less clear once we fill in the details of the example to include the requirement that Lin has no independent reasons to bring about the desired outcome. However, I do not take this to be evidence that the claim is unsound. Instead, I take it to be evidence that our pre-theoretical intuitions are not very firm on this particular question. As such, the question is not one that is best settled by an appeal to pre-theoretical intuitions. If I am right, then we may need to invoke theoretical motivations for thinking that Lin is in being irrational. This is what I attempt to do below. However, my eventual aim is not to show that a desire has a rational significance that coin-flips lack. After all, one of the advertised aims of this dissertation is to show that this conclusion remains unestablished. Instead, I will be offering an alternative explanation of why Lin is irrational for acting on her coin-flip rather than on her desire.

To summarise, I wish to grant that there is something rationally problematic about Lin's decision to act upon her coin-flip rather than her desire. However, I will argue that this does not show that desires have a rational significance that a coin-flip lacks. To this end, I invite the reader to consider yet another version of the Lin example:

**Example 6.4 (Chocolate Coin-flip):**

Lin is at the supermarket trying to decide between two unfamiliar flavours of ice cream: chocolate and vanilla. She assigns head-side-up to the proposition 'purchase chocolate

---

<sup>84</sup> This comports with my claim (in chapter 5) that an intention is a commitment-incurring attitude. Since I claim that the ability of an attitude to incur rational commitments is tied to the fact that it is a reason-receiving attitude, then the claim that an intention is a commitment-incurring attitude fits hand in glove with the claim that an intention is a reason-receiving attitude.

rather than vanilla' and tail-side-up to the proposition 'purchase vanilla rather than chocolate'. When she flips the coin, it lands head-side-up—favouring the intention to purchase chocolate rather than vanilla. However, for no reason whatsoever, Lin decides to flip the coin a second time. On the second coin-flip, the coin lands tail-side-up—favouring the intention to purchase vanilla rather than chocolate. Lin then adopts the intention to purchase vanilla rather than chocolate based on the second coin-flip.

I believe that the agent in the preceding example is guilty of irrationality. In fact, I will argue that she is guilty of two separate instances of irrationality. The first instance of irrationality occurs when she decides to flip the coin a second time, and the second occurs when she adopts the intention to act on the second coin-flip despite having no reason to disregard the first. Let us begin with the irrationality tied to her decision to flip the coin a second time. Flipping the coin a second time constitutes an instance of irrationality because doing so amounts to the adoption of a selection procedure that is poorly suited for its purpose. Recall, Lin resorts to flipping a coin in order to determine if she should purchase chocolate or vanilla. A single coin-flip is well suited for this purpose, since it is an instance of single-answer sortition (i.e., a random selection procedure that yields one and only one answer). However, by deciding to flip the coin a second time, Lin initiates a multiple answer selection procedure. If the coin is flipped twice, the space of possible results is <heads, heads>, <tails, tails>, <heads, tails>, and <tails, heads>. This means that there is a 50% chance (2 in 4) that the coin-flips would yield a combination of both heads and tails, thereby failing to provide a definite answer to the question of whether she should purchase chocolate or vanilla. In fact, the Chocolate Coin-flip example is just such a case. I sum, given that flipping the coin twice is poorly suited for the purpose of arriving at a single definite answer of what to do, I claim that Lin's decision to flip the coin a second time is irrational.

The second instance of irrationality occurs when Lin decides to act on her second coin-flip, despite having no reason to disregard the first coin-flip. As I already alluded to in chapter 4, an agent has a rational obligation to extricate herself from any pre-existing forms of irrationality. For the agent in the Chocolate Coin-flip example, one option would be to switch to a 2-out-of-3 selection procedure. This would involve flipping the coin a third time and allowing the third coin-flip to have the tie-breaking vote. A second option would be to un-deputize one of the two coin-flips. I will have more to say about the notion of deputization in the next section. But to anticipate the discussion to come, *deputization* is the mental act of bestowing an extra-mental process with the authority to guide one's actions. The mere act of throwing a coin up in the air and observing what side it lands on, even if intentionally performed, does not qualify as a selection procedure. A coin-flip only counts as a selection procedure if an agent has decided to take it as a guide for her actions. This decision is one an agent may reverse by simply deciding to no longer take the coin-flip as a guide for her actions. In so doing, she strips it of its standing as a selection procedure. Hence, Lin may extricate herself from the irrationality incurred by her reliance on a sub-optimal selection procedure—i.e., one that juxtaposes two coin-flips—by stripping one of the two coin-flips of the authority to guide her actions. Since Lin fails to take either of the above courses—or any other that would extricate her from her pre-existing irrationality—and instead opts to simply act on the second

coin-flip, despite the fact that there is no reason to privilege the second over the first, Lin is again guilty of irrationality.

### 6.5. The Juxtaposition of Coin-flip and Desire

It should now be clear where my argument is heading. I wish to argue that the agent who relies on a selection procedure that juxtaposes a desire and a coin-flip is irrational for the same reason as an agent who relies on a selection procedure that juxtaposes two coin-flips. In short, it is the sub-optimality of relying on a selection procedure that has a 50% chance of yielding conflicting results (when one could just as easily rely on one that always yields a single definite result) that accounts for the intuition that there is something irrational about Lin's decision to disregard her desire in favour of her coin-flip. She is employing a multiple-answer selection procedure when she should be using a single-answer one.

However, this reply will need to be enriched further if it is going to provide us with a satisfactory solution to the challenge posed by the Chocolate Desire example. This is because the Chocolate Desire example is not simply meant to show that an agent who decides to flip a coin even though she already has a desire is guilty of irrationality, but also that an agent is irrational for deciding to act on a coin-flip when there is a desire present. To this end, it is important to emphasise that the temporal ordering of the desire versus the coin-flip does not matter to the question of which takes priority. Even if the desire comes after the coin-flip, it still seems to take rational priority. For example, let us suppose that when Lin is first confronted with the decision to purchase chocolate or vanilla ice cream, she finds herself with no desire either way. She therefore decides to flip a coin. But suppose that immediately after flipping the coin, she finds herself with a desire to purchase chocolate rather than vanilla. Suppose further that Lin still decides to act on her coin-flip. In such a case, Lin is once again guilty of irrationality. Hence, it seems as though Lin's desire always takes rational priority over the coin-flip. The takeaway of the Chocolate Desire example is this: irrespective of temporal ordering, whenever one juxtaposes a desire and a coin-flip, the desire always takes rational priority.

I wish to grant that whenever we juxtapose a desire and a coin-flip, the desire takes priority in the manner just suggested. However, I will offer an alternative explanation of why this is so, one that does not rely on or entail the claim that a desire has more rational significance than a coin-flip. Let us begin by noting that, in the context of the present discussion, both a desire and a coin-flip constitute a type of selection procedure for deciding between two options. That is, if we assume that an agent confronted with the choice between intending to purchase chocolate rather than vanilla or intending to purchase vanilla rather than chocolate, then the fact that she desires to purchase chocolate rather than vanilla plays the role of a selection procedure for determining which of the two intentions she should adopt. The same is true of the coin-flip. However, an agent who desires to purchase chocolate rather than vanilla is passive with respect to the adoption of the selection procedure instantiated by her desire. That is to say, the desire is not something the agent chooses to have. Rather, it is something that, in a certain sense, *happens* to the agent. The sense in question is analogous to the sense in which a perceptual appearance is not something an agent chooses to have but is rather something that happens to her.

There are a number of ways to make sense of the notion of passivity currently at play. One common (albeit controversial) way is in terms of a control principle, according to which an agent is passive vis-à-vis her perceptual appearances and desires because they are not under her control. Apart from certain exceptional circumstances, we do not choose to have the perceptual appearances or desires we do in fact have. Another way of making sense of the notion of passivity currently at play is in terms of the distinction between agential and sub-agential processes, according to which an agent is passive in relation to her desires and perceptual appearances because they represent the perspective of a sub-agential system rather than that of the agent as a whole. Which (if either) of these ways of making sense of the claim that an agent is passive with respect to her desires will not matter for our present purposes. All that matters is that an agent is passive with respect to her having a certain desire in a sense that is analogous to how she is passive with respect to the having of a perceptual appearance. Moreover, the sense in question is one that tracks liability to rational appraisal. That is to say, an agent is passive with respect to her desires and perceptual appearances in a way that entails that she is not liable to rational appraisal for having (or failing to have) a desire or perceptual appearance.

By contrast, an agent is active with respect to the adoption of a coin-flip as a selection procedure. As before, there are a number of ways we may choose to make sense of the notion of “active” being presently being invoked. We may (albeit controversially) make sense of it in terms of a control principle, according to which Lin is active with respect to her coin-flip because the adoption of a coin-flip as a selection procedure is under her control. Alternatively, we may say that the adoption of the coin-flip as a selection procedure is an agent-level rather than sub-agential decision. That is to say, while sub-agential processes determine which desires we happen to have, it is agential-level processes that determine whether or not we employ a coin-flip as a selection procedure. Whatever story we choose to tell, the crucial point for our present purposes is that the adoption of a coin-flip as a selection procedure is something for which an agent is liable to rational appraisal.<sup>85</sup> And therein lies the solution to our puzzle. Why is Lin guilty of some irrationality when she decides to flip a coin even though she already has a desire? Because she is relying on a sub-optimal selection procedure—i.e., she is employing a multiple-answer form of sortition when she should be using a single-answer one. Why is it that the desire always takes priority over the coin-flip, irrespective of the temporal order in which they appear? Because of the two selection procedures, the coin-flip is the only one for which the agent is liable to rational appraisal. Since she is not (nor can be) expected to stop having the desire she has, she can only fulfil her rational obligation to free herself from the irrationality incurred by adopting a sub-optimal selection procedure by un-deputising the coin-flip. This means that she will always be rationally obligated to act on her desire rather than on her coin-flip, when the two are in competition. The takeaway of the present reply is this: a desire takes

---

<sup>85</sup> I mention these two candidate explanations of why a desire is passive and of why a coin-flip is active in order to underscore that the sense in which I am describing a desire as passive is the counterpart to the sense in which I am describing a coin-flip as passive. Presumably, there are many ways in which we can describe a desire or coin-flip as active or passive, and so it is important to ensure that the senses are not somehow orthogonal to each other. By offering these candidate explanations of why desires are passive and coin-flips are active, I hope to make it clear that the candidate explanations are opposite sides of the same coin (no pun intended). This should give us some reason to think that I have identified a single sense of passive versus active, rather than senses that are somehow orthogonal to each other.

rational priority over a coin-flip not because the former has a rational significance that the latter lacks, but because the decision to employ a coin-flip as a selection procedure is something for which an agent is liable to rational appraisal while the having of a desire is not.

## 6.6. Passive versus Active Selection Procedures

The conclusion of the immediately preceding section is that a desire takes rational priority over a coin-flip (qua competing selection procedures) not because desires have a rational significance that coin-flips lack, but because an agent is active with regards to the latter (in a sense that tracks liability to rational appraisal) but passive with regards to the former. However, the following example seems to pose a challenge to this view:

### Example 6.5 (Chocolate Advice):

Lin is at the supermarket with her friend, Aasif, trying to decide between two unfamiliar flavours of ice cream. Since neither Lin nor Aasif have had ice cream before (and they have no idea what ingredients work well together) they have no reason to think that one flavour is better than the other. They are in a hurry to leave and Lin, not having any reason to think that either flavour is better than the other, is having a hard time making a decision. In order to speed things along, Aasif suggests that Lin purchase chocolate rather than vanilla. Moreover, let us suppose that Aasif arbitrarily picks chocolate and that Lin is aware of this fact. It follows that Aasif's suggestion fails to evidentially favour purchasing chocolate rather than vanilla. (Like the coin-flip, this is equivalent to a guess by Lin.) However, before adopting any intention and unrelated to Aasif's advice, Lin comes to have a desire to purchase vanilla rather than chocolate. However, Lin decides to disregard her desire and act on Aasif's suggestion.

Once again, it seems that it would be irrational for Lin to purchase chocolate rather than vanilla based on Aasif's advice and rational for her to purchase vanilla rather than chocolate based on her desire. However, unlike the coin-flip, Lin is passive relative to Aasif's advice. Hence, if my explanation of why desires take rational priority over coin-flips is correct, Aasif's advice should be rationally on par with Lin's desire. However, this does not seem to be case since Lin once again seems to be rationally obligated to act on her desire (rather than on Aasif's advice). This appears to show that my diagnosis of why desires take rational priority over a coin-flip is incorrect.

The present objection takes as a premise the assumption that Aasif's advice (like a desire) does not involve an active expression of Lin's agency. In response, I wish to argue that, in all relevant respects, Lin's decision to act on Aasif's advice is like deciding to act on a coin-flip rather than like deciding to act on a desire. What the present objection gets right is that there is a sense in which Lin is passive relative to Aasif's advice—to wit, she has no control over what advice Aasif gives. However, Lin is passive relative to the coin-flip in precisely this sense—to wit, she has no control over which side the coin lands on. Hence, there is no important difference between the coin-flip and Aasif's advice in this respect. However, I hold

that there is also a sense in which Lin is active relative to both the coin-flip and Aasif's advice, and that this is the very sense in which Lin is passive relative to her desire.

When I say that Lin is active relative to a coin-flip, I am not merely talking about her act of throwing the coin up in the air and observing what side it lands on. (Given that the set of bodily movements in question are intentionally performed, throwing the coin up in the air and observing what side it lands on does indeed involve the active expression of someone's agency. However, this is not what I have in mind when I describe Lin as active relative to her coin-flip.) Rather, I am talking about her deputizing the coin-flip with the authority to guide her actions. This deputization is an active expression of Lin's agency and it is analytically prior to her adoption of the intention that complies with what the coin-flip says. Once Lin deputizes the coin-flip with the authority to tell her what to do, the fact that the coin lands head-side-up means that she is committed to adopting the intention to purchase chocolate rather than vanilla. Since Aasif's advice is external to Lin's cognition, it must also be deputized by Lin with the authority to guide her actions. Thus, Aasif's advice can only guide Lin's action if Lin deputizes it with the required authority, and this deputization is something for which Lin is liable to rational appraisal.

By contrast, an agent's desires do not require deputization in order to guide her actions. In other words, while an agent only counts as acting on a coin-flip or as acting on someone's advice if she deputizes that coin-flip or that advice, an agent may count as acting on a desire without any deputization whatsoever. That this deputization is unnecessary for an agent to count as acting on a desire is established by the following consideration: deputizing a coin-flip or someone else's advice with authority is only something that cognitively sophisticated agents can engage in. For example, my pet hamster cannot do it. However, even cognitively unsophisticated agents (like my pet hamster) can act on a desire. Thus, acting on a desire cannot require deputizing it. I take this to show that there is something extra that is present in the case of acting on a coin-flip or acting on someone's advice that is not present in the case of acting on a desire. A detailed analysis of what this something extra amounts to is a project worth pursuing. However, only two points about this something extra are immediately relevant. First, we can safely say that the something extra is something that only cognitively sophisticated agents exhibit. Hamsters cannot exhibit it. Second, the something extra is something for which an agent is liable to rational appraisal. For example, we can rationally criticize someone for allowing inferior advice to factor into her deliberation. Thus understood, I refer to the something extra as the deputization of an extra-mental process with the authority to guide one's actions. I claim that this deputization is an active expression of someone's agency; it is something for which an agent is liable to rational appraisal.

## 6.7. The Zero-Sum Structure of Intention

In section 6.2, I highlight an important disanalogy between rationally permissible beliefs and rationally permissible intentions. I will now attempt to offer an explanation of this disanalogy. Of course, the claim that there is a disanalogy of the kind I have identified may be true even if my preferred explanation of why the disanalogy exists is false. Hence, the arguments I have offered in defence of the claim that there is a disanalogy between rationally permissible belief



and rationally permissible intention may be assessed independently of any of the considerations presented below.

Let us refer to any deliberation that involves the weighing of reasons in order to determine if **P** is true as *theoretical deliberation* and any deliberation that involves the weighing of reasons in order to determine if **P** is to be made true as *practical deliberation*. Theoretical deliberation always involves a choice between three doxastic attitudes: believing, disbelieving, or withholding belief and disbelief. This means failing to believe that **P** does not entail disbelieving **P** and failing to disbelieve **P** does not entail believing **P**. Withholding belief and disbelief always remains an option. (For simplicity, I will henceforth refer to the attitude of withholding belief and disbelief towards **P** as withholding **P**.) Withholding **P**, as I am using the expression, is not the same as failing to adopt an attitude of belief and disbelief towards **P**. One may fail to adopt an attitude of belief and disbelief towards **P** because one has simply not considered **P**. In such a case, one neither believes nor disbelieves **P**. But one is not withholding **P** either. One simply has not taken any attitude towards **P**. In short, withholding **P** is as much an attitude towards **P** as believing or disbelieving **P**.

I believe that the disanalogy between rationally permissible belief and rationally permissible intention, highlighted in section 6.2, is tied to the absence of a practical analogue to withholding. Consider the following example:

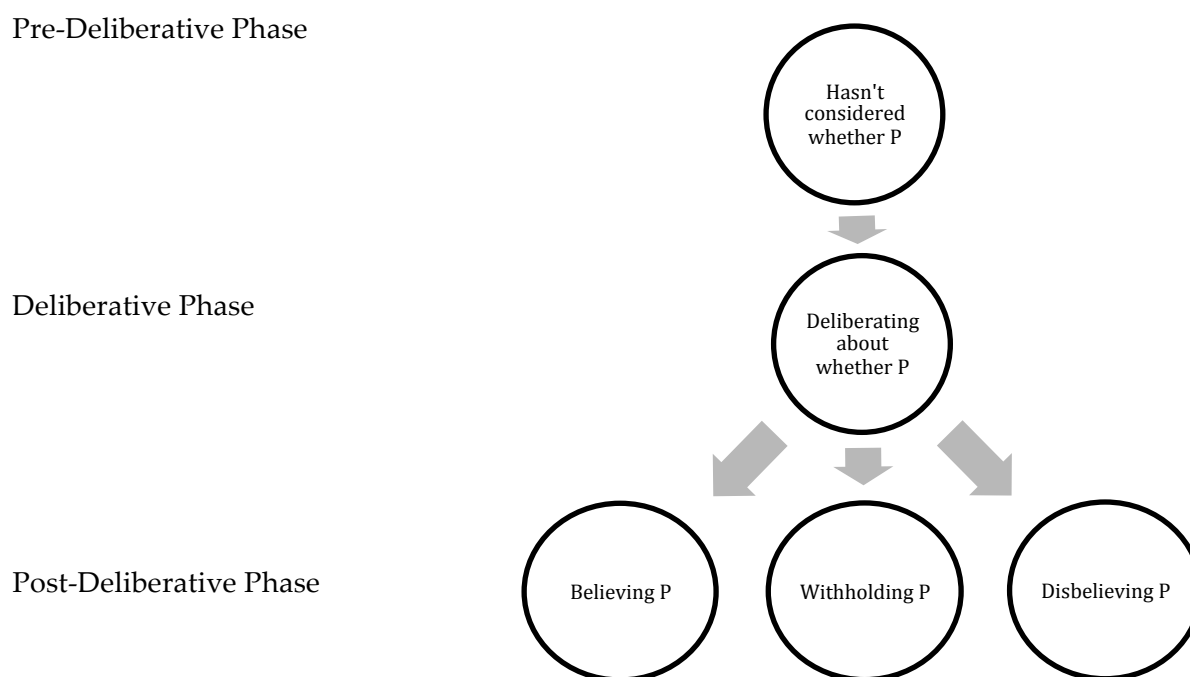
**Example 6.6 (Lorry Driver Intention):**

A lorry driver, Parvati, is driving down a long stretch of road and is trying to decide if she should stop for a break. To put the point in action theoretic terms, she is deliberating about whether or not the following outcome is to be brought about: 'I am stopping for a break'. After weighing all the considerations in favour of stopping for a break (e.g., she would be able to get some much deserved rest), and all the considerations in favour of not stopping for a break (e.g., she is more likely to make her delivery on time), she comes to the conclusion that the sum of her reasons are inconclusive either way. She has just as much reason to stop for a break as she does for not stopping for a break. Parvati decides to flip a coin: head-side-up representing 'stop for a break!' and tail-side-up representing 'do not stop for a break!'. When the coin lands head-side-up, Parvati pulls the lorry off to the side of the road and removes the key from the ignition.

Parvati's decision has a zero-sum structure. If she adopts the intention to continue driving, she has ipso facto adopted the intention not to stop for a break. The same is true, *mutatis mutandis*, if she adopts the intention to stop for a break. There is no attitude of practical withholding that she may adopt as an end point of her deliberative process. In sum, while there are three possible doxastic attitudes one can take towards **P**—believing **P**, disbelieving **P**, and withholding **P**—there are only two possible volitional attitudes one can take towards an outcome—intending to bring about **P** and intending not to bring about **P**. On the present analysis, not intending to bring about **P** is not an attitude towards bringing about **P** but rather the absence of an attitude towards bringing about **P**.

The claim that there is no practical analogue to withholding may seem surprising, if not unsound. After all, isn't there a difference between adopting the intention not to bring about **P** and simply failing to intend to bring about **P**? And is not failing to intend to bring about **P** the practical analogue to withholding **P**? I answer yes and no, respectively. There is indeed a difference between intending to not bring about **P** and not intending to bring about **P**. However, the latter is the practical analogue not of withholding, but of failing to adopt any doxastic attitude towards **P**. We may gain a better handle on why this is so by conceiving of theoretical deliberation in terms of the following three-tier flowchart:

**Figure 6.1 (Theoretical Deliberation Flowchart):**



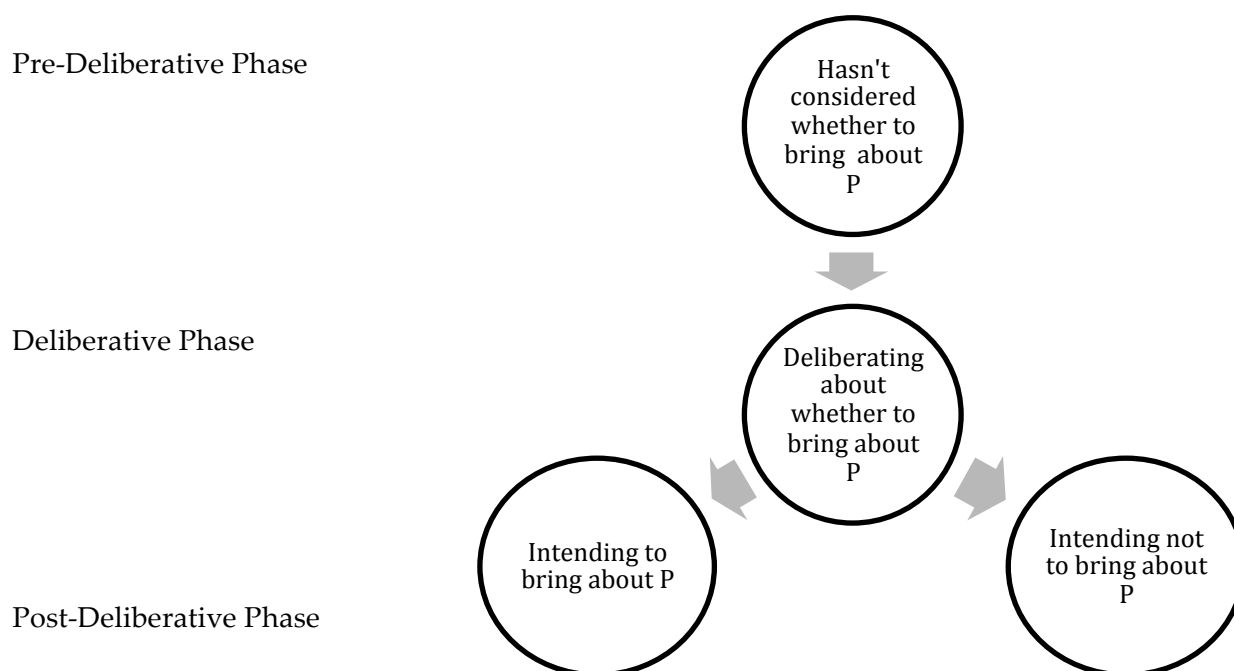
The top tier of Figure 6.1 (i.e., the pre-deliberative phase) stands proxy for an agent's thoughts prior to considering whether or not **P** is true. It does not designate an attitude towards **P**, but rather the absence of any attitude towards **P**. The middle tier of Figure 6.1 (i.e., the deliberative phase) stands proxy for when an agent is in the process of weighing reasons in order to figure out the right course of action. I wish to remain non-committal on the question of whether or not deliberating about **P** is itself an attitude towards **P**.<sup>86</sup> What I do maintain is that an agent who is deliberating about **P** is in a different psychological state, with respect to **P**, than an agent who has adopted one of three possible doxastic attitudes one may take towards **P**—

<sup>86</sup> My suspicion is that deliberating about **P** involves one or more non-doxastic attitude(s) towards **P**, but the specific conception of deliberation we have will not matter for our present purposes.

believing **P**, disbelieving **P** and withholding **P**. Like the pre-deliberative phase, the deliberative phase does not entail the adoption of any particular doxastic attitude towards **P**. It is only the bottom tier of Figure 6.1 (i.e., the post-deliberative phase) that entails the adoption of a doxastic attitude towards **P**.

By contrast, practical deliberation may be conceived of in terms of the following flow chart:

**Figure 6.2 (Practical Deliberation Flowchart):**



Consider once again the lorry driver. She is in the process of deliberating about whether she should continue driving or take a break. During this time, it is true that she has neither decided to continue driving nor to stop for a break. However, this process of deliberation, I have been urging, is the practical analogue not to withholding **P**, but to deliberating about whether or not **P** is true. Now, fast forward to the point at which the driver has finally made a decision, or the post-deliberative phase. It is at this point that the disanalogy between theoretical and practical deliberation becomes apparent. The lorry driver has only two options. She can either continue driving or stop for a break. If she adopts the intention to continue driving, she has ipso facto adopted the intention not to stop for a break. The same is true, *mutatis mutandis*, if she adopts the intention to take a break. There is no attitude of practical withholding that she may adopt as an end point of her deliberative process. In sum, while there are three possible doxastic attitudes one can take towards **P**—believing **P**, disbelieving **P**, and withholding **P**—there are

only two possible volitional attitudes one can take towards an outcome—intending to bring about **P** and intending not to bring about **P**. Not intending to bring about **P** is not an attitude towards bringing about **P** but rather the absence of an attitude towards bringing about **P**.

Let us consider once again the example of the practical analogue to a modus ponens inference (first introduced in chapter 5, section 5.6, and reproduced below):

(E): I intend to make an omelette.

(F): If I intend to make an omelette, then it is necessary that I intend to break eggs.

(G): I intend to break eggs.

It is clear that there remains a need for commitment-involving attitudes (i.e., attitudes that display the kind of rational significance displayed by belief) in the practical sphere. For example, the intention to make an omelette incurs a rational commitment to do whatever is necessary to make an omelette (e.g., break eggs), so long as one persists in intending to make an omelette. Thus, the claim that intentions are rationally significant remains secure.

However, it is less clear that there is a need for an attitude that provides rational support, of the kind provided by perceptual appearances, in the practical sphere. Apart from incurring rational commitments (in the manner that intentions do), the only role we need a practical attitude to play is that of a selection procedure. Consider an agent who has equal amounts of rational support both for and against making an omelette. (This, of course, includes cases in which one has no rational support either for or against making an omelette.) According to the zero-sum view, one does not require any further rational support to settle the question of whether or not one should make an omelette. A coin-flip may do the job. In other words, while there remains a need for a selection procedure, the procedure in question need not be rational. It need not represent the selected outcome as good. It need not put the agent in touch with some consideration (independent of the selection procedure itself) that speaks in favour of the selected outcome.

I believe that attitudes with imperative content play the role of such a selection procedure. They do this not by representing a certain outcome as good, but by conformatively favouring a certain outcome. Rationally or evidentially speaking, they provide us with a verdict when we need one, and nothing more. This, of course, is not to suggest that desires are random in same sense that coin-flips are. There are a host of phylogenetic and ontogenetic factors that contribute to us having the particular desires that we have. Some of these factors may be known to us, some may be revealed via psychoanalysis, and some may forever remain outside our ken; but it seems mistaken to label them (or the role they play in giving rise to our desires) random. Nevertheless, while the aetiology of a desire may be quite different from that of a coin-flip, the role that the former plays in practical deliberation is (like the role of the latter) merely one of sortition. This explains why the latter may be substituted for the former, and is itself explained by the zero-sum structure of intention.

## 6.8. Conclusion

This dissertation began with the question: what is the rational significance of desire? The answer we arrive at is little or none, depending on whether or not one thinks a coin-flip or guess is rationally significant. More precisely, if we define rational significance along the lines suggested in chapter 4—i.e.,  $X$  is rationally significant if and only if  $X$  plays a role (qua psychological state) in generating a rational obligation to adopt or refrain from adopting a belief or intention—then a desire has no rational significance. In the present chapter, I have argued that a desire provides no more rational support than a coin-flip. Insofar as a coin-flip fails to provide rational support, it follows that a desire fails to provide rational support. However, this conclusion is not as bleak as it may initially seem. For I have also argued that there is no need for rational support (of the kind provided by perception) in the practical sphere. Specifically, I claim that while it is not rationally permissible to adopt a belief if one recognises that one has no rational support for it, it is rationally permissible to adopt an intention even if one recognises that one lacks any rational support for it. Moreover, I have attempted to explain the aforementioned disanalogy in terms of the zero-sum structure of intention. One implication of this difference between the practical and the theoretical is that cognitivist arguments that presuppose or exploit an analogy between desire and belief or desire and perception are much too quick. Specifically, the cognitivist strategy for arguing that desires provide rational support by appealing to an analogy between desire and belief or perception suffers from a fatal flaw, since the practical and theoretical inferences are disanalogous in precisely this respect: namely, there is no need for rational support (i.e., the kind of rational significance displayed by perception) in the practical sphere. It is this fact that explains why a desire may do all the work we need it to do even though it has no more rational significance than a coin-flip.

## BIBLIOGRAPHY

---

- Altham, J. E. J. [1986], "The Legacy of Emotivism", in Graham Macdonald and Crispin Wright, eds., *Fact, Science and Morality: Essays on A. J. Ayer's Language, Truth and Logic*. (Oxford, UK: Blackwell Publishing): 275-288.
- Anscombe, E. [2000], *Intention*. (Cambridge, MA: Harvard University Press).
- Audi, R. [1986a], "Acting for Reasons", *Philosophical Review*, 95, 4: 511-546.
- Audi, R. [1986b], "Intending, Intentional Action, and Desire", in Marks, J., (ed.) *The Ways of Desire*. (Chicago, IL: Precedent Publishing).
- Audi, R. [1993], *The Structure of Justification*. (Cambridge, UK: Cambridge University Press).
- Baier, K. [1958], *The Moral Point of View*. (Ithaca, NY: Cornell University Press).
- Blackburn, S. [1998], *Ruling Passions*. (Oxford, UK: Oxford University Press).
- Beaver, D. [2001], *Presupposition and Assertion in Dynamic Semantics*. (Stanford, CA: CSLI Publications).
- Bond, E. J. [1983], *Reason and Value*. (Cambridge, UK: Cambridge University Press).
- Botterill, G. and Carruthers, P. [1999], *The Philosophy of Psychology*. (Cambridge, UK: Cambridge University Press).
- Boyd, R. [1988], "How to be a Moral Realist", in G. Sayre-McCord (ed.) *Essays on Moral Realism*. (Ithaca, NY: Cornell University Press): 181-228.
- Braithwaite, R.B., [1932-1933], "The Nature of Believing", *Proceedings of the Aristotelian Society*, 33: 129-146.
- Brandt, R. B. [1969], "Rational Desires", *Proceedings and Addresses of the American Philosophical Association*, 43: 43-64.
- Brentano, F. [1874], *Psychology from an Empirical Standpoint*. (London, UK: Routledge and Kegan Paul).

- Broome, J. [1997], "Reason and Motivation", *Proceedings of the Aristotelian Society, Supplementary Volume*, **71**: 131-46.
- Broome, J. [2000], "Normative Requirements", in J. Dancy (ed.) *Normativity*. (Oxford, UK: Blackwell): 78-99.
- Broome, J. [2001], "Are Intentions Reasons? And How Should We Cope with Incommensurable Values?" in C. W. Morris and A. Ripstein (eds.) *Practical Rationality and Preference: Essays for David Gauthier*. (Cambridge, UK: Cambridge University Press): 98-120.
- Broome, J. [2002], "Practical Reasoning", in J. L. Bermudez and A. Millar (eds.) *Reason and Nature: Essays in the Theory of Rationality*. (Oxford, UK: Clarendon Press): 85-112.
- Broome, J., [2004], "Reasons", in Wallace, R. J. et al. (eds.) *Reasons and Value: Themes from the Moral Philosophy of Joseph Raz*. (Oxford, UK: Clarendon Press).
- Broome, J. [2007], "Wide and Narrow Scope", *Mind*, **116** (462): 359-70.
- Brunero, J. [2007], "Are Intentions Reasons?" *Pacifica Philosophical Quarterly*, **88**: 424-444.
- Brunero, J. (2010). "The Scope of Rational Requirements." *Philosophical Quarterly*, **60** (238): 28-49.
- Burge, T. [2003], "Perceptual Entitlement", *Philosophy and Phenomenological Research*, **67** (3): 503-548.
- Byrne, A., [2009], "Experience and Content", *Philosophical Quarterly*, **59** (236): 429-51.
- Byrne, A. and Hájek, A. [1997], "David Hume, David Lewis, and Decision Theory", *Mind*, **106**: 411-28.
- Castañeda, H. N. [1975], *Thinking and Doing: The Philosophical Foundations of Institutions*. (Boston, MA: D. Reidel Publishing Co.).
- Chellas, B. F. [1971], "Imperatives", *Theoria*, **37**: 114-129.
- Chisholm, R. [1980], "A Version of Foundationalism" in Wettstein, et al (eds), *Midwest Studies in Philosophy V*. (Minneapolis, MN: University of Minnesota Press).
- Christensen, D. [1994], "Conservatism in Epistemology", *Noûs*, **28** (1): 69-89.
- Collins, J. [1988], "Belief, Desire, and Revision", *Mind*, **97** (387): 333-342.

- Cummins, D. [1996], "Evidence For the Innateness of Deontic Reasoning", *Mind and Language*, **11**: 160-190.
- Cummins [1996], "Dominance Hierarchies and the Evolution of Human Reasoning", *Minds and Machines*, **6**: 463-480.
- Dancy, J. [2000], *Practical Reality*. (Oxford, UK: Oxford University Press).
- Dancy, J. [2004], *Ethics without Principles*. (Oxford, UK: Oxford University Press).
- Darwall, S. [1983], *Impartial Reason*, (Ithaca, NY: Cornell University Press).
- Davidson, D. [1963], "Actions, Reasons, and Causes", reprinted in D. Davidson, [1980], *Essays on Actions and Events*. (Oxford, UK: Oxford University Press): 3-19.
- Davidson, D. [1969], "How is Weakness of Will Possible?", in Joel Feinberg (ed.), *Moral Concepts*. (London, UK: Oxford University Press).
- Davidson, D. [1978], "Intending", *Essays on Actions and Events*. (Oxford, UK: Clarendon Press): 83-102.
- Davidson, D. [1984a], "Moods and Performances", *Inquiries into Truth and Interpretation*. (Oxford, UK: Clarendon Press): 109-122.
- Davidson, D. [1984b], "Communication and Convention", *Inquiries into Truth and Interpretation*. (Oxford, UK: Clarendon Press): 265-280.
- Davidson, D. [2001], "A Coherence Theory of Truth and Knowledge", *Subjective, Intersubjective, Objective*. (Oxford, UK: Clarendon Press).
- Davidson, D. [2004], "Paradoxes of Irrationality", *Problems of Rationality*. (London, UK: Oxford University Press).
- Davies, M. [1992], "Perceptual Content and Local Supervenience", *Proceedings of the Aristotelian Society*, **92**: 21-45.
- Dreckmann, F. [1999], "Animal Belief and Their Contents", *Erkenntnis*, **51**: 93-111.
- Dretske, F. [1999], "Machines, Plants and Animals: The Origins of Agency", *Erkenntnis*, **51**: 19-31.



- Dretske, F. [2003], "Skepticism: What Perception Teaches", in Luper (ed.), *The Sceptics*. (Hampshire, UK: Ashgate Publishing, Limited): 105–118.
- Dretske, F. [2005] "Is Knowledge Closed Under Known Entailment?" in Steup, M. and Sosa, E. (eds.) *Contemporary Debates in Epistemology*. (Malden, MA: Blackwell Publishing).
- Bernard, L. C., et al. [2005], "An Evolutionary Theory of Human Motivation", *Genetic, Social, and General Psychology Monographs*, **131** (2): 129-184.
- Brittan, G., [1999], "The Secrets of Antelope", *Erkenntnis*, **51**: 59-77.
- Edwards, J., [1957], *Freedom of Will*, ed. P. Ramsey. (New Haven, CT: Yale University Press).
- Engel, A. K. [2003], "Temporal binding and the neural correlates of consciousness", in A. Cleeremans (ed.), *The Units of Consciousness: Binding, Integration and Dissociation*. (Oxford, UK: Oxford University Press): 132-152.
- Engel, P. [2004], 'Truth and the Aim of Belief', in D. Gillies (ed.), *Laws and Models in Science*. (London: King's College Publications).
- Epstein, S. and Pacini, R. [1999], "Some Basic Issues Regarding Dual-process Theories from the Perspective of Cognitive-experiential Theory", In Chaiken, S. and Trope, Y. (eds.) *Dual Process Theories in Social Psychology*. (New York: Guildford Press): 462-82.
- Erickson, R. [1975], *Sound Structure in Music*, (California: University of California Press).
- Evans, D. [2002], "The Search Hypothesis of Emotion", *British Journal of Philosophy of Science*, **53**: 497-509.
- Evans, J. [2008], "Dual Process Accounts of Reasoning, Judgement, and Social Cognition", *Annual Review of Psychology*, **59**: 255-278.
- Evans, J. [2009], "How Many Dual-process Theories Do we Need? One, Two, or Many?" in Evans, J. and Frankish, K. (eds.), *In Two Minds: Dual Processes and Beyond*. (Oxford, UK: Oxford University Press): 33-54.
- Fraassen, B. C. [1968], "Presupposition, Implication, and Self-Reference", *The Journal of Philosophy*, **65** (5): 136-152.
- Frankfurt, H. G. [1978], "The Problem of Action", *American Philosophical Quarterly*, **15** (2): 157-162.

- Fumerton, R. [2007], "Epistemic Conservatism: Theft or Honest Toil?", *Oxford Studies in Epistemology*, Vol. 2, by Szabo Gendler, T. and Hawthorne, J. (eds). (Oxford University Press).
- Garrard, E. and McNaughton, D. [1998], "Mapping Moral Motivation", *Ethical Theory and Moral Practice*, **1**: 45-59.
- Grice, H. P. [1957], "Meaning", *The Philosophical Review*, **66** (3): 337-388.
- Hare, R. M. [1963], *Freedom and Reason*. (Oxford: Clarendon Press).
- Harman, G. [1986], *Change in View*. (Cambridge, MA: MIT Press).
- Harman, G. [1990], "The Intrinsic Quality of Experience", *Philosophical Perspectives*, **4**: 31-52.
- Hawkins, J. [2008], "Desiring the Bad Under the Guise of the Good", *The Philosophical Quarterly*, **58** (231): 244-264.
- Heuer, U. [2004], "Reasons and Actions and Desires", *Philosophical Studies*, **121**: 43-63.
- Humberstone, L. [1987], "Wanting as Believing", *Canadian Journal of Philosophy*, **17** (1): 49-62.
- Hurley, S. [2003a], "Animal Action in the Space of Reasons," *Mind and Language*, **18** (3): 231-256.
- Hurley, S. [2003b], "Making Sense of Animals: Interpretation vs. Architecture", *Mind and Language*, **18** (3): 273-281.
- Kenny, A. J. [1963], *Action, Emotion and Will*. (London, UK: Routledge and Kegan Paul).
- Kolodny, N. [2005], "Why Be Rational?" *Mind*, **114** (445): 509-63.
- Kornell, N. [2009], "Metacognition in Humans and Animals", *Current Directions in Psychological Science*, **18** (1): 11-15.
- Korsgaard, C. [1986], "Skepticism about Practical Reason", *The Journal of Philosophy*, **83** (1): 5-25.
- Korsgaard, C. [1992], "The Normativity of Instrumental Reason", in G. Cullity, and B. Gaut, (eds). [1997], *Ethics and Practical Reason*. (Oxford: Oxford University Press): 215-254.
- Lewis, D. [1973], "Causation", *Journal of Philosophy*, **70**: 556-567.

- Lewis, D. [1988], "Desire as Belief", *Mind*, **97** (387): 323-332.
- Lewis, D. [1996], "Desire as Belief II", *Mind*, **105** (418): 303-313.
- Little, M. O. [1997], "Virtue as Knowledge: Objections from the Philosophy of Mind", *Nous*, **31** (1): 59-79.
- Marcus, R. B. [1990], "Some Revisionary Proposals About Belief and Believing", *Philosophy and Phenomenological Research*, **1**: 133-153.
- McDowell, J. [1994], *Mind and World*. (Cambridge, MA: Harvard University Press).
- McGinn, C. [1982], *The Character of the Mind*. (Oxford, UK: Oxford University Press).
- Meiland, J. W. [1970], *Nature of Intention*. (London, UK: Methuen & Co. Ltd.).
- Millar, A. [2004], *Understanding People: Normativity and Rationalizing Explanation*. (Oxford, UK: Clarendon Press).
- Milligan, T. [2007], "Whimsical Desires", *Ratio*, **20**: 308-319.
- Milliken, J. [2008], "In a Fitter Direction: Moving Beyond the Direction of Fit Picture of Belief and Desire", *Ethical Theory and Moral Practice*, **11**: 563-571.
- Moran, R. [2001], *Authority and Estrangement*. (Princeton, NJ: Princeton University Press).
- Moss, J. [2010], "Aristotle's Non-Trivial, Non-Insane View that Everyone Always Desires Things under the Guise of the Good," in *Desire, Practical Reason, and the Good*. (Oxford, UK: Oxford University Press): 65-81.
- Nagel, T. [1978], *The Possibility of Altruism*. (Princeton, NJ: Princeton University Press).
- Nerlich, G. [1967], "Presupposition and Classical Logical Relations", *Analysis*, **27** (3): 104-106.
- Nozick, R. [1981], *Philosophical Explanations*, (Cambridge, UK: Cambridge University Press).
- Oddie, G. [2005], *Value, Reality, and Desire*. (New York, NY: Oxford University Press).
- Parfit, D. [1984], *Reasons and Persons*. (Oxford, UK: Clarendon Press).

- Parfit, D. [1997], "Reasons and Motivation", *Proceedings of the Aristotelian Society, Supplementary Volume*, **71**: 99-130.
- Parfit, D. [2006], "Normativity" in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Volume 1. (Oxford, UK: Clarendon Press): 325-380.
- Peacocke, C. [1992], *A Study of Concepts*. (Cambridge, MA: The MIT Press).
- Pettit, P. [1987], "Humeans, Anti-Humeans and Motivation", *Mind*, **96** (384): 530-533.
- Plantinga, A. [2000], *Warranted Christian Belief*. (New York, NY: Oxford University Press).
- Plantinga, A. [1993], *Warrant: The Current Debate*. (New York, NY: Oxford University Press).
- Price, H. [1989], "Defending Desire-as-Belief", *Mind*, **98** (389): 119-127.
- Prinz, J. and Clark, A. [2004], "Putting Concepts to Work: Some Thoughts for the Twenty-first Century", *Mind and Language*, **19** (1): 57-69.
- Quine, W. V. O. [1951], "Two Dogmas of Empiricism", in *From a Logical Point of View*. (New York, NY: Harper & Row).
- Quinn, W. [1993], *Morality and Action*. (Cambridge, UK: Cambridge University Press).
- Railton, P. [1986], "Moral Realism", *Philosophical Review*, **95**: 163-207.
- Railton, P. [2006], "Normative Guidance", *Oxford Studies in Metaethics*, **1**: 3-34.
- Raz, J. [2010], "On the Guise of the Good", in *Desire, Practical Reason, and the Good*. (Oxford: Oxford University Press): 111-137.
- Ross, P. W. [2002], "Explaining Motivated Desires", *Topoi*, **21**: 199-207.
- Scanlon, T. [1998], *What We Owe to Each Other*. (Cambridge, MA: Harvard University Press).
- Schall J. D. [2001], "Neural basis of Deciding, Choosing and Acting", *Nature Reviews Neuroscience*, **2**: 33-42.
- Schall, J. D. [1999], "Weighing the Evidence: How the Brain makes Decisions", *Nature Reviews Neuroscience*, **2**: 108-109.

- Schroeder, M. [2004], "The Scope of Instrumental Reason." *Philosophical Perspectives (Ethics)*, **18**: 337-62.
- Schroeder, M. [2009], "Means End Coherence, Stringency, and Subjective Reasons", *Philosophical Studies*, **143** (2): 223-248.
- Schueler, G. [1995], *Desire: Its Role in Practical Reason and Explanation of Action*. (Cambridge, MA: MIT Press).
- Schueler, G. [2003], *Reasons and Purposes*. (Oxford, UK: Clarendon Press).
- Searle, J. [1983], *Intentionality*. (Cambridge, UK: Cambridge University Press).
- Searle, J. R. [1985], *Expression and Meaning: Studies in the Theory of Speech Acts*. (New York: Cambridge University Press).
- Searle, J. R. and Vanderveken, D. [1985], *Foundations of Illocutionary Logic*. (Cambridge: Cambridge University Press).
- Sellars, W. [1954], "Presupposing", *Philosophical Review*, **63**: 197-215.
- Setiya, K. [2007], *Reasons without Rationalism*. (Princeton, NJ: Princeton University Press).
- Setiya, K. [2010], "Sympathy for the Devil," in *Desire, Practical Reason, and the Good*. (Oxford: Oxford University Press): 82-110.
- Sklar, L. [1975], "Methodological Conservatism", *Philosophical Review*, **74**: 186-191.
- Slooman, S. A. [2002], "Two Systems of Reasoning", In T. Gilovich, D. Griffin, and D. Kahneman (eds.) *Heuristics and Biases: The Psychology of Intuitive Judgement*. (Cambridge, UK: Cambridge University Press): 379-98.
- Smith, M. [1988], "Reason and Desire", *Proceedings of the Aristotelian Society*, **88**: 243-258.
- Smith, M. [1987], "The Humean Theory of Motivation", *Mind*. **96**: 36-61.
- Smith, M. [2004a], "Humeanism, Psychologism and the Normative Story", in *Ethics and the A Priori*. (Cambridge: Cambridge University Press): 146-154.
- Smith, M. [2004b], "The Possibility of Philosophy of Action", in *Ethics and the A Priori*. (Cambridge: Cambridge University Press): 155-177.

- Soames, S. [1989], "Presupposition", in Gabbay, D. and Guenther, F. (eds.), *Handbook of Philosophical Logic*. (Dordrecht: Reidel), 4: 553-616.
- Stalley, R. F. [1972], "Intentions, Beliefs, and Imperative Logic", *Mind*, 81: 18-28.
- Stalnaker, R. [1972], "Pragmatics", in Davidson, D. and Harman, G. (eds.), *Semantics of Natural Language*. (Reidel: Dordrecht): 389-408.
- Stalnaker, R. [1973], "Presuppositions", *The Journal of Philosophical Logic*, 2: 447-457.
- Stalnaker, R. [1974], "Pragmatic Presuppositions", in Munitz, M. and Unger, P. (eds.), *Semantics and Philosophy*. (New York: New York University Press): 197-214.
- Stalnaker, R. [1998], "On the representation of context", *Journal of Logic, Language and Information*, 7: 3-19.
- Stampe, D. [1987], "The Authority of Desire", *The Philosophical Review*, 96: 335-81.
- Stenius, E. [1967], "Mood and Language-Game", *Synthese*, 17: 254-274.
- Steup, M. [1996], *An Introduction to Contemporary Epistemology*. (Upper Saddle River, NJ: Prentice Hall).
- Stocker, M. [1979], "Desiring the Bad: An Essay in Moral Psychology", *Journal of Philosophy*, 76 (12): 738-753.
- Strawson, G. [1994], *Mental Reality*. (Cambridge, MA: MIT Press).
- Strawson, P. F. [1950], "On Referring", *Mind*, 59: 320-344.
- Strawson, P. F. [1952], *Introduction to Logical Theory*. (London, UK: Methuen).
- Strawson, P. F. [1954], "A Reply to Mr. Sellars", *Philosophical Review*, 63: 216-231.
- Tenenbaum, S. [1999], "The Judgment of a Weak Will", *Philosophy and Phenomenological Research*, 59 (4): 875-911.
- Tenebaum, S. [2003], "Accidie, Evaluation, and Motivation", *Weakness of Will and Practical Irrationality*. (Oxford: Clarendon Press).
- Tenenbaum, S. [2006], "Direction of Fit and Motivational Cognitivism", in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*. (Oxford, UK: Oxford University Press): 235-264.

- Tenenbaum, S. [2007], *Appearances of the Good*. (Cambridge, MA: Cambridge University Press).
- Tenenbaum, S. [2010], *Desire, Practical Reason, and the Good*. (Oxford, UK: Oxford University Press).
- Thompson, M. [2008], *Life and Action*. (Cambridge, MA: Harvard University Press).
- Travis, C. [2004], "The Silence of the Senses", *Mind*, **113** (449): 57-93.
- Tye, M. [1995], *Ten Problems of Consciousness*. (Cambridge, MA: The MIT Press).
- Velleman, D. [1992], "The Guise of the Good", *Noûs*, **26** (1): 3-26.
- Vahid, H. [2004], "Varieties of Epistemic Conservatism", *Synthese*, **141**: 97-122.
- Vranas, P. B. M. [2008], "New Foundations for imperative Logic I: Logical Connectives, Consistency, and Quantifiers", *Noûs*, **42**: 529-572.
- Vranas, P. B. M. [2010], "In Defense of Imperative Inference", *Journal of Philosophical Logic*, **39**: 59-71.
- Wallace, R. J. [2001], "Normativity, Commitment, and Instrumental Reason." *Philosophers' Imprint*, **1** (4): 1-26.
- Warnock, G. [1976], "Imperatives and Meaning," in Lewis, H. D. (ed.), *Contemporary British Philosophy: Personal Statements, Part 4*. (London, UK: Allen and Unwin): 294-304.
- Watson, G. [1987], "Free Action and Free Will," *Mind*, **96**: 145-72.
- Way, J. [2010], "Defending the Wide Scope Approach to Instrumental Reason", *Philosophical Studies*, **147** (2): 213-33.
- Zagzebski, L. [1996], *Virtues of the Mind*. (Cambridge, UK: Cambridge University Press).
- Zajonc, R. [1980], "Preference Needs No Inference", *American Psychologist*, **35**: 151.