

REASONING, BELIEF AND THE QUEST FOR TRUTH

by
Marianna Bergamaschi Ganapini

A dissertation submitted to Johns Hopkins University in conformity with the
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

May, 2017

ABSTRACT

In this dissertation, I offer a new explanation for the fact that we can't, as a result of reasoning, come to believe something simply because we want to, i.e. the fact that we can't believe at will. On my view, reasoning is the inferential process necessarily guided by the aim of arriving at a conclusion (e.g. forming a belief) sufficiently supported by normative reasons. I also defend the view that believing a proposition is tantamount to being disposed to use it as a default premise in reasoning. These two claims combined show that the impossibility of believing at will is the result of our capacity to reason. The limited power we have over our beliefs is thus a condition of possibility for being the type of reasoners we are.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Steven Gross. I feel lucky to have worked with him for all these years, and my work here was immensely improved by his invaluable insights and constructive objections. His support and availability have been key to allow me to complete this dissertation.

I like to thank the following people for their incredibly useful comments on the various presentations and chapter drafts that led to this dissertation: Justin Bledin, Kathrin Glüer-Pagin, Conor McHugh, Jonathan Way, John Waterman, Åsa Wikforss, and Michael Williams.

My deepest gratitude also goes to those who, many years ago now, encouraged me to apply to a U.S. doctoral program in Philosophy: Furio Cerutti, Wayne Davis, Stefano Poggi, and Achille Varzi.

Without my family and friends, I wouldn't have been able to write this dissertation. In particular, I would like to thank my parents, Myriam Bergamaschi and Luigi Ganapini, for their endless love and encouragement throughout all these years.

Most importantly, none of this would have been possible without Andrea Pedferri. His strength and unwavering support really motivated me to succeed in doing what I love.

This dissertation is dedicated to my 4-year-old daughter, Cecilia Pedferri-Ganapini. From the day she was born she has been my main source of inspiration in writing and thinking about philosophy.

CONTENTS

Abstract	ii
Acknowledgements	iii
Contents	iv
Introduction	1
Chapter 1: Defending Full Belief	31
Chapter 2: Belief and Its Consequences	77
Chapter 3: Beyond Rule-following: Reasoning With Purpose	125
Chapter 4: Reasons vs. Correctness: When is Reasoning Good?	163
Chapter 5: What Guides Belief	189
Conclusion	221
Bibliography	222
Curriculum Vita	223

INTRODUCTION

Suppose a demon offers you eternal happiness if you believe that Napoleon lived and died in New York City. Unfortunately, since you have no evidence to support it, you can't form this belief—at least not consciously and as a direct result of deliberation. The goal of my dissertation is to account for this puzzling fact, namely that our beliefs are not subject to the will as our actions are (at times this phenomenon is called doxastic involuntarism, transparency, exclusivity or, more simply, the fact that we can't believe at will – pick your favorite). My solution comes in three steps. First, I defend the claim that believing means being disposed to use a proposition as a default premise in reasoning. Second, I focus on reasoning and define reasoning as the inferential process guided by the constitutive aim of coming to a conclusion supported by sufficient normative reasons. Because it plays the role of a default premise in reasoning, belief is also generally expected to play the role of a normative reason. Arguably, normative reasons are either true or at least epistemically supported propositions (or mental attitudes). Hence, as a result of the constitutive aim we have when reasoning in general and the unique role belief has in reasoning, when forming beliefs through reasoning we are forced to (try to) form beliefs that are epistemically supported, i.e. beliefs that we are willing to use in further reasoning. And this is why, through reasoning, we can't come to believe what we want, but only that for which we take to have sufficient evidence. I conclude that the limited power we have over our beliefs is actually a condition of possibility for reasoning at all.

It is important to point out that this dissertation does not deal *directly* with issues about norms or normativity. I will not talk about the truth norm of belief, or the

normative relation between truth and belief. Here I don't deal with epistemic norms or epistemic normativity in general. I don't openly advocate for nor argue against evidentialism or pragmatism. I don't claim that we should form our beliefs in this or that way.

Contrary to what happens in the rest of the dissertation, though, this introduction will be devoted to norms, epistemic normativity, and what theorists believe we should care about when we form beliefs and direct our inquiry. In the initial pages here, I will explore the allegedly special relation many believe there is between truth and belief, and why this relation is seen as the source of epistemic normativity. In fact, my larger goal is to undermine this picture by offering a view of belief and reasoning that is substantially different from the widespread thesis that truth is the defining, constitutive, or essential aim or standard for belief. My goal, that is, is to eventually undermine a multifaceted truth-centered view of belief, reasoning and epistemic normativity. That might in turn have an influence on the debate between evidentialists and pragmatists possibly favoring the latter. Indeed, the claim that only evidential considerations can function as normative reasons to believe, is usually linked to the idea that belief's very nature is normatively directed at truth. And since I offer reasons to be skeptical of that claim, this may take away some of the support evidentialism enjoys.

Perhaps this will also open the door for those promoting a weak notion of truth. Lynch (2009), for example, has argued that its normative or teleological relation with belief is what defines our concept of truth. The deflationist theory of truth, he then adds, fails in this task and needs to be rejected. Since I hope to give enough reasons to doubt

Lynch's first premise, the result of my dissertation may turn out to be congenial to deflationist views of truth.

However, important as they are, I won't fully explore those issues here, and I will limit my scope to providing a new picture of belief, reasoning, and epistemic rationality that is independent of the notion of truth.

THAT SPECIAL RELATION BETWEEN TRUTH AND BELIEF ¹

It all starts with Bernard Williams' saying that "belief aims at truth" (1973: 136), an expression introduced in his analysis of the impossibility of believing at will. Many interpret it as a metaphor for the claim that truth is the standard of correctness of belief. As Gibbard writes, "[f]or belief, correctness is truth. Correct belief is true belief. My belief that snow is white is correct just in case the belief is true, just in case snow is white." (2005: 338) If so, then belief is governed by the following truth norm:

TN) it is correct to believe that p (if and) only if p is true.

This standard of correctness is, for many, a distinctive, individuating feature of belief: belief is the attitude that has truth as its standard of correctness, and belief's relation with truth distinguishes it from other attitudes (e.g. intention, supposition).

In philosophy and cognitive science, mental attitude types (e.g. beliefs, desires) are often defined in terms of how they are connected to other mental states. This relation is explained by looking at attitudes' input (or upstream) and output (or downstream) roles, namely the sort of inputs they are sensitive to and the sort of consequences they tend to have (Fodor, 1985; Nichols and Stich, 2003). When applied to belief, it has been

¹ I leave aside any reference to the normativity of content, and how that can be used to argue for the normativity of belief.

often argued that beliefs are individuated by their output role in causing action and the formation of other attitudes. Recently, however, it has been argued that belief's role in the mind is not enough to account for the uniqueness of belief *unless* such a role is framed in a way that mirrors beliefs' special relation with truth (see chapter 2 for an argument against this claim). And this has in part motivated the strong support Williams' phrase about belief's aim has lately enjoyed.

There are two, non-necessarily opposing, interpretations of belief's aim, though: a normative and a non-normative. I will spend the next few pages delving into the normativist reading, explaining its strengths while also pointing out the worries it has raised. After that, I will say a few things about the non-normativist, aim-focused, interpretation of the aim of belief.

NORMATIVISM

Recently, Gibbard (2005) pointed out that 'correctness' is normative: "Correctness, now, seems normative. More precisely, as we should put it, the concept of correctness seems to be a normative concept." I will follow Wedgwood in explaining what it means for TN to be normative. TN is normative in the sense that it is a constitutive feature of TN that it plays a regulative role in certain practices: if one engages in one of those practices, and "makes judgments about which moves" conform or do not conform to TN, "one is thereby committed to regulating one's moves within the practice by those judgments". That means, again following Wedgwood on this, that it is "irrational" for one to engage in this practice, while simultaneously making the judgment that move M violates TN and yet to make move M. It is irrational because, doing so,

involves mental states that “intuitively conflict with each other” (Wedgewood, 2002: 268). Thus, TN has – according to the Normativist – a constitutive regulative role in the practice of belief-formation. On this view, TN is normative with respect to belief-formation: engaging in belief formation commits one to conforming to it (I will come back to this below).

Before delving into the specifics of Normativism about belief, it is worth mentioning that this account of belief may be connected to the broader view – called Intentional Normativism – that *all* our primary mental states (e.g. emotions, intentions, desires) are constitutively normative, in the sense that those mental attitudes are individuated based on their respective *constitutive* norms (Velleman, 2000; Lynch, 2009).² More specifically, Intentional Normativism is a complex view usually seen as subscribing to a cluster of different claims. The first of these claims is that a norm N is constitutive of an attitude only if it is part of what a belief, desire, intention, or emotion is that it is an attitude governed by N. I will have more to say about what that means later, but for now let me point out that Intentional Normativism can be a *metaphysical* thesis concerning the nature of intentional states. Or, it can take the form of a *conceptual* thesis

² All Normativists understand normativity as non-reducible to the descriptive. Intentional Normativism itself, though, can be understood as a reductivist thesis: the intentional is actually reducible to the normative. The reductive normativist claims that there is a correct reductive account of the nature of intentional states, and any such reduction must refer to normative properties and relations which are themselves non-reducible. Brandom (1994), for instance, argues that intentional states must be explained in terms of normative moves in the game of giving and asking for reasons. Intentional states are thus explained as deontic statuses (e.g. commitments, entitlements) which are then explained in terms of interpretative attitudes (e.g. attributing a commitment, scorekeeping). There is also a non-reductionist form of Normativism. Such an account may claim merely that the normative and the intentional are essentially interdependent. In other words, on this view any adequate account of the intentional requires that one also mentions the normative and vice versa, without either of the two being reducible to the other.

about what is built into our concepts of intentional states³. That means that if N is constitutive of, say, belief, then N can either be constitutive of the concept of belief or part of the essence or nature of the attitude of belief.

Relatedly, N is a *sui generis* norm: it does not derive from the application of a more general normative standard. It is, in contrast, a standard that is unique, namely it only governs a specific mental state. In addition, it is a common assumption that N has an inescapable authority that silences any other non-*constitutive*, non-*sui generis* norm that may apply to mental attitudes, and determines a decisive reason for having or not having that attitude (Wedgwood, 2002: 268).

It is also important to notice that, although Normativism as such is not committed to any specific claim about what norm N has to be, it is often the case that constitutive norms are linked to the correctness conditions usually attributed to different attitudes (e.g. truth for belief, good for intention, dangerousness for fear). This, however, can be explained in various ways depending on how you interpret the normative force of those *norms* and the term ‘correctness’. There are two main camps, at least for the norm of belief: deontic and axiological Normativism. Before looking at the deontic reading of belief’s correctness condition, let me briefly present axiological Normativism.

The axiological account explains correctness in terms of ‘good’, in the sense that *for any S, p: if S believes p, then that belief is good if p is true, and that belief is bad if p is false*. This version of TN makes room for the intuition that there is a value in having true beliefs whereas having false beliefs is, in some sense, bad. As we will see below, this approach squares nicely with the teleological view of belief, namely that by its very

³ Alternatively, it can be a semantic thesis concerning the meaning of intentional terms. I take this from Wedgwood (2009) who also argues that the semantic and conceptual theses could be true even if the corresponding metaphysical thesis is false.

essence beliefs aim at truth, and that such a goal is a valuable one. One worry, though, is that *correctness* and *good* seem to be very different properties, and it is doubtful they can be easily compared. We talk about ‘good’ students, not ‘correct’ ones: the two terms don’t seem to match that well. In addition, the axiological reading leaves it unclear how the value of truth individuates beliefs exclusively, and how it competes with other valuable goals one may have. Guesses may be valuable when true as much as beliefs, so it seems that interpreting TN in axiological terms fails to make it a unique, distinctive feature of belief. In addition, it is unclear how TN now competes with other values: beliefs that make me happy are also good as they satisfy my goals. So how can we distinguish the various senses of “goodness” of belief (Owen, 2003; McHugh, 2014; see also chapter 5)? As a result, this interpretation of TN does not seem to account for two key Normativist claims: that belief’s correctness is unique to it and is not comparable to other values or normative properties of belief.⁴

DEONTIC NORMATIVISM

TN is formulated by employing the term ‘correct’. But what kind of normative property is correctness? According to the deontic reading TN should read as follows (Wedgwood 2002, Boghossian 2003, Engel 2002, Shah 2003):

(TN) One ought to believe that p (if and) only if p is true

⁴ To address some of these worries, Wedgwood (2002) defends the claim that correctness is a basic normative concept in its own right. On this reading, TN is distinct from and irreducible to either deontic or axiological interpretations.

That is, for any S, p : if p is true then S ought to believe p, and if p is false then S ought not believe p.⁵

It is a natural claim to make that TN is part of the concept of belief (Velleman, 2000: 16; Boghossian, 2003; Adler, 2002; Gibbard, 2003; Shah, 2003; Shah and Velleman, 2005; Engel, 2013). As Boghossian (and others) puts it, “it’s a condition on understanding what it is for S to believe that p that one understand that S ought to believe that p only if p” (Boghossian, 2003: 40). Or: “Marco could not be said to understand what it is for Ebenezer to believe that Mallory reached the summit unless he understands that [...] Ebenezer ought to believe that Mallory reached the summit only if he did” (Boghossian, 2003: 38). That is, if Marco denies that having a belief entails being under a specific obligation that has to do with truth, either he is making a conceptual mistake, or he doesn’t know how to use the term ‘belief’.

I believe there are a few reasons why TN may seem very attractive. First, it is in the line with our practice to say that we ought to believe only true propositions, or that we have a duty to keep our beliefs in line with the truth. Not only do we blame people for their false beliefs, and usually praise true beliefs, but we also believe that that evaluation seems to have an effect on how people regulate their beliefs. As Nolfi (2013: 98) puts it, “the fact that a believer is capable of changing the way she regulates her beliefs in

⁵ Daniel Whiting (2010) employs the weaker notion of permissibility, defending the following version of TN:

TN) One is permitted to believe that p only if p is true

Engel (2013:208) argues that such a norm “is an *ideal of reason*, in the sense that it tells you what you ought to ideally believe [...] and thus it belongs to the category of the *ought-to-be* rather than to the category of the *ought-to-do*.” According to Sellars, “one ought to feel gratitude for benefits received, though feeling grateful is not something which one does. [...] One ought, however, to criticize (an action proper) oneself for not feeling gratitude and to take steps (again an action proper) to improve one’s character.” (SM, 76) The feeling of gratitude is a state in which one ought to be in such and such circumstances. If I have reasons to believe I am not feeling the appropriate gratitude in such and such situation, then there are things I ought-to-do.

response to and in a way that is guided by epistemic praise or criticism explains (at least in part) why she is an appropriate target of prescriptive evaluation for believing as she does.”

If that is so, then it should not come as a surprise, as Adler points out (2006), that we find the following Moorean claims problematic (see also Littlejohn, 2010):

- 1) p, but I don't believe that p
- 2) I believe that p, but I have evidence p is false

1&2 flout – and explicitly so – the biconditional truth norm of belief. Since such a norm is part of the concept of belief and we are committed to it, 1&2 are bound to strike us as odd, if not flat out absurd.

Second, TN may be able to explain how we are motivated to deliberate about beliefs. I will bring up this issue again in chapter 5, so now I will offer a brief description of it. Doxastic deliberation is a form of reasoning directed at deliberating about whether to believe p. Recently, Shah (2001; 2003; 2006; and Shah and Velleman, 2005) has argued that in doxastic deliberation the question *whether to believe that p* is transparent⁶ to the question *whether p is true*. This is how Shah puts it: “when asking oneself *whether to believe that p*, [one] must ... immediately recognize that this question is settled by the answer to the question *whether p is true*” (Shah 2003: 447).

For Shah, that means that truth immediately rules out all other (non-epistemic) considerations from having a motivating role in doxastic deliberation. As he puts it, “only

⁶ There are two distinct meanings of ‘transparency’. Among others, Evans and Moran (2001) talk about transparency, but use the term ‘transparency’ in relation to the first personal access we have to our own mental states. In contrast, Shah (2003) sees it as a feature of doxastic deliberation that has nothing to do with how we access our mental states. Therefore, the two meanings of ‘transparency’ should not be confused, and here I will only focus on the latter.

truth-regarding considerations move an agent in such [doxastic] deliberation” (2003: 468). The point here is that, even if non-epistemic considerations may at times appear in deliberation, at least as a matter of psychological impossibility, they can’t move us to form the belief in question as a result of deliberation. Only epistemic considerations, i.e. considerations that serve to answer the question *whether p is true*, can motivate one to form the belief that p. In addition, doxastic deliberation doesn’t always result in forming an actual belief but can, at times, conclude in suspension of judgment (or withholding of belief). And non-epistemic considerations don’t seem able to motivate us even to *withhold* beliefs as a result of deliberation either.⁷ That means that, when you ask yourself whether to believe that p, and you reach the conclusion that you have strong evidence that p is true, non-epistemic considerations cannot make it the case that you suspend judgment about p. All considerations that have nothing to do with the truth of p (i.e. pragmatic considerations) cannot motivate us in deliberation about whether to believe that p. This phenomenon is called, by Shah, the *transparency of doxastic deliberation*: in doxastic deliberation pragmatic considerations cannot motivate us, qua reasons, to form or withhold a belief that p. (But there are many other labels for it: doxastic involuntarism, exclusivity or, more simply, the fact that we can’t believe at will. In chapter 5 I will address it directly and call it, for simplicity, the “Explanandum”.)

It is, of course, possible that non-epistemic (i.e. pragmatic) considerations may cause the formation of a belief in non-deliberative contexts. What is special about the motivational role of epistemic considerations, though, is that in *deliberation* they can function as the *reasons for which* one forms a belief. It is notoriously hard to explain

⁷ Withholding of belief (or suspension of judgment) involves or implies the absence of the belief that p and the belief that not-p.

what *reasons for which* are, and I won't attempt at a definition here. Here Shah's own formulation: "R is a reason for which X ϕ s only if R is capable of disposing X towards ϕ ing through R's role in X's deliberation whether to ϕ " (2006:485). Another way to put this is to say that R is a reason for which X ϕ s only if R is capable of *consciously* and *directly* motivating R to ϕ when X deliberates on whether to ϕ and uses R as a premise in such deliberation. Thus, pragmatic considerations can play a role in motivating you to form an *intention* to engage in all the necessary actions that could make it the case that you end up forming a belief; but in deliberation there is no direct motivational link between consciously entertaining a pragmatic consideration as a premise and the formation (or suspension) of a belief.

Theorists usually appeal to intuitions which, they claim, strongly indicate that any consideration that appears to us as having nothing to do with the truth of a given proposition 'p' is incapable of playing the role of motivating us to form/withhold the belief that p as a result of deliberation.⁸ Shah (2003: 2006, and Velleman, 2005) offers an inference to the best explanation for transparency. He argues that the best explanation for transparency is that belief is a normative concept. A "normative concept" is a concept that contains/expresses/entails a norm/evaluation.⁹ In any deliberative belief-forming process we apply the concept of belief. The concept of belief is a normative concept which contains the norm TN. Now, Shah has a specific view about what is required to

⁸ For defenders of transparency (or exclusivity) see, for instance, Archer (2015), Kelly (2002), Steglich-Petersen (2009). Also Bernard Williams points out that we don't seem to be able to believe something "just like that" and as a result of the fact that we want to believe it (1970: 108). For worries about the claim that transparency is a genuine phenomenon see McHugh (2013a), Sharadin (2016), and Zalabardo (2010).

⁹ We need to distinguish between thin and thick normative concepts. Thin concepts are concepts like GOOD and RIGHT which only express or contain evaluations. Paradigmatic examples of thick concepts are various quality-concepts such as COURAGE, CRUELTY, TRUTHFULNESS, KINDNESS and many others. But also MURDER is a thick concept: it has both a descriptive content and a normative component. BELIEF is also a thick normative concept for Shah.

have a normative concept, such as, BELIEF. He says that having a normative concept means understanding and *accepting* its normative component. For Shah accepting a norm means that one is necessarily motivated by that norm.¹⁰ Not only that: for Shah acceptance is a dispositional state to conform to a norm N, not follow norms incompatible with N, or assess others' and one's own performances based on whether they conform or not to that norm (Shah, 2003: 467). Thus, in the case of belief, accepting the truth norm means having a disposition to conform to it, not following norms incompatible with it, and assessing others' and your beliefs in relation to such a norm.

On this view, the conceptual nature of TN explains the fact that when we apply the concept belief in deliberation, we also necessarily apply only belief's truth norm to our belief-forming process. That is, any time we deliberately form a belief, we ask ourselves whether we should *believe* that p. As a result, we necessarily employ the concept of belief and thus, since we accept TN, we are motivated to apply TN and *no other norm incompatible with that*.¹¹ As transparency does not occur outside the deliberative context when one does not apply the concept of belief (Shah, 2003: 467), it follows that transparency is "something that is demanded by the nature of first-personal

¹⁰ This is an expressivist understanding of normative concepts based, I believe, on Gibbard's account (1992; 2003). Take the concept/term LEWD. On Gibbard's proposal, when someone uses the term "lewd" referred to x she means that a certain attitude (called it L-censoriousness) toward x is warranted because x passed the limits of sexual display. So applying the concept/term LEWD has some motivational implication: when you apply it you accept some norm. Gibbard says: "I myself am not a licensed user of the term [lewd] because I do not find it reasonable to elaborate standards of warrant for feelings of L-censoriousness toward sexual displays. I understand the feeling, but I do not think that standards of warrant apply to it" (Gibbard, 1992: 281).

¹¹ I take it that, roughly, fact E is evidence for hypothesis H if E is a sign or reliable indicator that H is true even though it can be defeated by evidence to the contrary. Shah is careful to point out that, to get transparency in deliberation, we also need to have some understanding of what type of consideration counts as evidence for the truth of some proposition. If someone were totally confused about what sort of thing counts as evidence, that confusion could undermine transparency. Clearly, this is compatible with the possibility that one may be mistaken about what specifically counts as evidence for what.

doxastic deliberation” (Shah, 2003: 447). That means that if there were an agent for which the deliberative question about belief was not transparent to truth, that agent would be applying a different norm rather than the truth norm, and thus she would not be deliberating about belief but about some other attitude. The ability of TN to explain transparency is one of the key strengths of Normativism about belief.

There is a third reason why TN enjoys so much support: the fact that TN is the constitutive, conceptual norm of belief purportedly explains why epistemic norms apply to belief. More specifically, TN is used to explain the source of epistemic normativity and the necessary or inescapable authority of epistemic norms, such as:

EN) One ought to believe that p if and only if one has sufficient evidence that p.

EN) If one believes that p and believes that if p then q, then one ought to believe that q.

If the Normativist is right, beliefs are subject to those and other epistemic norms simply by virtue of being the sorts of mental states that they are, namely mental states constituted by the truth norm. In addition, the Normativist argues that, because of the role such a norm has vis-à-vis any other non-constitutive norm that may apply to belief, the inescapable authority of epistemic norms follows from the inescapable authority of TN (Wedgwood, 2002).

The advantages of endorsing Normativism do not stop here. To be sure Intentional Normativism is distinct from Content Normativism, namely the claim that the content of a mental state is normative. The relation between the two views is indeed a matter of debate. However, at least according to Boghossian, the normativity of content derives from the fact that belief is normative (Boghossian 2003). That means that introducing TN as the constitutive norm of belief would kill two birds with one stone for

the Intentional Normativist: not only it would account for belief and its relation with truth, but it would also offer an explanation for the normativity of content, leading to a full blown normative picture of the mind.

Finally, one may wonder what the link is between the debate on the aim of belief and the question about (normative) reasons for belief. One may argue that strict evidentialism, i.e. the normative claim that only evidence can be a reason for belief, seems to follow from the normative nature of belief. Indeed, it may seem reasonable to say that if belief aims at truth and TN is constitutive of belief, then a consideration counts as a reason to believe only if it is an epistemic consideration. As a reply, pragmatists will either try to deny that belief aims at truth or insist that, even if TN is the constitutive norm of belief, it is not immediately clear why it would exclude non-epistemic reasons, i.e. reasons that depend on belief's relation to the agent's ends, whether cognitive or not. The Normativist has a ready reply to that last point, though. In a recent paper Shah draws the consequences of his position on transparency and the concept of belief, and argues that from that, we can derive the strict evidentialist claim that only evidence can be reason for belief (for a skeptical look at this see Steglich-Petersen, 2008).¹² In a nutshell his argument is that if we cannot but be motivated by and only by evidence concerning the truth of *p*, then pragmatic considerations can never play the role of motivating reasons. Now, if we combine this observation with the argument concerning the 'deliberative constraint of reasons' – the claim that a consideration can be a reason for

¹² A different tack is to argue for evidentialism independently of the aim of belief. In his "The Ethics of Belief", Clifford presents an argument for the claim that "It is wrong always, everywhere, and for anyone to believe anything on insufficient evidence." In his view, it seems that we have an epistemic obligation to have beliefs that are sufficiently supported. In turn, we have a moral requirement to satisfy our epistemic obligations. Evidentialism is thus established on moral grounds on this view.

someone to F only if it is capable of being a reason for which she Fs¹³ – we get that only evidence can be reason to believe. Hence evidentialism. If right, all that would add even more grist to the Normativist's mill.

PROBLEMS WITH DEONTIC NORMATIVISM

In the literature on this topic, Shah's explanation of belief's transparency has been shown to be quite problematic for various reasons. Recall that Shah maintains that in deliberation, only truth-related considerations can function as reasons for which one believes. Nevertheless does the application of a concept in deliberation imply that I necessarily commit myself *only* to the norm that it contains? Shah thinks so. However, the controversial point here is that to deliberate about belief is not just to ask whether to believe that p would be correct (i.e. true), but whether I *have most reasons* to believe that p. Now, it seems that the mere fact that a norm N is constitutive of a concept of Φ does not mean that in deliberating whether to Φ I can't take into account considerations that are in contrast with N. Consider that when I deliberate whether to assert or promise something, I apply the relevant concepts and thus acknowledge their constitutive norms. I acknowledge that my promise to Φ is correct only if I intend to Φ . However, when deliberating whether to promise I might also take into account considerations opposed to the constitutive norm of promising. Similarly, for assertion, I can deliberate to assert something I know to be false – and thus contrary to assertion's constitutive requirement – for moral or prudential reasons (Steglich-Petersen, 2006). Why is that not the case for belief?

¹³ Shah (2006) attributes the deliberative constraint to Bernard Williams (1981) and argues for its plausibility in his paper.

To explain why this happens, an analogy between belief and chess is often invoked. It is constitutive of being a chess-player that one tries to win over one's opponent. Similarly, *it is constitutive of being a believer* that one cares, above all, about the truth of one's beliefs. This is because there is a norm that governs respectively chess and belief; for chess, it is the norm that says that correct, good, or successful chess-playing means checkmating your opponent. Likewise, a belief is correct only when true. Intentionally flouting these norms and thus no longer caring about what makes chess or belief correct, means failing to be a believer or chess player. This is why chess-players must care about checkmating their opponent if they want to count as chess-players to begin with, and so each time they make a move, they have such a goal at the forefront of their mind. Analogously, we – believers – care about truth when we form a belief. It is constitutive of being a believer to do so. Unfortunately, though, the fact that I care about the truth of my beliefs in deliberating about them does *not* solve the question of whether I *ought* to care. Indeed, why should we care about being believers in the first place? Why should we make ourselves subject to TN and form beliefs rather than smeliefs (Enoch, 2006)? The Normativist story does not offer an answer to this question.

A second, unrelated worry against TN is: explaining transparency by referring to TN requires subscribing to a form of motivational internalism in which accepting the truth norm means being necessarily motivated by it. However, this already commits Shah to assuming an internalist position not everyone may be comfortable with. In addition, his view requires that to explain transparency the truth norm is able to silence any norm that opposes it when applied in deliberation. That is, Shah does not deny that other norms beyond the truth norm may apply to belief. To explain transparency, though, he is

committed to saying that when those norms are in opposition to the truth norm, they will be trumped by it. Again, though, it is unclear why we should accept that the truth norm has such a strong role (McHugh, 2013). Finally, as Steglich-Petersen (2006: 507) puts it, “[i]t is doubtful whether a consideration which necessitates motivation should be considered a normative consideration at all.” If deliberation about belief exhibiting transparency is a result of applying the norm of truth, then necessarily such a norm cannot be violated. However, it seems a requirement of following a norm that one has the freedom to violate it. Thus, the inescapability of transparency casts doubts on the very existence of the truth norm (for a reply to this see Engel, 2013).¹⁴

A different kind of objection against deontological TN has to do with the impossibility of being guided by it (Glüer & Wikforss, 2009: 44-45). According to this objection, TN can’t really guide agents in the formation and revision of their beliefs because, in forming a belief about *p*, one can conform to TN only by first forming beliefs about whether *p* is true. That means that the formation or revision of the belief that *p* does not come *because* of the norm TN has been followed, but actually arrives prior to following that very norm. Thus, TN does not provide guidance, and its existence does not make any difference vis-à-vis belief’s creation, maintenance, or revision. That threatens one of the assumptions for adopting TN: namely, that it does provide guidance and regulate beliefs.

A common Normativist reply to this is that even if TN does not provide any direct guidance, it can still guide belief regulation indirectly via some subjective norm of rationality, either logical or epistemic (Boghossian, 2003; Wedgwood, 2002). The

¹⁴ For other objections to Shah see, in particular, Buleandra (2009), Steglich-Petersen (2006; 2008), Zalabardo (2010), and McHugh (2013). Nolfi (2016) for an original take on this issue.

subjective oughts are grounded in the objective ought of TN. Glüer & Wikforss (2013: 87-89) reply that that particular route opens the following dilemma for the Normativist: either these subjective norms are derived from TN because they are reliable indicators of truth but then they are contingent norms; or they are not contingent, but then they do not derive from TN. If those norms are contingent, then it is possible that in a different world, other norms may apply to belief. They become purely instrumental norms. Still, many Normativists are wary of saying that rationality is purely instrumental. However, if epistemic norms are constitutive and not instrumental, then they can't derive from or depend on the truth-norm because there are situations in which these norms diverge. For instance, in evil demon scenarios the truth norm and epistemic norms would give rise to contradictory oughts (for a reply see Wedgwood, 2013).

It gets worse for the Normativist. The claim that TN is a norm that can guide belief-regulation also faces the following concern: no formulation of TN seems to offer a plausible description of our belief-revising practice. In particular, Bykvist and Hattiangadi (2007) argue that of the several possible formulations of TN, none of them seems to be a plausible candidate for the truth norm of belief. For simplicity, I will consider the most common one:

TN: For any S, p: S ought to believe that p if and only if p

Notoriously, the problem with this formulation is that it requires that one believes every truth there is, even those truths that are trivial or uninteresting. In addition, Bykvist and Hattiangadi raise the worry that TN forces you to believe blindspot propositions because they are true, and yet if you believe them they will become necessarily false. As a result, for those propositions, TN is impossible to follow (for a reply see Whiting,

2010). Things don't get better if we eliminate the injunctive half of TN because now, on this formulation, if something is true one is under no requirement to believe it, even if one has evidence for it.¹⁵ On top of this, TN faces "ought-implies-can" worries. The jury is still out on whether one has voluntary control over one's beliefs. If one does not have voluntary control over one's beliefs, though, it is plausibly said that belief is *not* subject to deontic evaluation. Consequently, any deontic formulation of TN is simply false (Alston, 1988; against this argument see, for instance, Steup, 2000).

Finally, Bykvist and Hattiangadi (2013: 103-4) present an objection to any Normativist formulation of the truth norm. They argue that it is unclear why 'correctness' need be understood as a normative term at all. As they claim, "to view 'correct' as an essentially normative term does not square well with common usage. For example, it does not fit well with the fact that 'correct' is standardly used to mean either 'in accordance with acknowledged or conventional.'" Mozart's C major sonata just is a piece of music that is correct only if certain notes are played in a certain order (Rosen, 2001: 619). Here 'correct' means that there is a certain standard that is individuating this particular piece of music, and yet 'correct' here is *not* a normative term. That means that standards of correctness can be constitutive of a practice without being normative: they give rise to no requirements to conform to them. In the context of belief, Bykvist and Hattiangadi (2013: 4) suggest, "it seems natural to assume that [...] 'correct' just means 'in accordance with truth' or, simply, 'true'." That is true for guesses as well: we say that they are correct when true. And yet we do not *require* them to be true. This seems to

¹⁵ For other problematic formulations of TN see Bykvist and Hattiangadi (2013).

undermine the Normativist story that, in the case of beliefs, we should understand ‘correct’ as a normative term (see also Dretske, 2001: 247).

CONSTITUTIVISM

Normativism argues that belief is a mental attitude constituted by a norm. Although, as we just saw, the standard view is that the norm of belief is the truth norm, other norms are possible: the knowledge norm (Engel 2004), an evidential norm (Adler 2006), norms of rationality (Zangwill 2005), and norm of justification (Gibbons, 2013). Each of these norms, though, is seen as a *constitutive norm of belief*. Unfortunately, we are rarely told what that is supposed to mean.

Non-constitutive norms are norms for a type of action, or practice, that exists independently of the norms itself. Rules of etiquette are a clear example of that: they regulate some practice which, however, exists independently of them.¹⁶ In contrast, constitutive norms “create” the very practices, or activities, they regulate. In the literature, games are the primary examples with which the notion of constitutive norm is used. Games are activities with explicitly defined rules concerning the permissible moves within the game itself (Dummett 1959, Searle 1969, Williamson 2000). To bring this into focus, we need also to distinguish *two* different ways in which the label ‘constitutive norm’ has been used in the literature. It starts with the distinction between constitutive norms and regulative norms. Regulative standards prescribe conducts within a certain game, whereas constitutive ones explain how to engage in those games. Consider the

¹⁶ Papineau (2013: 66) has recently sketched a non-constitutivist position on the norm of belief: “I do not of course wish to deny that there are some valid prescriptions that apply to the adoption of beliefs. However [...], I shall maintain that all such prescriptions arise from considerations of moral value, or personal value, or possibly aesthetic value, and not from any distinct species of doxastic value.”

following example. I am playing basketball and at one point while throwing the ball I commit a foul. My act was improper because it violated a rule of basketball. However, my act still counted as *playing* basketball. Let's take an example with speech acts. I promised to wash the dishes, but haven't the least intention of keeping the promise. Here I have promised all right, but the act is improper because it is not sincere. My act breaks a regulative rule, i.e. the rule of sincerity. In such cases, even though I did perform a speech act (i.e. a promise), I did it *improperly*: I violated a regulation.

Turning now to constitutive norms, we said that there are two senses in which a norm can be constitutive. According to Searle a constitutive rule/norm for an activity A is a rule/norm you *must* obey in order to partake in activity A. This is the first conception of "constitutive" that we see in the literature. Similarly, for Rawls the notion of constitutive norms makes sense of the idea that for a practice nothing counts as part of the practice if it is "outside the stage-setting" set by the constitutive norms of that practice (Rawls, 1995: 25).

There is also a second conception of constitutive norm which is the notion of 'constitutivity' that is most important to us here. On this view a constitutive norm of, say, a game is a *regulative* standard that defines that game. As such, it can be violated without automatically failing to play that game. And yet if players repeatedly (and overtly) violate constitutive norms of a game, that plausibly means that they actually stopped playing that game, and that they are now playing a different game. Regulative, constitutive norms can thus be violated, but only up to a certain point.¹⁷

¹⁷ Importantly, there is a difference between intentionally and overtly failing to conform to norm N and being a deceiver/cheater. Someone who intentionally wants to deceive others represents herself as being correctly applying N. The deceiver thus intentionally and covertly breaks N while representing herself as following it.

All this is, unfortunately, quite vague. The most pressing question for us is: what distinguishes a constitutive regulative norm from a simply regulative norm? Here is a first pass definition: it is usually said that a constitutive norm governs an activity *necessarily*. Going back to belief then, TN is a constitutive norm of belief only if it is a necessary norm of belief. However, being a necessary norm is not enough for being a constitutive norm. An important feature of any constitutive norm is that it is non-derivative: constitutive norms, so to speak, derive from the nature of an attitude or an activity ϕ , they don't come from some general principles. A constitutive norm of ϕ is thus a norm that is part of the nature or essence of ϕ . Still and all, not all necessary properties are also essential properties: a property may be necessary of ϕ without itself being an *essential* property of ϕ . Famously, Aristotle thought that being able to laugh is a necessary property of human beings, but not an essential one. Similarly, one may worry that the truth norm, although it applies to belief necessarily, it does so as a result of a more general principle (e.g. a moral requirement). So if Normativism is right and the norm TN is an essential property of belief then it seems that TN can't *just* be a necessary norm of belief: it must be part of its essence (Steglich-Petersen, 2008: 277-278).

It is notoriously hard to say what essential properties are, though. A way to try to solve this is to require that TN is the *only* necessary norm of belief. The thought here is that a necessary norm is constitutive/ essential of ϕ only if it is the only necessary norm of ϕ . Value monism seems indeed to be a widespread assumption when talking about constitutive norms.

The problem with this tack, though, is that it is too restrictive as it does not apply well to other cases of practices that are norm-constituted. Consider the following case. It

is commonly thought that knowledge is the constitutive normative requirement of proper assertion. This has been formulated in terms of the Knowledge Rule of Assertion (KR): “one must: assert p only if one knows that p” (Williamson, 2000: 243). Now, assuming that knowing entails believing, then also believing that p is a normative requirement for properly asserting that p. Accordingly, we can formulate a belief norm for assertion: (B) You must: assert that p only if you believe that p. In that KR is a constitutive norm of assertion, then it governs assertion necessarily. It follows that if KR entails B, then also B is a necessary rule of assertion. Then, however, contrary to the prevailing wisdom, KR is not unique and – according to the suggested restriction – non-constitutive.

To address this, we can try adding the following element to the definition: in order for a necessary norm N to be constitutive, N has to be basic. N is a *basic* norm of ϕ only if any other necessary norm of ϕ is reducible to it. Therefore, TN is constitutive of belief only if it is a basic, necessary norm of belief. This works well for the Normativist: indeed it seems plausible that the truth norm is basic with respect to other necessary norms of belief such as the norms of rationality and evidential norms (Wedgwood, 2013; for worries about this strategy see Glüer & Wikforss, 2013: 87-8). Finally, a constitutive norm of ϕ is also individuating when it governs only ϕ as its constitutive norm. As a result, TN is not individuating unless it is the constitutive norm *only* of belief.¹⁸

Is this a good definition of constitutivity? Can it explain in what sense TN is essential or part of the nature of belief? Here is a worry: the fact that TN is a basic, individuating, necessary norm doesn't *per se* establish that TN is *essential* to belief, i.e. it's part of its nature/essence. As Paul Horwich writes (1998: 188-193):

¹⁸ It is unclear that the truth norm is individuating of belief. Arguably, some speech acts, like assertion, guess, conjecture, are *also* governed by the truth norm as their constitutive norm. And what's more, the mental state of perceiving is constitutively governed by the truth norm.

[A] property may have normative implications without itself being normative. [...] It might be possible for our most basic normative principles to have the form $\neg (x) (Dx \rightarrow Nx)$ – where “D” describes some state of the world and “N” specifies what ought or ought not to be done in that situation. Thus the normative implications [...] leave it entirely open that its nature is completely non-normative.

The ‘essence’ of killing isn’t normative because we can explain it in entirely non-normative terms; and yet killing has necessary normative implications that derive only because of some extrinsic normative principles that apply to it.

This worries the Normativist. Assume one could offer a description of belief in non-normative terms. Assume such a description could specify some state D of belief that has some normative implications TN such that TN would be necessary and basic while applying only to belief. To make this possibility more concrete, imagine that TN (or any other basic and necessary norm of belief) could be the result of the role of belief in reasoning (which may be described as a purely descriptive property of belief) and the goal or norm of reasoning. In this case, TN would *not* be part of the essence or nature of belief, and this is a real threat for the Normativist project.

As a reply to this worry, the Normativist could try to argue that TN is a conceptual truth: it is part of the concept of belief that TN governs it. However, as I mentioned above, this strategy has raised serious worries. And even if we think of beliefs as states that we ought or ought not to be in under such-and-such circumstances, the nature or essence of those states may not depend on their conceptual status. So even if BELIEF is a normative concept, the nature of belief itself may not be normative at all.

Alternatively, the Normativist could simply deny that we can offer an account of belief in descriptive terms, so we need to appeal to a necessary, basic norm to explain

what belief is (Velleman, 2000; Shah and Velleman, 2005).¹⁹ I will, however, argue against this reply in chapter 2 and again in chapter 5. Though I will not go as far as arguing that belief's role in reasoning is purely descriptive and has normative implications, it is also true that the Normativist does not possess any convincing positive argument for why TN should be regarded as the constitutive norm of belief.

TELEOLOGY: WHEN THE AIM OF BELIEF IS AN ACTUAL AIM

Some theorists don't read Williams's claim that belief aims at truth as a metaphor. They actually believe that beliefs or believers want the truth, and this is because of the nature of belief. Still, the teleological and normative views of belief and its aim are fully compatible, and hybrid views have been offered. Some theorists maintain that the teleological nature of belief explains its constitutive norm: belief is correct iff when true because only true beliefs satisfy the constitutive aim of belief (Velleman, 2000). This view is associated with an axiological reading of the norm of belief, and here 'correct' means good (Steglich-Petersen, 2006). Alternatively, one may defend both the teleological and normative approach as to belief but deny that the norm of belief derives from its constitutive aim (Shah and Velleman, 2005).

The hybrid view notwithstanding, however, the most interesting position to adopt, dialectically speaking, is to see the teleology in opposition to the normative view. In this dissertation, ergo, I will treat teleology as an alternative to the normative view. An attractive feature of the teleological approach, when detached from the normative, is that

¹⁹ Wedgwood makes a slightly different point when he says that explaining the nature of belief requires that one appeals to concepts such as 'ought', 'should', 'correct', 'right', 'wrong', 'rational', etc., and that these are considered the "paradigmatic normative terms" (Wedgwood, 2009b). Whether or not he is right about that, it remains to be seen whether this is really an argument for Normativism: as Rey points out, one can use such terms to refer to natural laws and so without any normative implications (Rey, 2007: 76).

it fits well with a naturalistic account of belief because goals and functions can be accounted for in naturalistic terms. This also leads to an instrumentalist (and thus naturalistically accepted) view of epistemic normativity and epistemic rationality. That is, belief aims at truth, and does so by its very nature. Epistemic rationality, or having the support of epistemic reasons, is what secures that goal, is the means to the end-goal of truth. That being the case, we care for epistemic rationality because of our quest for truth.²⁰

There are two compatible ways to explain belief's aim for the teleologist. Some teleologists take that belief's aim is a goal or function of the system that produces beliefs. This function could be a biological one, emerging out of the survival benefits of having true beliefs (Millikan, 1984; Sullivan-Bissett, 2017). Others see belief's aim to be a personal goal of the believer, akin to an intention to have only true beliefs and, possibly, to believe what's true (similarly Sosa, 2011 has argued belief is a kind of performance, directed at a certain aim).

David Velleman combines both these views arguing that the essential feature of belief is the function of the mechanism that regulates it. And so to be a belief is to be brought about by a mechanism whose function is to track the truth. This 'mechanism' could be either some sub-personal feature of the mind or a conscious intention of the believer. This intention emerges especially when the believer engages in doxastic deliberation. Thus, belief aims (subpersonally) at truth in the sense that belief is necessarily influenced, formed, and caused by some truth-tracking mechanism. Also, belief aims at truth in the sense that, in forming a belief, the believer always has the goal

²⁰ Berker (2013) and Kelly (2003) for arguments against these epistemic consequentialist approaches.

to form a belief that is true (Steglich-Petersen, 2006). As it happens, a belief-producing mechanism – whether it's an intention or a subdoxastic mechanism – may fail and create a belief that is false and irresponsive to evidence. In this case, such a belief-producing mechanism has failed to fulfill its function (Velleman, 2000:254-255).

I believe it is very intuitive to say that truth is by and large desirable and that our cognitive systems may have evolved or have the function to track it. As I will mention in chapter 1, though, there is now a strong body of evidence that shows that humans often do not conform to principles of reasoning and don't evaluate arguments based on those principles (Rey, 2007: 75). That is, people don't seem to be very good at spotting bad reasoning and generally their reasoning seems to be defective. This constant deviation from the norm is, according to some, not merely a problem of performance but the result of an underlying competence-problem (Stein, 1996). This has led some to become skeptical of the idea that we are truly after truth after all (Stich, 1990).

Perhaps less controversial is the claim that agents generally aim at having true beliefs. We want truth, it seems, and our deliberative processes of belief formation are guided by this goal and intention. On the face of it, this seems to square nicely with the fact that we are not in full control of our beliefs, and only evidence really motivates us in doxastic deliberation. And yet I am skeptical that we aim at truth. Of course, by and large we do think that truth is a good thing, and we definitively expect our beliefs to be true because of their role as premises in reasoning. However, I doubt that we *necessarily* aim at truth when deliberatively forming beliefs. Thus, as I argue in Chapter 5, the teleologist account of belief is severely misguided (see also Owen, 2003), and we need a better story of what happens in deliberation.

THE DISSERTATION: RETHINKING THE AIM OF BELIEF

In the dissertation, I argue that belief has a necessary aim. I interpret that aim as a goal of the believer (not necessary an intention, but a desire that may be unconscious). I argue that the believers' goal is epistemic justification. Such a goal is not part of the nature of belief, but is still necessary to form beliefs in deliberation. The nature of belief, in my view, is to play a specific role in reasoning. From that, and from the necessary goal of reasoning, I derive that in doxastic deliberation we aim at forming sufficiently justified beliefs. This positive thesis has some negative upshots for both the Normativist and the Teleologist view: although it is possible that various normative requirements govern beliefs, they are not necessarily linked to truth, and they are not constitutive of belief. As I argue in chapter 2, there is no need to define belief based on its constitutive goal or norm of truth. In addition, in chapter 5, I show that the Teleologist's attempt to explain what happens in doxastic deliberation is doomed to fail. Combining this with the other worries the normativist and teleologist views face, I conclude that Williams' claim that belief aims at truth needs to be now seriously reconsidered.

Here is a summary of all the chapters:

(Chapter 1) Defending Full Belief. Here I defend the claim that we use full beliefs in reasoning against those who think that full beliefs should be reduced to partial beliefs (or eliminated all together) and that full beliefs cannot rationalize action. In contrast, I offer reasons to believe that we do reason using full beliefs, and that belief-based reasoning is a *heuristic reasoning method* that is less cognitively demanding than

reasoning with partial beliefs. I also show that, even though full belief-based reasoning may at times deliver sub-optimal choices, it still has the power to rationalize action.

(Chapter 2) Belief and its Consequences. In this chapter, I argue for the claim that belief has a unique role in our mental lives. To do so I attack the recent claim that the output role of belief, its ability to influence action and other attitudes, is shared by other cognitive attitudes such as (propositional) imagination, acceptance and supposition. My argument comes in various steps. Initially, I defend the claim that *only* belief (and desire) can motivate action. Second, even assuming that other cognitive attitudes can motivate action, that is not enough to show that their consequences are the same as belief's. Belief is an 'unqualified' cognitive attitude, whereas all these other attitudes *must* come with some kind of constraint on their use that limits the type of inferences or reasoning patterns they can enter into. This has two main upshots: for belief, the possibility to be defined in terms of its unique role in inference and reasoning is still open. And we do not need to refer to belief's relation with truth to uniquely individuate belief.

(Chapter 3) The Aim of Reasoning. Proposed here is a new constitutive account of reasoning. Reasoning is an inferential process that usually starts with some mental attitudes (e.g. beliefs, emotions, and desires) as premises and ends with dropping, forming, or withholding another mental attitude. On my view, reasoning is also a process guided by the aim of arriving at a conclusion supported by at least sufficient reasons. This aim is constitutive of reasoning in the sense that it is necessary for reasoning, and makes reasoning into a unique type of inferential process. I show that the strongest proposal currently on the market, i.e. the rule-following view, is insufficient to make sense of reasoning (Boghossian, 2014; Broome, 2013).

(Chapter 4) Reasons vs. Correctness: When is Reasoning Good? In this second to last chapter I argue again for the Reason View: reasoning is the process of attitudes revision that aims at conclusion supported by at least sufficient reasons. Here I argue for the Reason View by expanding on some of the arguments presented in the previous chapter. In particular, I will show that the Reason View can make sense of what we mean by ‘good reasoning’ and that the Reason View is more plausible than its immediate competitor, i.e. the Correctness View: the view that purports that reasoning aims at forming attitudes that are correct.

(Chapter 5) What Guides Belief. Finally, I take up the puzzling phenomenon often called ‘Doxastic Involuntarism’, i.e. the fact that, psychologically, we can’t believe at will. First, I clarify what that exactly means, and then I demonstrate that what guides belief-formation is the constitutive aim of having beliefs supported by sufficient *epistemic* reasons, and this is the result of the general aim we have when reasoning toward attitudes as well as the unique role belief has in securing that goal. In other words, it is metaphysically impossible to come to believe what we want by means of reasoning: doing so is simply contrary to the very existence of reasoning. Hence, the limited power we have over our beliefs is actually a condition of possibility for being the type of rational agents we are.

DEFENDING FULL BELIEF

Chapter 1

The goal of this chapter is to argue for the following cluster of claims: (1) at times we do reason *as if* the propositions we use as premises were true (even if we are not certain they are true); (2) this way of reasoning can produce rational outputs (given certain conditions), and (3) this is our default way of reasoning; it follows that, (4) believing is a *default* disposition to use a premise as 1 (true) in reasoning. In the next chapter, I will offer more arguments for the claim that belief is a default premise by distinguishing belief from other (full) attitudes that present their content as true (e.g. assumption, supposition, imagination).

This is the plan: first, I present a non-reductionist picture according to which full beliefs are metaphysically different from partial beliefs (or credences). Then, I focus on reasoning and support the view that full belief is a disposition to use a proposition as true in reasoning. My next step is then to show that that full belief-based reasoning is a heuristic reasoning method that is less cognitively demanding and possibly more efficient than reasoning with credence. Now, since generally our mind uses systems of reasoning that are fast and frugal as default, and it also *rational* to use the more efficient and cost effective means to reach our goals, the natural hypothesis that arises is that belief-based reasoning is our default strategy of reasoning, i.e. a strategy of reasoning we use in normal conditions.

CLAIM

The aim of this chapter is to argue for the following claim:

CLAIM: part of the functional role of full belief is to be a default disposition to treat a proposition *p* as *true* in reasoning.²¹

It is common in philosophy to describe full beliefs by at least referring to the role they are disposed to play in reasoning. As Scanlon puts it, to believe that *p* “involves recognizing it as having the status of something that is to be relied on in further theoretical reasoning [...], and to be relied on as a premise in practical reasoning.” (Scanlon, 2007: 104) Littlejohn maintains that “what beliefs are supposed to do is represent how things are so that we might rely on them for the purpose of deliberation.” (2012: 238) Similarly, Williamson claims that “one believes *p* outright when one is prepared to use *p* as a premise in practical reasoning.” (2000: 99). Fantl and McGrath claim that if you believe that *p* “then you are prepared to put *p* to work as a basis for what you do, believe, etc.” (2010: 143). Weatherson adopts “the functionalist idea that to believe that *p* is to treat *p* as true for the purposes of practical reasoning” (2005: 421). Accordingly, “[...] if a person consistently refused to rely on *p* as a premise, and rejected arguments relying on it, then it would be plausible to say that he did not really believe that *p*.” (Scanlon, 2007: 91)

Relatedly, it is often said that believing that *p* entails being disposed to use ‘*p*’ as a premise in one’s reasoning *while assigning value 1 to ‘p’*. When we treat a proposition as true, we exclude, in our reasoning, the possibility that it is false. We don’t use it, that is, as *probably* true: we are simply not interested in what would happen if the proposition in question were false. At times, we do that because we are in fact certain of the truth of that

²¹ CLAIM is far from being a new position in philosophy, so my aim here is simply to give new strength to that position. The way I do that is by combining views on the nature of belief offered by Frankish (2004), Ross and Schroeder (2014) and Wedgwood (2012). However, what I say in this chapter represents a new take on the issue because, first, it can’t be reduced to any of those views in particular, and second, it is supported by arguments that have not been offered by those authors to argue for their own views.

proposition. Most of the time, though, we believe propositions of which we are not certain.

Despite being very popular, CLAIM has been recently criticized in various ways. In particular, some have argued that full beliefs should be reduced to credences, or that beliefs don't play a role in reasoning and, if they do, they can't rationalize it anyway. In contrast, I plan to argue for the plausibility of CLAIM²². Before doing that, though, I need to make some preliminary remarks.

For starters, note that belief's disposition to show up in reasoning means that, *if* a belief were explicitly represented and its content available to the subject then it would, other things being equal, take part in reasoning. So the mere fact that a proposition is *implicitly* believed does not make it ready to motivate and enter in reasoning. Implicit beliefs are disposed to become part of reasoning only if they become explicit. Here a rough formulation of the implicit/ explicit distinction: explicit beliefs are propositions that are contained in the so-called "belief box". In contrast, implicit beliefs are beliefs whose content is not contained in the belief box but that is easily or swiftly derivable from the other explicit beliefs the agent has (See Dennett, 1978). Explicit belief is a state with such and such representational content and the functional role constitutive of beliefs. Presumably then, with implicit beliefs, there is no such state one is in. Rather, one could easily or swiftly come to be in such a state.²³

I talk of '**dispositions**' so it is important to briefly clarify what I mean. I intend dispositions to be functional states of the cognitive system. Thus, beliefs are states of the

²² In the next chapter I take up a different objection to CLAIM that says that full belief is not the only attitude that can be defined as a default disposition to treat a proposition *p* as true in reasoning.

²³ Pinning down exactly what explicit/implicit means proves to be quite difficult (see Field, 2000). In particular, there is an intuitive distinction to be made between implicit beliefs and beliefs one might quickly adopt in the right circumstances

mind which, when occurrent, play a causal role in the processes leading to action (and to the formation of other mental states), and action provides at best good but defeasible evidence for their existence. Dispositions here are *ceteris paribus*: dispositions may not manifest and when so it is because the situation is in some way abnormal. For belief, in this chapter I argue that belief is a disposition to use a proposition as true premise in reasoning by default, i.e. in normal contexts unless some additional information shows up.

Now **reasoning**. I will have a lot to say about reasoning later in the dissertation but here is a rough explanation of it. I intend reasoning to be a process. By “reasoning process” I mean a process of thinking that proceeds from premises to conclusions (or the other way around) and in which the subject takes that there is a rational link between the premises and the conclusions. Thus, on this view a reasoning process occurs only when the subject is aware of the premises and the conclusion and takes the conclusion to rationally follow from the premises. I assume ‘rationality’ to apply fairly broadly as to include the possibility that what is rational to do is determined by a calculation of expected utilities, or that what is rationally to (partly) believe is determined by axioms of probability. And by ‘aware’ I mean that their content is available for action and report.

By this I don’t mean that when going from premises to conclusion, the subject is actively and intentionally employing some rule of reasoning (e.g. modus ponens, inference to the best explanation). Also, I don’t intend to say that the subject needs to form a belief that there is a rational link between premises and conclusion. I mean something much weaker, namely that, if asked or challenged, the subject is in a position to mention the premises of her reasoning as supporting the conclusion. In addition to that,

I can grant that those reasons may not actually be what caused that particular conclusion. Here is what I mean. In the famous ‘Linda test’, after reading a personality sketch about a woman called Linda, subjects were asked to assess the probabilities that Linda was a feminist activist *and* a bank teller vs. only a bank teller. In this case, it is said that most of the subjects reach their answers by using heuristic reasoning for which they ignore the starting point. They don’t know they are using a heuristic and they don’t know what they are basing their reasoning on. However, if asked to justify their answer they are likely to come up with one. That is, they are able to offer reasons why they believe what they do, even if these did not actually motivate them to form that belief in the first place. My point is that their coming up with reasons for their conclusion is the outcome of a thinking process called ‘reasoning’.²⁴

Note that by talking about ‘premises’ and ‘conclusion’ I don’t mean to say that the only type of reasoning possible is inferential reasoning. As I explain below, I assume that, in the practical realm, reasoning includes also decision-theoretical decision making processes. What defines reasoning – in any of its forms – is the commitment of the agent to taking her decision as rational given other attitudes she has.

Finally, I don’t mean reasoning to require that the subject intentionally sets out to perform a bit of reasoning and arrive at a conclusion as reasoning can happen rather spontaneously in the sense that I can spontaneously (i.e. no conscious effort/decision in doing it) arrive at a conclusion given some beliefs I have.

What reasoning is not, I claim, is the processing that takes place, for instance, in the visual module. The output of such a module is a representation which I am usually aware

²⁴ This is similar to Mercier and Sperber’s (2011: 58) definition of “reasoning proper”. They write, “[w]hat characterizes reasoning proper is indeed the awareness not just of a conclusion but of an argument that justifies accepting that conclusion.”

of, namely a representation usually available for report and for action. However, in this case I am not aware of the premises of the inferential process that produced such a representation. What's more, the following is no reasoning: the production of beliefs and intentions based on processes that use premises we ignore. Sometimes this kind of processes is called 'intuition' or System 1, although these labels are often used to mean different things.

I assume that many of our attitudes are the result of such a thinking process that does not count as reasoning in my sense. I am actually happy to concede that this is the way in which we often come to form beliefs and intentions and that reasoning happens only in certain circumstances. All I care here is to show that our beliefs are at least dispositions to use a proposition as premise *when* we engage in reasoning. When we are not reasoning (in my sense), such a disposition will not be activated or manifested.²⁵

Here is a possible worry concerning my definition of reasoning. It seems that, by restricting the definition of belief to a form of reasoning that requires some sort of metacognition and possibly language, I am implicitly stating that such a definition does not cover doxastic states of animals and very young children. And the worry is that either my definition is too restrictive and cannot really account for what beliefs are or I am committing myself to the controversial claim that animals and babies don't have beliefs.

I'd like to remain agnostic concerning the questions of whether or not animals and babies have *full* beliefs. Possibly, only adults and children with some metacognitive

²⁵ Keith Frankish (2004) has recently argued that system 1 runs on partial beliefs or credences rather than full beliefs. I'd like to remain agnostic on whether this type of cognitive process makes use of, among other attitudes, credences and/or full-beliefs. What I don't necessarily subscribe to is Frankish's claim that partial beliefs have a place only in unconscious, System 1, reasoning.

capacities have full beliefs. This is a possibility that I can't explore here.²⁶ Alternatively, one could argue that animals and children have beliefs and that CLAIM actually points to a subset of beliefs which only adults and older children have. This is in fact what I am inclined to subscribe to. de Sousa's "assent" (1971), Dennett's "opinion" (1978), Sperber's "reflective belief" (2000, 1996), Rey's avowal (1988) and Frankish's "superbelief" (2004) all seem to point in the direction of an attitude expressed in language, involved in reasoning as a high-level activity, and present only in adult humans. A somewhat analogous distinction is drawn by Sosa (2015) when he sets belief apart from judgment. Also, Brandom's (1994) equating assertion with commitment suggests that we have a high level, language-based doxastic attitude that is distinct from low level doxastic dispositions.

Also, as I will mention again in chapter 3, the metacognitive capacities that are required for reasoning may actually be very minimal. More specifically, one may not need to actually *metarepresent* one's beliefs when reasoning, but only have their content available for report and be able to use it when, for instance, the conclusion that those beliefs support is challenged. This latter ability may itself not require the concept of 'reason' either. That is, taking a premise to support a conclusion may only require a fairly basic understanding of what counts as a reason. For such a basic understanding all is needed is sensitivity to the normative standard of belief and intention, i.e. being able to reliably distinguish correct vs. incorrect performances. It is possible that even young children may have that kind of 'sensitivity' too. If a child is able to assess the validity of a claim or belief based on its causes, we may be in a position to grant her an implicit

²⁶ See Frankish (2004) for a possible way to develop this idea.

understanding of reasons. More concretely, if a child reliably spots normatively appropriate vs. inappropriate beliefs by accepting or refusing them, this may be an indication she knows what norm governs beliefs and claims and what counts as a (good) reason for believing or claiming. If this suggestion is on the right track, then perhaps we can venture the following hypothesis: 3- and 4- year old children do have an implicit grasp of what a reason for belief is. And so they *may* be able to engage in a form of reasoning that requires this kind of implicit grasp. Of course, this is only a suggestion at this point. But if it is correct, then the notion of reasoning I am adopting here may not be so restrictive as it originally appeared to be and can be extended, if not to animals, at least to small children.

A final point. What I am trying to do here is to vindicate a notion of belief that belongs to folk-psychology. However, one may worry that a mature cognitive science will eventually get rid of folk-psychological notions all together. For instance, Jeffrey (1970: 71-72) argues for the eliminativist position that, strictly speaking, full beliefs do not really exist. Here is a famous passage:

By 'belief' I mean the thing that goes along with valuation in decision-making: degree-of-belief, or subjective probability, or personal probability, or grade of credence. I do not care what you call it because I can tell you what it is, and how to measure it, within limits...Nor am I disturbed by the fact that our ordinary notion of belief is only vestigially present in the notion of degree of belief. I am inclined to think Ramsey sucked the marrow out of the ordinary notion, and used it to nourish a more adequate view.

However, as Jerry Fodor (1990: 156) has put it: if such eliminativism is true, then “practically everything [we] believe about anything is false and it’s the end of the world.”

My hope is that my argument will make the possibility of elimination look less attractive, and support the claim that at least those creatures who engage in reasoning must have full beliefs or something very *similar* to that (in light of the suggestions that whatever scientific theory cognitive science will end up being, there will have to be some “non-trivial coordination” with the folk psychological model of the mind (Godfrey-Smith, 2004: 157)).

ASSUMPTION

CLAIM naturally leads to the following metaphysical picture of the mind:²⁷

ASSUMPTION: full beliefs are metaphysically different from partial beliefs (or credences) and can't be reduce to them.

Full beliefs are all-or-nothing attitudes in the sense that they are not graded. For instance, either one believes in some proposition *p* or the belief is absent. All-or-nothing doxastic attitudes come in three forms: belief in '*p*', disbelief in '*p*', or suspension of judgment about '*p*'.²⁸

In contrast, credences or partial beliefs are attitudes that come in degrees. Talk in terms of levels of credence seems very plausible. Indeed, it seems intuitive to many that we generally talk of some beliefs as being held with a greater degree of confidence than others and that when we use introspection to assess what we believe we notice that we have much more confidence in some things we believe than in others. So, for example, I

²⁷ It is not impossible to subscribe to CLAIM and refuse ASSUMPTION by maintaining that full beliefs are reducible to credences. Possibly, Frankish (2004) holds this view.

²⁸ Since disbelief in '*p*' could be seen as a belief in not-*p*, it is usually said that this model of belief is actually a binary one. And some people reduce suspended judgment to belief in the sense that suspension of judgment about *p* is nothing but the absence of belief that *p* and disbelief (which is belief in not-*p*). That means that belief is really the only full doxastic attitude. Nothing I say here hangs on this point.

am much more confident that the Earth revolves around the Sun than that the Universe started with a Big Bang. This point has led some philosophers to postulate the existence of a doxastic attitude called “partial/graded belief”, or “credence”. Credences come in degrees: some propositions we are absolutely certain are true and others we are certain are false. Then there are propositions we take to be true or false to a degree. And since it is usually thought that one can assign numerical value to credences, we can say that credences vary between 1 (certainty that the claim is true) and 0 (certainty that the claim is false).

One common way to quantify degrees of belief is to associate them with a relative disposition to place a bet: one’s credence in p is the amount of money one is willing to pay for a bet that yields \$1 if p is the case and \$0 if not (Ramsey, 1926; Savage, 1954). So one can measure the degree of one’s belief in p in terms of how much one is willing to bet.²⁹ Credences can thus be defined as attitudes that dispose the agent to bet based on her confidence that something is the case. To illustrate: I am very confident that my official first name is Marianna. High confidence does not require certainty, though. Although I am pretty sure my name is Marianna, I am not willing to bet \$1 that that is in fact my name but only, say, \$0.9. Accordingly, if we were to assign a numerical value to confidence, then my degree of belief that my name is Marianna would be 0.9 whereas my degree of belief that my name is *not* Marianna would be 0.1.

One way to read ASSUMPTION is that full beliefs and credences are two metaphysically distinctive psychological states. This is a controversial claim as many philosophers tend to adopt a reductivist view of full belief and credence according to

²⁹ The relation between degrees of beliefs and betting disposition may not be metaphysical but only a causal connection that holds in ideal situations. However, see Christensen (2004) for an argument against this.

which one of the two states is derivative. One way to argue that credences or degrees of belief are derivable from full beliefs is to say that when we talk of degrees of belief what we are really referring to are flat-out beliefs about probabilities. Here probabilities are intended as chances relative to an implied fixed background. So, for example, when we say that someone is 50 per cent confident that a coin toss will come up heads, what we mean is that they flat-out believe that, if tossed enough number of times, there is 50% chance the coin will come up heads. (Or perhaps: if tossed enough times, it will come up heads 50% of the time). Alternatively, chances can be interpreted as evidential probabilities. For example, one might believe that there is an 80 % chance that I will have pancakes this morning. This is a belief about the relationship between one's total evidence and the occurrence of an event.³⁰

Gilbert Harman (1986) has proposed a different type of reduction of credence to facts about full belief. On his view degrees of beliefs are emergent properties, an “epiphenomenon resulting from the operation of rules of revision.” (Harman, 1986: 22) And the fact that beliefs can be held strongly or weakly is a matter of how hard it would be for us to give them up. Harman (1986: 22) writes, “[...] it may be that P is believed more strongly than Q if it would be harder to stop believing P than to stop believing Q, perhaps because it would require more of a revision of one's view to stop believing P than to stop believing Q.” Degrees of beliefs are thus explained in terms of how strongly or weakly we hold a full belief and thus how hard it is for us to give it up. Importantly, on this view, degrees of belief are usually not explicitly represented. Of course, that is not to deny that we can have full beliefs about probabilities. However, having full beliefs about

³⁰ The standard objection to this reductivist account is that it seems overly demanding to expect that to have graded beliefs one needs to have and employ concepts like CHANCES and PROBABILITIES.

probabilities is not necessary to have degrees of belief, at least according to Harman. And it may not be sufficient either: I believe it is possible that I may assign a degree of probability to an event that does not correspond to the actual confidence I have about the truth of that event. Presumably, confidence, i.e. (on Harman's view) how strongly I hold a belief and how easily I would give it up, may not be a luminous condition. Accordingly, there may be cases in which I fail to assess, or fail to assess accurately, my degree of confidence in some proposition I believe.

I am somewhat skeptical about reducing credences to full belief. But my goal here is to make room for the idea that we do reason with full beliefs, so I don't have to assess any proposal of reduction that goes the other way.³¹ I will only focus on attacking the position that takes credence or confidence as basic and full belief as derivative. On this view, the only doxastic attitude that has metaphysical status is credence.³² This view comes in two flavors: the Certainty view and the Threshold view. The Certainty view says that believing is assigning probability 1 to a proposition. Since this is a reductivist view, it says that one can simply reduce full belief to certainty. As such, however, it is quite problematic: it seems just false that when we fully believe we are also certain of what we believe. And if this reduction were on the right track, then we would have to

³¹ It seems that only an on-off model of belief can fit with philosophical analyses of knowledge which is usually taken to be justified true belief plus something else. Credences, in contrast, are not assessable in terms of truth and falsity (although they can be assessed in terms of how close/far to truth they are). So it seems that they can't be an element that constitutes knowledge. However, Moss (2013) has argued that credences can constitute knowledge. I will not consider this issue here. I won't discuss implications ASSUMPTION may have on various epistemological views.

³² Sturgeon (2008) and Richard Foley (2009) argue for the metaphysical primacy of credence and offer a reduction of full belief to credence. Unfortunately, I can't do justice to the complexity of this debate here. For instance, I can't evaluate Christensen's proposal (2004: 98-100) that binary belief is merely degreed belief that counts as "high" within a context.

conclude that we in fact have very few full beliefs (which is contrary to what we intuitively think).

If belief cannot be reduced to credence 1, then the alternative proposal is the Threshold View: there is a threshold t such that a rational agent believes p if and only if $cr(p) \geq t$. Here the threshold is both vague and contextually variable.

Notoriously, the Threshold View has problems with the following closure principles:

Deductive closure: Rationality requires of S that if S 's total belief set logically entails some proposition p , then S believes p

Conjunction closure: Rationality requires of S that if S believes that p and believes that q , she believes their conjunctions.³³

If one thinks of credence as basic, and of full belief as reducible to credence, one will naturally expect the closure constraint to fail. A classical example, often used in the literature, is the Lottery paradox: given that you have high credence that your ticket will lose, you believe that it will lose (Kyburg, 1961). By the same token, you also believe that to be true for everyone else's ticket. But you also have high credence that one ticket will win and you believe it. However, since you rationally believe that each individual ticket will lose, given you follow the conjunction principle, you also rationally believe that all tickets will lose. Another example is the Preface Paradox (Makinson, 1965): if high credence is enough for full belief, then it is plausible that when writing a book an author rationally believes each of the assertions in her book. However, the author also takes herself to be fallible and thus rationally believes that the conjunction of all her assertions in the book is false. This is problematic if one holds the view that one is

³³ These principles are sometimes stated in permissive rather than obligatory terms. This difference won't matter here.

rationally required to believe the conjunction of the propositions one believes. To deal with this, some Threshold theorists are willing to bite the bullet, reject the conjunction principle and thus allow that a rational agent can have each of the “lottery beliefs” and believe each of claims in her book *without* holding their contradictory conjunction (Christensen, 2004). I won’t assess the plausibility of this reply here, but this strategy is important for my next point.³⁴

There is another quite pressing worry against the Threshold view, a worry that is extendible to any reductivist view (no matter the direction of the reduction). For simplicity, I’ll call it the ‘motivation objection’. There are two ways to voice this objection. One way focuses on the idea that, if we want to keep talking about full belief we need to identify the *distinctive* causal role it plays, a causal role that can’t be reduced or explained away by credences. Otherwise, the objection goes, why don’t we just drop the belief-talk altogether? A key premise of this argument is that there is nothing else that could justify keeping the idea of belief alive other than it having a *distinctive* motivational role. So such an objection takes the following structure: (i) belief’s defining element is its motivational role.³⁵ (ii) The Threshold view says that the distinctive

³⁴ For a different objection to the Threshold view see Friedman (2013) who argues for the rationality of suspending judgment on high-credence propositions.

³⁵ Mark Kaplan (1996) offered a view that may constitute a reply to (i). On Kaplan’s view, belief is not a disposition to act in a certain way but a disposition to make a sincere assertion. Here is how he puts it: “You count as believing [i.e. believing flat-out] *P* just if, were your sole aim to assert the truth (as it pertains to *P*), and your only options were to assert that *P*, assert that $\sim P$ or make neither assertion, you would prefer to assert that *P*.” (Kaplan 1996: 109) The problem with this definition of belief is that it does not account for irrational beliefs. Imagine I have a belief about my great abilities at tennis and imagine that I also I have no sufficient reasons to believe that. Regardless of that, I am happy to have such a belief and have no interest in discovering how accurate it is. However, if my aim is to speak the truth, I’d prefer to avoid asserting that I am a great tennis player and for the reason I alluded above: I am aware I have no strong enough reason to take my belief to be true.

element of belief can be fully captured in terms of credences. (iii) On that view, the very notion of belief becomes idle.

The Threshold view says that full belief kicks in once a certain threshold has been reached. Importantly, though, nothing really changes in that case because, as a matter of fact, we continue to reason based on our level of credence. That is, having a full belief does not change anything in what we are disposed to come to believe/do or not believe/do as a consequence of reasoning. As pointed out by Keith Frankish, on the Threshold view the role of full belief in relation to action is idle. Beliefs are motivationally idle, he claims, because they “possess just as much causal power as the states of confidence in which they consist” (Frankish, 2004: 64). On the Threshold view, the causal power of belief is nothing above and beyond the causal power of the level of credence we are in when we have a full belief. What’s more, beliefs have the same causal power of states of confidence that are close enough to the ones they consist in. Assume the threshold for belief is 0.8 credence. Presumably, in many cases a state of confidence of 0.8 has the same causal power of a state of confidence 0.75. That means that, from the causal point of view, whether or not I fully believe a proposition is actually irrelevant. Nothing that wasn’t already there gets to be added once I reach the level of full belief.

However, the objection concludes, we usually take beliefs to be “explanatory salient psychological states” (Frankish, 2004: 64). We cite them when explaining action and we use them to predict what people will do. A theory of belief should try to account for that. Otherwise, one is left to wonder why we should keep the idea of full belief at all. Why don’t we just drop the talk of full beliefs and refer only to degrees of confidence (Stalnaker, 1984: 91)?

I mentioned above that there are two ways to spell out the motivational worry against the Threshold view. I explained the first way, and now here is the second. The second motivational worry against the Threshold view is that folk-psychology indicates that there is in fact a distinctive causal role that belief plays, and reducing belief to credence fails to capture that. So the reduction can't be done. One way to argue for this is the following. Beliefs are dispositions to conform to certain principles of reasoning, conjunction closure being one of them. If this is so, then reasoners are *disposed* to have conjunctively closed beliefs. That is, when we believe two propositions we are, all else being equal, disposed to believe their conjunction.³⁶ However, on a Threshold account, unless one has credence *one* in all the conjuncts, one is disposed to assign lower probability to a conjunction than to any of the conjuncts that make the conjunction. But the Threshold view also maintains that below a certain threshold of confidence one does not have a full belief. That means that I may believe two or more propositions but fail to believe their conjunction.

According to the folk conception of belief there must be something amiss with a view of full beliefs that does not take into account the following: people are disposed to believe the conjunction of propositions they believe and, when they don't believe, it is because of some conflicting dispositions.³⁷ (Something similar could be said for other principles of reasoning like deductive closure.) Similarly, it seems natural that our folk-psychological conception of belief is that one *ought* to believe the conjunctions of the

³⁶ Wedgewood (2012), Frankish (2004), and Kaplan agree that beliefs have such dispositions. Christensen (2004, chap. 4) and Sturgeon (2008, section 5) deny that beliefs are so disposed. See Kolodny (2008) for an argument on why this dispositional view is a version of the Myth of Coherence.

³⁷ Or, perhaps, the fact that they have not considered the question at all.

propositions one believes. But this is not the case for credences. So again, something seems amiss.

There is a way to reply to this that – *contra* folk-psychology – simply denies that full beliefs have such dispositions. This reply can appeal to the fact that there is now a strong body of evidence that shows that humans often do not conform to principles of reasoning and don't evaluate arguments based on those principles (Rey, 2007: 75). That is, people don't seem to be very good at spotting bad reasoning³⁸ and generally their reasoning seems to be defective. This constant deviation from the norm is, according to some, not merely a problem of performance but the result of an underlying competence-problem (Stein, 1996).

I believe that the evidence of the “heuristics and bias” view of reasoning is very strong and don't intend to deny that humans can be faulty reasoners. However, there may be good reasons to resist the stronger claim that humans are *intrinsically* irrational. It is true that there is strong evidence that humans reason based on heuristics that lead to biases rather than using principles of ‘correct’ reasoning. Heuristics are mental shortcuts we deploy in various contexts to arrive at various conclusions. The reason we deploy them, consciously or not, has to do with the fact that heuristics are fast and frugal tools to form beliefs and make decisions. Importantly, though, even if we agree that we are prone to using heuristics, the evidence we have does not establish that we *always* employ heuristics unless we are tutored to use correct reasoning. Indeed, the move from

³⁸ Here is a standard example of this deviation from rationality: in a series of experiments students were asked to evaluate whether the following argument is deductively valid: ‘All roses are flowers. Some flowers fade quickly. Therefore some roses fade quickly.’ In the experiment a very large majority of students wrongly deemed this argument to be valid. The explanation for this is that they adopted a heuristic: first they noticed that the conclusion is true and then moved from the conclusion being true to thinking that the argument is valid.

‘evidence indicates that we often use heuristics’ to ‘evidence indicates that we always do it’ seems problematic.³⁹

What’s more, besides being a problematic generalization, it seems that there are positive reasons to resist accepting a skeptical conclusion about human reasoning competence. There is evidence that, although in many situations we are indeed bad reasoners, there are contexts in which we do much better. That is, there is evidence that, when we use reasoning in dialogical or argumentative settings, performance improves substantially. For instance, people often employ *modus tollens* reasoning in argumentative contexts (Pennington & Hastie 1993:155). Results in the Wason selection task get a lot better when people exchange ideas in a group setting. Studies have shown that we are fairly good at spotting argumentative fallacies (e.g. slippery slope) and react to them appropriately (by refusing to endorse their conclusion) (Hahn & Oaksford 2007; Neuman 2003; Neuman et al. 2006; Weinstock et al. 2004). That means that we are able to produce good reasoning when set in the right context even without any tutoring.

If confirmed, this body of evidence may prove useful to lessen the skepticism some have about human reasoning. The strategy is to locate the source of error *not* in our reasoning competence, but in the performance affected by external conditions.⁴⁰ Accordingly, failures of rationality can be explained by the fact that subjects may not reason well when placed in contexts not conducive to good reasoning. Now, contexts are

³⁹ As Gigerenzer has put it, sometimes the heuristic and bias camps seem to believe that “the untutored mind is running on shoddy software, that is, on programs that work only with a handful of heuristics” (1991: 235). As Samuel and al. (2002) point out, though, “[a]t most, what could be plausibly claimed is that we have reason to think that, in many instances, human beings use normatively defective heuristics. The further claim that these normatively problematic heuristics are the only cognitive tools that untutored folk have available is vastly stronger than anything the available evidence will support.”

⁴⁰ This strategy is common to those who try to argue against the idea that humans are fundamentally irrational.

‘right’ for reasoning when no countervailing force is present such that prevents reasoning and thus belief from following its due course. These countervailing forces may be tiredness, sloppiness, and a ‘desire’ to cut corners and save energies and resources. So, on this reading, in contexts that are not ‘right’, our disposition to reason well is blocked by opposing forces. And this would give support to the idea that beliefs are indeed disposed to conform to principles of reasoning *other things being equal*. And as we saw, this would be a problem for the Threshold view.⁴¹

And even if the evidence provided may not be enough to show that beliefs are disposed to conform to principles of reasoning, still it seems that – when the right *motivation* is in place – beliefs do conform to these principles. Traditionally, psychologists and philosophers alike believe that the point of reasoning is to discover the truth and enhance knowledge. In contrast, it has been recently suggested that the function of reasoning is argumentative. That is, on this view reasoning abilities are geared toward producing arguments and justifications for our claims, beliefs, and actions. Notably, though, reasoning is not only designed to produce arguments. It also has the function of monitoring them. More specifically, the idea is that there are mechanisms used to filter the information coming from communication in order to avoid misinformation. Such mechanisms check the reliability of the communicated information by looking at the source of the information, at the content of the information, and at the processes through which it was formed. This could explain why, when we are in the context of communication and argumentation, we are much better at spotting bad arguments: we use

⁴¹ I am inclined to support this stronger reading and in the next pages I will assume it to be true. One reason for my support is that it seems to me hard to explain how we can make sense of each other’s behavior if we accept the hypothesis that we are fundamentally irrational. Unfortunately, I will not offer an argument for that here. For more arguments against attributing intrinsic irrationality to humans see Cohen (1981) and Davidson (2004).

our reasoning abilities to check whether the information sent is reliable or not, whether it aligns with what we know, whether it forces us to revise some of the things we believe, and so on (Mercier & Sperber, 2011).

To achieve the goal of convincing others (who may not be easily persuaded) it is usually better to have good arguments. More specifically, if those we communicate with are able to spot potentially bad arguments and our goal is to convince them, the best way to do that is to try to have arguments that could pass the test. This would explain why at times we see conformity to principles of reasoning: when we are in argumentative context, we are motivated to reason well.

As a result, we can't exclude that there may be cases in which beliefs will behave in a way that is not always parallel to how credence would - when rational - behave. That means that it is possible that beliefs may follow their own motivational course, which they don't necessarily share with credences. So *if* it is true that beliefs at times show an independent motivational role, then it seems that one can't reduce beliefs to credences above a threshold. If this is correct, then they must be two metaphysically distinct attitudes which is what the Threshold view denies.

Now, of course, this is highly speculative at this point and it is doubtful that evidence from the reasoning-literature can be used to directly support the irreducibility of full belief to credence. That is, I don't think I have provided any definitive argument against the Threshold view. My main goal was only to argue for the plausibility of ASSUMPTION, and indicate that the possibility of reduction may be less attractive than initially thought.

Before ending this section, I will briefly mention an objection that has been moved against ASSUMPTION by Scott Sturgeon (2008: 47; 148), an advocate for the Threshold view, who argued that a view like ASSUMPTION can't explain why credence and full belief seem to go hand in hand when producing actions:

Whenever someone goes to the fridge, say, because they believe that it contains beer, there is a clear and everyday sense in which they go to the fridge because they are confident that it contains beer. [...] Coarse and fine belief yield everyday action in harmony, marching in step throughout everyday practice. This cries out for explanation; and it does so in spades on the Divide- &-Conquer approach [read ASSUMPTION]. After all, that approach has it that confidence and binary belief are quite different things. But then it's surprising that each marches in step with the other as a source of everyday action. Why on earth should that be? Why should strong confidence go with binary belief in the production of ordinary acts; and vice versa? The Divide-&-Conquer strategy has no internal resource to answer this question.

And again:

The Threshold View prompts the natural idea that coarse and fine belief march in step as the causal source of action because coarse belief is nothing but sufficient confidence. If that were so, coarse and fine belief would causally march in step just as they seem to in practice—they would generate action in parallel; for that is how causal powers of determinable and determinate relate to one another. This strongly suggests that the Threshold View is on the right track.

As a reply, let me point out that the fact that ASSUMPTION has no “internal resources” to explain why high confidence and belief march together in everyday action may not be a problem for ASSUMPTION. That is, there is in fact a plausible way to make sense of this phenomenon: beliefs and high credences tend to occur together because they are sensitive to or caused by the same kind of input. More specifically, they are both sensitive to the evidence available to the agent who has them. So it is not very

mysterious why they tend to show up simultaneously. Also, as I will explain below, full beliefs are employed in reasoning patterns that approximate Bayesian ideals, the same ideals credences strive for. As a result, it is not mysterious why they usually march together as causes of actions.

Another way of developing Sturgeon's worry is to ask: Why do we have these different kinds of states at all, if they march hand-in-hand? Why do we need both? What do we gain by having both? What are the distinct functions such that we use up resources maintaining both full beliefs and credences? In the next sections, I will try to offer an answer to these questions. The bottom line will be: we need both because, although they often lead to similar outcomes, they actually require different level of cognitive sophistications and, depending on the situation, they may lead to divergent actions.

PROBLEMS WITH BAYESIAN RATIONALITY AND MODELING

If there are both credences and full beliefs in the mind, what is their role? Bayesianism has provided a powerful account of the role of credences in reasoning. My goal in this chapter is to convince the skeptical reader that we need to take also the role of full belief in reasoning seriously. In this section, I will mention some worries Bayesianism has to face. In particular the last point I will make here will open the doors for an account of belief based on its role in reasoning.

The Bayesian account of rationality has two components: one component is an account of theoretical rationality; and the other component is an account of practical rationality. For theoretical rationality, probability calculus is the standard tool for calculating ideally rational degrees of belief in the sense that an ideally rational agent's

degrees of belief must be probabilistically coherent and thus obey the laws of probabilities.⁴² When it comes to practical reasoning, Bayesianism offers a specific model of rational choice which is the decision theoretic model. The norm of decision theory in its modern form is: when choosing among acts one should choose the act with the highest value of expected utilities. For example, when one is deciding whether to bring an umbrella to work or leave it at home, one considers the utility of getting wet, of staying dry while not carrying an umbrella, and of staying dry while carrying the umbrella, as well as one's credence in rain vs. not-rain. This reasoning captures the structure of the considerations involved in instrumental or means-ends reasoning. In typical cases, an agent faces a choice among means that lead to different competing ends, which she values to different degrees. And, in typical cases, none of the means available to the agent will lead with certainty to some particular end. So the agent's judgment about what to do must be sensitive both to judgments about which ends she cares about and how much, and to the likely result of each of her possible actions. That means that it is rational to choose an act only if it maximizes expected utility with respect to the confidence we have in certain outcomes and the value of those outcomes in relation to the choice we could make.⁴³

⁴² Degrees of confidence have epistemic rational constraints: coherence and conditionalization. Subjectivist Bayesianism maintains that incoherent beliefs are subjected to a Dutch book, in the sense that there is a finite set of bets the agent is willing to accept that (subjectively) guarantees the agent will lose money. Similarly, a partial belief is irrational and thus subject to a Dutch book if its degrees don't change in proportion to the probability of the propositions learnt by the believer (Ramsey (1926) and de Finetti (1937)). A different view, called Probabilism, does not accept conditionalization. I take that the Dutch Book argument is a pragmatic argument *for* the claim that ideally rational degrees of belief conform to the probability calculus. However, it is a matter of debate whether the Dutch book argument is just a pragmatic argument or whether it is also epistemically valid.

⁴³ This conception of decision theory depends on the idea that it is beyond the agent's control whether a state of nature obtains. In addition, one should be able to establish exactly the value of each available act if

There are various motivations to adopt this view. Bayesian decision theory (BDT) is a theory that, in the words of David Lewis (1981: 5), is “simple, elegant, powerful, and conceptually economical.” What’s more this theory is used in microeconomics, and so also enjoys some empirical support. Finally, it has been argued that violating BDT means having irrational preferences: if one does not maximize expected utilities one will end up with preferences that are intransitive or violating the sure-thing principle (Savage, 1954).

Still, there are numerous worries that afflict the Bayesian model. I won’t press those worries too much, but it is important to mention them.

First, it is a question about rationality. That is, are Bayesian standards really the ideal of rationality? I will tackle this question again (in relation to full-belief reasoning) later in the chapter. For now, let me explain what the issue amounts to. The issue is not whether Bayesian probability or decision theory is a good standard of rationality in the ideal sense. The pressing question is whether it is a good tool to apply to the human mind. Here is what I mean. Bayesianism tends to be concerned with modeling reasoning processes of certain kinds of ideal agents, setting the standard at the computational level. Assuming that humans can reach those levels of rationality, the pressing question is whether those standards are rational for humans, and whether they would be really useful when implemented by humans to resolve the challenges posed by the environment. This is important because Bayesian standards have been adopted to constrain a theory at the algorithmic level, which means that whatever theory we may have about mechanisms/algorithms the brain uses, it will have to be constrained by the goal of

it were performed in that state of nature. Note, however, that it is possible to construct decision theory without appealing to the idea of state of nature. I will not consider this possibility here, though.

Bayesian rationality which is set at the computational level (at least for “methodological” Bayesians following Marr (1982). See also Bower & Davis, 2012).⁴⁴

The problem seems to be that Bayesianism sets the standards of rationality so high that it is perhaps impossible for us ordinary humans to meet them. For instance, Bayesian Decision Theory seems to operate under the assumption that in any decision-making situation the agent is simply given the options from which she is to choose. Bayesians don't say what, according to them, justifies this feature. But in reality the conditions in which an agent is confronted with “the whole set of alternatives from which he will choose his action” is limited to only a small set of decisions we make (Simon, 1983: 22; Douven, 2002). More commonly, when deciding what to do, searching for options is one of the most important aspects of deliberation. Hence, Bayesian standards are good for ideal agents, not for us. (However, in the Bayesian camp there have been some attempts at ‘de-idealized’ Decision Theory so that its normative requirements could be more easily met.)⁴⁵

If so, then it is not obvious that Bayesian standards should be adopted to constrain a theory at the algorithmic level, or at least not always. Perhaps other models should inform the computational level, models that are not in accordance with Bayesian rationality. As a result, that should have some effect on the constraint Bayesian models can impose on the algorithmic level.⁴⁶

⁴⁴ As Bower & Davis (2012: 393) point out “this approach is in no way committed to the claim that the mind and brain compute in a Bayesian-like way at the algorithmic level.”

⁴⁵ For instance, Paul Weirich (2004) developed a Bayesian decision theory compatible with various human cognitive limitations.

⁴⁶ Some have argued for a notion of bounded rationality which roughly says that it is better or adaptably more efficient to not be fully Bayesianly rational. The assumption behind this is that bounded rationality and Bayesian rationality often ‘recommend’ different behaviors, so that doing X may conform to bounded rationality but not to Bayesian rationality.

Second. There are worries about the fact that people are *not* approximating Bayesian ideals, even assuming these are the correct standards of rationality. It seems human credences are in fact not in accord with what Bayesian rationality dictates. For instance, according to Bayesians we should be assigning probability 1 to all logical truths in order to be rational and this sounds like a task we fail at. This goes back to the worry about humans being intrinsically irrational I mentioned before. As before, the fact that we fail at a task does not rule out that it is possible to describe the computational level in Bayesian terms. But perhaps, if we are too distant from that ideal, then it could be said that we really lack the necessary competence to become rational in the Bayesian sense.

Third. A further worry has to do with the computational – the actual calculation – complexity of the Bayesian model. For instance, to calculate expected utilities one may need to engage in some complicated math which most people can't do and certainly can't do fast. Actually, there is some skepticism that finite agents like us are really in a position to make those calculations at all. What's more, Harman (1986: 25-26) points out that, if we reasoned with degrees of belief, we would have to make extensive use of the conditionalization rule, which can be a process of considerable mathematical complexity. As he says it, “[o]ne can use conditionalization to get a new probability for P only if one has assigned a prior probability not only to E [the evidence proposition], but to P & E. If one is to be prepared for various possible conditionalizations, then for every proposition P one wants to update, one must already have assigned probabilities to various conjunctions of P together with one or more of the possible evidence propositions and/or their denials.” The problem with this, he claims, is that it leads to a “combinatory

explosion” that finite agents like us would not be in a position to compute. Harman concludes that humans can’t reason with partial beliefs.

One common response Bayesians in cognitive science make to critics (who say it’s not computationally feasible etc.) is that using heuristics can be seen as an implementation at the algorithmic level of the Bayesian ideal, one that approximates the Bayesian ideal more-or-less, at least in ecologically relevant circumstances.⁴⁷ On this view, Bayesian models really focus on predictions rather than on how and whether Bayesian reasoning *actually* happens. That is, Bayesian models (of various kinds) are supposedly good predictive tools, but lack power to say what goes on in our psychology (Jones & Love, 2011).

Four and most importantly. The jury is still out on whether Bayesianism is a good model for what goes on in the mind/brain (at the algorithmic level). One fair question, however, is whether these processes involving credences are actually *reasoning* proper, as I called it above. Perhaps probabilities are used in learning mechanisms, behavior-production mechanisms of various kinds, but not in reasoning. If, as I will try to show in the following chapters, reasoning proper is a process that requires that the agent takes the premises to support the conclusion (i.e. to be ‘reasons’ in the general sense of the term), then we still lack an account of whether this reasoning can happen with credences.

In the literature on Bayesianism, the question of whether credences play a role in real agents’ reasoning processes does not seem to be often directly addressed. Recently, Staffel (2013) has offered an attempt to model reasoning with credences. More specifically, she has argued that “degrees of belief, just like outright beliefs, can function

⁴⁷ See, for example, Griffiths et al. reply to Bower and Davis in *Psychological Bulletin* 2012; Sanborn et al. 2010. Importantly, these heuristics are qualitatively different from the heuristics the bounded rationality approach usually refers to, which are for the most part non-probabilistic in nature (Gigerenzer et al., 1999).

as attitudes that we reason *from* and attitudes we reason *to*.” (3537) However, what she means by ‘reasoning’ is a fairly low level process (akin to System 1), so I think her view won’t do for my purpose here. Frankish has also dealt with the same question in an interesting way. In relation to practical reasoning (2004: 31) Frankish points out that, “our *conscious* reasoning very rarely takes a Bayesian form. We generally prefer to reason from unqualified premises to unqualified conclusions, employing classical inference schemata, such as the practical syllogism.” His point is that, in general, we don’t seem to reason (in the proper sense) by employing strategies that resemble Decision Theory. In contrast, Frankish’s proposal is that high level reasoning (which he equates to System 2) is done with full beliefs; low level, unconscious reasoning (System 1 for him), operates with credences. I am sympathetic to his proposal, and in light of this my goal in the remainder of the chapter is to show that we reason with full beliefs, that we do so rationally, and that this is likely to be our default *reasoning* strategy for making decisions. This is, of course, compatible with the possibility that low-level (and some high-level) reasoning processes are computed with partial beliefs.

FULL BELIEF AND DECISION-MAKING

In this section of the paper I will argue for CLAIM by showing that, given how we reason, make decisions and bring about actions, there must be an attitude that has the dispositional role that CLAIM attributes to belief. Saying that there is no such attitude or that belief’s dispositional role is not captured by CLAIM, means leaving unexplained how it is possible that we reason the way we do. Or so I will argue.

Ross and Schroeder (2014: 267) recently pointed out that,

[f]or some propositions we have] a defeasible or default disposition to treat them as true in our reasoning—a disposition that can be overridden under circumstances where the cost of mistakenly acting as if these propositions are true is particularly salient. [...] But if we concede that we have such defeasible dispositions to treat particular propositions as true in our reasoning, then a hypothesis naturally arises, namely, that beliefs consist in or involve such dispositions.

Thus, to reason with credences we need a full attitude that treats its content as true even when one is not certain of it. Decision-theoretic reasoning requires an attitude like belief as background information (Lance, 1995). When I consider the partition of relevant states in the Bayesian framework, I must also treat a number of background propositions as they were certain for me even if they are not. Ross and Schroeder (2014: 265) offer the following example:

Suppose Renzo has rented a DVD from a store on Canal St., and the DVD is due before the store closes twenty minutes hence. Renzo is deciding which train to go to the store by, the Broadway train or the Canal St. Express. He reasons as follows: If I take the Canal St. Express, it will cost me \$3, but I'll get to the store on time and so I won't be fined. Thus, I'll be out \$3. If, on the other hand, I take the Broadway train, it will cost me \$2. If it stops at Canal St., I'll get to the store on time and won't be fined, so I'll be out only \$2. But if the Broadway train doesn't stop at Canal St., I won't get to the store on time, and I'll be fined \$5, so I'll be out \$7. Since *it's as likely as not* that the Broadway train won't stop at Canal St., it isn't worth risking the fine to save \$1 on the train, so I'll take the Canal St. Express.

Now, Renzo is taking into account two possibilities: either the train stops at Canal street or it does not. However, he is not certain of this disjunction, i.e. he has non-zero credence that the train may never leave the station or crash before even getting to Canal street. And yet he does not take into account these possibilities and treats the disjunction as certain. Similarly, Renzo is not certain that what the actual outcomes of the actions

will be as he can't rule out the possibility that the price of the train may have changed. However, it seems reasonable for him to take this proposition as if he were *certain* of it and thus exclude the possibility of its falsity. As Ross and Schroeder (2014: 266) put it, "After all, if he were to take into account every relevant possibility in which he has nonzero credence, then his decision problem would be unmanageable, and the store would be closed long before he could decide between his options." Our cognitive resources are limited and we can't take into account the possibility of the negation of each proposition that figures in our reasoning and about which we are uncertain. So some propositions will have to be taken as certain. And since we define beliefs as the attitude that treats a proposition as certain in reasoning, it seems that we can't calculate expected utilities without full beliefs (also Wedgwood, 2012: 323).

Ross and Schroeder, however, only focus on the fact that reasoning with credences requires the use of full-beliefs as background conditions. What remains unclear (in the sense that they don't talk about it) is whether beliefs play the role only of background conditions or whether there is actually a form of reasoning that employs **only** full beliefs.

For simplicity, assume that what we are after in pragmatic reasoning is maximal satisfaction of our desires. There are two ways in which we can decide what to do through reasoning.⁴⁸ We can base our decisions on propositions that are the reasons (we think) we have. Alternatively, we can assign subjective probabilities to some propositions and reason from those based on expected utilities. That means that, at times, we engage in careful probabilistic reasoning as the one described by – among others – Bayesian

⁴⁸ Sometimes we don't pick any strategy and we decide randomly or perhaps based on trivial factors (e.g. we buy what product we saw first on the shelf at the supermarket). Sometimes we do that because the stakes are either very low or incredibly high and we simply don't know what to choose. Here, however, I am interested in those cases in which what we do is the result of a reasoning that produces an actual choice.

decision theory. At other times, we simply start our reasoning by adopting as premises *flat out attitudes* about how the world is or will be.

Reasoning based on full attitudes is reasoning in terms of reasons. By ‘reasons’ I mean considerations that support my conclusion. When I treat something as a reason, I am motivated to act on it. So when I decide what to do based on (what I take to be) reasons I am motivated by the fact that there is a state of affairs that by being *true* satisfies a desire that I have (or make an action appropriate or required). And since I believe it to be true, I am motivated to act accordingly. In contrast, when we employ decision theory, no (apparent or real) reasons motivate us to act, but the motivational power of reasoning is the product of the fact that the decision we reach maximizes expected utilities.⁴⁹

Now, common sense strongly suggests that we in fact are motivated by what we take to be reasons. Our everyday experience and the way we talk about reasoning seem to indicate we often reason by making reference to reasons. So, intuitively, this gives some initial support to the idea that we employ full beliefs to make decisions.⁵⁰ Beyond common sense, there is an on-going research program in psychology dedicated to determining what methods we use to make decisions.⁵¹ Research strongly indicates that, when deciding what to do, we do not make choices only by calculating expected utilities but we rely on a variety of different methods. It appears that these methods are constantly used because of their ability to deliver reliable conclusions (in some contexts) while

⁴⁹ Bayesian decision theory does not use talk of reasons and explains rationality as determined by calculation of expected utilities. Bayesians tend to consider reason-talk as a mere remnant of an obsolete folk psychology, or at best as a loose way of talking about what to do.

⁵⁰ Also, from an internalist perspective, it is unclear that all our reasoning is about desire-satisfaction or maximization of utilities. It is possible that desire may be excluded from (some) reasoning patterns.

⁵¹ For some early work in this program, see Simon, 1956; Tversky, 1972; Kahneman and Tversky, 1982.

using little energy and time. This shortened reasoning can take various forms so, in what follows, I'll run through a few paradigmatic cases drawn from current research in psychology and simple observation/description of our practices.

Here is one. Full belief reasoning often makes use of explicit *rules* adopted in advance of making any individual decision.⁵² When we make decisions about what to do in a specific situation, we simply adopt the rule that applies to the situation without calculating pros and cons. These are sort of pre-established rules of conduct that get applied when facing repetitive decisions. So if I notice I am out of bread, I'll go out and buy it. If I see a red light, I stop.

Alternatively, instead of using explicit rules, we decide what to do by using one of our desires as premise. Here is an example of this type of reasoning:

If I go to the library I'll write at least 2 pages of my dissertation.

If I stay home I'll get nothing done.

I really want to get some dissertation work done today.

So I should go the library.

Importantly, in this form of reasoning (as in the previous rule-based one) we employ full beliefs. In addition, here we use full desires as well. Instead of assigning different values to different outcomes as in decision theory, here one simply assumes there is only one valuable outcome, and ignores any other desire or goal one may have. In the example above, I only see as valuable the fact that I will be able to write my dissertation. Although there are various things we value and, perhaps, we can even quantify how much we value them, at times we also assign 'maximum' value to some restricted set of things. We can

⁵² Some argue that these are not explicit rules, that they don't take the form of an enthymeme but instead are expression of material inferential relations of the following kind: "it is raining, I shall open the umbrella" (Brandom, 1994).

call these attitudes ‘full-desires’. As in the case of belief, to have a full desire is to have a *defeasible* disposition to act in order to bring about the content ‘p’. That means that, for instance, if I fully desire to work on my dissertation I’ll be disposed to use that as a premise in my practical reasoning (when p is relevant).⁵³ There are going to be cases in which it becomes clear that ‘p’ is not the most important goal or that there are other goals that are just as important. In *those* situations, I will have to adopt a different reasoning strategy, and look at the actual value that writing my dissertation has for me as opposed as to the value that I may assign to different outcomes. This is why the strategy of adopting full attitudes in practical reasoning is defeasible.

Here is a third reasoning strategy we employ. Sometimes, when we are torn between two options, it may not be very practical to adopt a single desire as the unique value to make our decision, especially when both options satisfy that desire perfectly well. Instead, we will have to run through a bunch of possible good options to find a tie-breaker. So, for instance, when we are confronted with two options, going to restaurant A or B, we often frame our decision by running through various parameters of preferences (e.g. price, taste, location) till we get a tie-breaker.⁵⁴ That is, in this case there may be a number of things we value and we just need one of them to make the difference. This is a form of ‘one-reason decision making’⁵⁵ that is quite crude because it does not take into account that there may be other options we care for down the line and which we don’t

⁵³ I got this idea straight from Frankish (2004: 95) where he talks about ‘goals’. His view is that by “adopting a goal, x, we commit ourselves to taking x, together with other relevant goals and premises, as input to conscious intentional practical reasoning.”

⁵⁴ I got this example of heuristic reasoning from Weisberg (2013: 7). As Weisberg points out, another useful strategy is the take-the-best heuristic theorized by Gigerenzer (Gigerenzer and Goldstein, 1996; Gigerenzer et al., 1999; Gigerenzer and Selten, 2001).

⁵⁵ Even in this case, when choosing A over B, the agent hits upon a desirable feature that A has but B lacks, it is natural to say her *reason* for choosing A is that it has this desirable feature (while B lacks it).

have the chance to examine (Newell and Shanks, 2003; Gigerenzer et al., 2008). There are ways in which we obviate this problem as, for instance, we may rank parameters in order of importance (e.g. price is more important than location, but taste is more important than price, and so on)⁵⁶.

These quite common ways of reasoning do not calculate expected utilities but often employ full beliefs (and full desires) as reasons for their conclusion. The reason we may decide to adopt one belief-based strategy over another could depend on various factors such as the information we possess, the time we have at our disposal, whether or not we have rules or full desires that can be relevant in that particular situation, etc.

Those mentioned above are only some examples of the ways we may reason with full beliefs, but many more are possible.⁵⁷ I limit my analysis to describing only those few cases trusting that they are enough to make the claim that reasoning based on full beliefs (and full attitudes, in general) is a common practice that agents adopt.⁵⁸

FULL BELIEF AS HEURISTIC

Here I argue that full belief-based reasoning is a heuristic. My argument comes in two steps: using full reasoning is less costly and perhaps more reliable in some contexts. For starters, using full belief when reasoning is less *costly* than using credences and assigning values to different outcomes. So even if – contrary to what some have argued (Harman, 1986) – calculating expected utilities may be computationally manageable and

⁵⁶ Lee and Cummins (2004) posit a more sophisticated method, the evidence accumulation method.

⁵⁷ Payne, Bettman, and Johnson (1993) survey some of these methods.

⁵⁸ Some have argued that we simply can't eliminate full beliefs if we want to make sense of some of our common practices. So, for instance, Buchak (2014) claims that belief has a distinctive role to play given our traditional understanding of our practices of holding each other responsible and assigning blame, and such practices can't be explained in terms of expected utilities calculation.

can be done by using various heuristics, it is presumably still a quite expensive calculation because of the cognitive load it requires (at least when done in reasoning and thus with some level of awareness). In contrast, reasoning in terms of full attitudes may cut some of those costs as it seems plausible that a reasoning strategy based on full beliefs may require less energy, cognitive effort, and time than a full-blown partial belief-based strategy (again: at least when some level of explicit reasoning is involved). Furthermore, there is evidence that employing full attitudes together with some of the heuristics mentioned above may prove to be more efficient than running through some decision theoretic strategy. As a result, it is possible that rationality requires of us to use full beliefs in reasoning and not credences, at least by default. The ultimate conclusion is draw from this is then that employing full belief is our default strategy in reasoning.

There are a few arguments in support of the first claim above. For starters, using belief as a premise is itself a heuristic (whether or not the reasoning employs any heuristics).⁵⁹ When one reasons based on partial beliefs less than 1 one forms a partial belief that p is, say, 0.8 and a partial belief that not- p is 0.2. Then one needs to calculate expected utilities for the possibility both that p and that not- p is true. That means quite a bit of calculation and a quite high number of options to keep in mind. In contrast, with full beliefs (and full attitudes more generally) one can simply evaluate one's options conditional on p while simply ignoring the possibility that not- p . That means fewer options to keep in mind. What's more, in this type of reasoning there is often no calculation to make at all: one simply draws an inferential step from one's belief and,

⁵⁹ To illustrate: the inferential reasoning I mentioned above is not usually considered a heuristic form of reasoning, but regardless of *that* using belief as premise is a heuristic.

possibly, one's desire. If so, then using full attitudes in reasoning is a simplification that ignores some information to cut costs.

What's more, calculating expected utilities with credence may actually be a method of reasoning that requires much more revision than deciding based on beliefs in the sense that if you calculate expected utilities you may need to revise your decision more often. Credences are fine-grain states and in some cases any new information coming in may change the level of credence one has (and thus may change the result of our expected utilities calculation). So to be fully rational one should be ready to re-calculate one's expected utilities each time new information is acquired. Here is an example. Say there is a big jar with 200 hundred red and blue marbles, but you don't know the ratio between the numbers of red and blue marbles. Say that you have to calculate expected utilities of some outcomes based on the credence you form that the last marble you draw will be blue. Now, you have drawn 20 marbles, 19 of which have been blue. As a result, your credence that the last marble you draw will be blue is 0.95. You now calculate expected utilities based on that. Imagine now you draw another 2 marbles and those are red. Now, your credence-level is diminished. It is possible⁶⁰ that this change in level of credence may require you to update any calculation made beforehand. And that seems quite costly.

In contrast, full beliefs are coarse grained. Going from believing in something to disbelieving it or suspending judgment usually requires a fairly substantial change in one's evidence. So the reasoning process in this case is often much more stable, it does not need constant revision. If I am right, then having to calculate expected utilities may be a quite precise way to proceed, but also more effortful than full belief-based reasoning.

⁶⁰ However, it is also possible one may use heuristic methods to decide when recalculate expected utilities so that one does not have to recalculate each time new information is acquired.

The *second* issue I mentioned above is that in some contexts reasoning with full beliefs may prove to be as reliable as calculating utilities. As it is often the case with heuristics, their reliability is tied to the ‘context’ in which they are used. And this is the case also for reasoning based on full attitudes: in normal conditions, it seems these are *effective* forms of reasoning. Their effectiveness is explained here in terms of their reliability in the production of decisions that largely conform to expected utilities. So reasoning by using as premises propositions we fully believe may be just as reliable a method as actually calculating expected utilities by using credences. When more is at stake, we consider further attributes or even calculate actual expected utilities. But in many cases, a restricted set of considerations suffices to settle the matter.

On top of this, it seems that using approximations is not only reliable but at times even *more* reliable than taking into account all the possibilities in a full rational way. Often addition of options increases the burdens of decision without increasing, much or at all, the likelihood of a good decision. Thus having many choices to choose from might well increase decision costs without improving outcomes. At the same time, note that the efficacy of heuristics is related to the context in which they tend to be used (or have evolved) and, when used out of context, they produce biases and suboptimal decisions (Gigerenzer and Brighton, 2009).

TWO MODELS OF RATIONALITY⁶¹

In this section I advance the hypothesis that acting by using belief-based reasoning may be at times the rational thing to do. That means that Bayesian standards may not be

⁶¹ The argument I offer here is in part a combination of two distinct arguments, one offered by Locke (2013), the other by Sargent (2009). However, the argument itself is original.

the only game in town when it comes to rational behavior and a more nuanced view is needed. Now, Bayesian models of decision-making are considered rational because they indicate how to maximize the benefits given the subjective probabilities and preferences of the agent. The ultimate goal is to satisfy the agent's desire. The rational way to do it is, on this view, to use subjective probabilities and preferences. In its standard formulation Bayesian Decision Theory can be stated in terms the following norm: we should perform an act only if that act has higher expected utility than any of the other available acts, given your credences in the events which bear on the utility of the acts.

This view seems to exclude the role of full belief in rationalizing action. Here is what I mean. The role of belief in producing action is not merely explanatory; belief should also be able to rationalize action, and make sense of it. Above I pointed out that belief explains actions by taking part in practices of reasoning that use *reasons* to motivate action. And many of these practices are commonly used as reasoning strategies. However, the role of belief in rationalizing action remains controversial. Here is a way to spell out this worry:

The rationality challenge: if a full-belief is a disposition to act with the certainty of truth, then full belief cannot take part in a rationalization of people's behavior.⁶²

If decision theory is a normative standard, as many claim it to be, acting as if *p* were certain, when one is not certain of it, is a violation of the principle of rational decision-making. Thus, if full believing is what CLAIM takes it to be then, whenever we act on a full belief we are either acting irrationally, or the rationality of the action depends solely on whether it maximizes expected utilities. So folk explanations of people's actions in

⁶² This is sometime taken to be an argument against the existence of full belief. If it is irrational to act on beliefs, and – as argued by Davidson - we should attribute rationality to agents, then we can't say that people ever (or at least very rarely) act on beliefs (Frankish, 2004: 54-55).

terms of full belief would not be able to rationalize those actions. And this is a problem because we naturally see belief as playing a role in rational action and rational decision-making. Indeed, according to the objection it appears we face a dilemma: either we give up CLAIM or we renounce to take belief to be playing the role in rational explanation we want it to play.

I will offer the following reply: at times reasoning with full beliefs doesn't produce choices that are optimal in Bayesian terms. However, I still believe there is room for arguing that belief can rationalize action in some instances.

A few clarifications. My view is not that it is sometimes rational to be irrational (although this is not unheard of). The view is that it is at times rational, in a bounded sense, to be irrational in an ideal sense. If the goal of practical reasoning is reaching a conclusion that satisfies one's goals or desires, then reasoning with full belief may be what in fact gets us there.

Also, for simplicity I will assume that a belief rationalizes action when the belief is true, i.e. when it actually counts as a reason for doing that action. In other words, I will leave aside those cases in which a belief may seem to rationalize an action even when false. So I will treat 'belief rationalizes action' as synonymous with 'belief provides a reason for action', and use 'rationalizing an action', 'making an action rational' and 'justifying an action' interchangeably. What's more, I am not committed to any general theory about what sorts of reasons, if genuinely held, would be sufficient to justify actions, and so I'll vaguely refer to 'sufficient reasons' hoping that this is enough for making my point. Finally, note that my aim here is to vindicate the thought that belief can rationalize action. Here I will try to carve out the rationalizing power of belief from

within decision-theory and show that, instead of getting rid of talking of belief as being a reason for action, perhaps we should integrate it in a larger decision-theoretical prospective.

Now, when does a fact or true proposition make an action rational? First, when it is a sufficient reason for it. Two, the agent who does the action has to be in some relation to that fact, i.e. she believes it to be true. More specifically, I will assume that if an agent S believes that p and p is true and is a reason to bring about an action X, then S *has* a reason to X. And finally, if the belief that p stands in an appropriate relation to X (e.g. it caused X), then we say that X was rational.

Now, the problem is that it is possible that X or the choice of doing X may not maximize expected utilities. So how can we say that X is rational? My proposal is that we say that action X is rational in normal contexts *if other conditions are in place*.

The proposal comes in two steps:

1. As a reasoning strategy to maximize utilities, it is rational to *use* belief-based strategies as a default way of reasoning in normal contexts.
2. In normal contexts it is rational – for a bounded-rationality prospective – to *act* on one's sufficient reasons even when doing so produces a less than optimal choice (for that particular action).

Importantly (1) is about what reasoning or premising strategy it is rational to use. Rationality is on-bounded here. In contrast, (2) is about action, and talks about bounded rationality. The two are connected, yet separate, as it may happen that I do the right action for the wrong reasons or by reasoning inefficiently.

Here is why (1) is, I believe, plausible. First, even in those cases in which reasoning based on belief does not offer the same outcome that reasoning by calculating expected utilities (and therefore is suboptimal), it may still be rational to choose to reason based on beliefs *because of the cognitive costs of reasoning otherwise*. Acting on beliefs, given the cost of calculating expected utility, may in the long run have a higher expected utility than performing difficult calculations. For calculating expected utility is itself an act; as an act it may not maximize expected utility (Sargent, 2009). In other words, at times it may be more rational to choose to reason with belief than having to engage in a complex reasoning calculus.⁶³

My proposal is that it is rational to use belief-based reasoning as a default. As Frankish (2004: 54) puts it:

Perhaps we all habitually lapse from the high standards of Bayesian rationality, ignoring the subtleties of confidence and acting on the basis of unqualified beliefs. Such a habit might be justified, given our cognitive limitations- in particular our lack of skill at conscious probabilistic reasoning. If accepting probable propositions flat-out helps us make our calculations more tractable, then the departure from strict rationality which it involves might be justified by the accompanying reduction in computational demands.

Now, default strategies are used in normal contexts. A context is normal unless *there are* reasons to think that the difference between one's degree of belief in p and certainty that p is such that it would make a significant difference in the expected utilities of the outcomes of my actions. And by 'significant' I mean that the difference is more than the actual cost of having to calculate expected utilities.

⁶³ Note that this does not mean that choosing any method that is less expensive in terms of cognitive resources (e.g. coin-flipping) is always the rational thing to do. This is because the method itself needs to be reliable enough to ensure that the outcome will be, if not optimal, at least good enough to counter-balance the risk of not basing one's choice on a careful decision-theoretical form of reasoning.

Here are two examples of two a-normal contexts:

BET. I am offered a bet on the truth of some proposition ‘p’ that I regard to be false even though I am not certain that it is false. However, the bet is practically a win-win for me: if I bet on what I believe to be false and I win, I get \$1000, otherwise I lose only \$1. I decide to bet.

SURGEON. Scott is a surgeon in the local hospital. He is ready to perform a kidney surgery on one of his patients. Scott has high credence that the kidney that needs to be removed is the left one. He also fully believes that.⁶⁴ If he were to act in accordance with his belief he would go ahead and perform the surgery. However, Scott is not certain of the truth of his belief. *Given the stakes*, Scott is quick to recognize that were he to be mistaken the consequences would be terrible. So he decides to stop by his office and check his patient’s dossier before performing the surgery.⁶⁵

BET is not a normal context because there are reasons to believe that the difference between my level of credence in not-p and having certainty that not-p is large *enough* to make a significant difference in the expected utilities of the outcomes of my actions.

In SURGEON, given that so much is at stake, there are good reasons for Scott to think that expected utilities in that context significantly vary depending on the level of confidence of his belief. Accordingly, since he is not certain of the truth of his belief he correctly decides to act not based on what he believes but on his calculation of expected utilities based on his level of credence and possible outcomes.⁶⁶

⁶⁴ One could try to argue that the relation between level of credence and belief is contextual, and thus in this context Scott does not believe that the kidney that needs to be removed is the left one. I am not convinced that belief can be contextual in this way.

⁶⁵ The idea for this example is taken from Brown (2008: 1144–1145).

⁶⁶ Note that it is not an objection to my view to say that in SURGEON Scott does not have a sufficient reason to go straight to the operating room. It is notoriously hard to say what makes a reason sufficient for

Note that an important result of this is that, besides having full beliefs about some propositions, to act rationally one also needs partial beliefs about those same propositions. When it becomes apparent that one's level of confidence in *p* is so distant from certainty to make a significant difference in the expected utilities, rationality requires to switch to a different method of reasoning.

Now (2): if one acts based on a sufficient reason one has in a normal context, then her action is rational. I take (2) to be the combination of two things, i.e. the rationality of one's *reasoning* strategy in maximizing utilities (in the long run) and the fact that one has sufficient reasons for doing an action. If correct, this produces the following principle:

(Bounded) Rationalization principle: It is rational to do action *X* if one believes that *p*, such a belief is connected to *X* in the right way, *p* is a sufficient reason to *X* and *X* is done in a normal context.⁶⁷

According to this principle, there are cases in which belief can't rationalize action, even if having that belief means having a sufficient reason for doing that action. The explanation for that is that in those cases the choice of using belief as reasoning strategy is irrational because the actual effort of calculating expected utilities is balanced out by the expected utilities of doing a particular action.

In conclusion, an action based on a belief – as defined in CLAIM – can be rational (in the bounded sense) when the belief in question constitutes sufficient reason it. Belief's rationalizing power is not reducible to credence's, though. In contrast, belief rationalizes

action and I have nothing to contribute to that debate here. All I am saying is that, *if* there are cases in which one has a sufficient reason to act and yet it does not seem rational to do so, then this is because the context is not normal.

⁶⁷ This may also mean that - in normal contexts and given the other conditions in place - it is irrational to calculate expected utilities, although it may be rational (again in the Bayesian sense) to act based on that calculation.

action by being sufficient reason for it in those contexts in which it is rational to *reason* based on beliefs.

BELIEF AS DEFAULT IN REASONING

In this final section I argue that there must be a default attitude, i.e. an attitude we adopt as a premise in reasoning in normal contexts, and that is belief. The reason why we need a default is the following. Both belief and credence are dispositions to act in a certain way. It may happen that they conflict and push in different directions. What is rational to do given my credences may be different from what is rational to do given my full beliefs.

So either neither of them motivates or enters reasoning by default or at least one of them motivates or enters reasoning by default. If neither of them motivates or enters reasoning by default then they both motivate or enter reasoning in relation to specific contexts of action. However, if this is the case, then neither of them can motivate automatically. The problem with this is that each time an attitude, say a belief, motivates action or enters reasoning the context has to be somehow represented (since we assumed belief does not motivate by default). But this representation is also a doxastic state (i.e. that the context is such and such), and so the worry of a regress surfaces. So there has to be an attitude that motivates by default. I think that is belief rather than credence.

Now, granted that reasoning in terms of full attitudes and reasons is indeed more convenient than reasoning in terms of partial beliefs, the next claim I want to make is that this is how we reason by default (when, of course, beliefs are available). This proposal is in line with a strong line of research in psychology that indicates that our mind cuts

corners in order to produce outcomes efficiently with the limited capacities it has. Although many of the decisions we make based on heuristics are not relevant here because in many of those cases our decisions are the outcome of an intuitive process and don't count as reasoning-based, it is worth explaining the background literature on heuristics.

For instance, in economics there is strong support for the claim that humans don't use decision theory to make financial decisions but appeal to various heuristics. More broadly, some psychologists have recently put forward the idea that we reason by using a suit of mechanisms that produce responses in a fast, automatic, and fairly effortless way. These mechanisms are the default way we reason as they don't require conscious applications of rules and conscious monitoring of the steps of our reasoning process. This suit of mechanisms is usually called System 1. In contrast, when the external conditions demand that, we employ a slower, more cognitively demanding way of reasoning, in which we consciously weigh the options we have. This second way of reasoning is System 2. Importantly, what this kind of research has indicated is that most of the time we seem to reason based on heuristic strategies that are efficient given the limited cognitive resources we have at our disposal. In other words, probably because of our limited resources, when we face uncertainty, we tend to simplify our choices. Given the various methods of reasoning we have, some of which more efficient than others, it appears that we use those efficient ones as our basic strategy.

Similarly, the tendency to simplify encompasses also the strategy we adopt in reasoning 'proper' (i.e. what I called 'reasoning'). As far as we can, we go for the shortest route. And since it appears that we do adopt full beliefs in reasoning, it is also

plausible to posit that this is our *default* way of reasoning. More specifically, if full belief-reasoning is somewhat a heuristic, then it is probably our default reasoning method. In contrast, when more precision is needed⁶⁸, we switch to a reasoning strategy that calculates expected utilities by conforming or trying to conform to (something like) Bayesian Decision Theory.

In conclusion, in this chapter I argued that belief plays a role in reasoning and motivates action by being treated as a *reason* for action. Reason-based (and thus belief-based) reasoning appears to be very common. On-going research in psychology shows that we often reason by using shortcuts and heuristics in which belief plays the role of premise. This way of reasoning, I argued, produces (bounded) rational actions in normal contexts, and it is the most rational way of reasoning in those contexts. As a result, this gives credit to the claim that belief is the default attitude we use to make choices.

⁶⁸ This point is actually quite controversial. Studies have shown that when the stakes are high we still tend to reason based on heuristics (e.g. investment behavior). One reason for that is that in many high stake contexts there is a problem of computational intractability that makes it the case that we do (and even perhaps should) use heuristic to decide what to do. So perhaps a somewhat more precise claim here is that we do use decision theory when more precision is needed and when we think that the problem at hand is cognitively tractable.

BELIEF AND ITS CONSEQUENCES

Chapter 2

In this chapter I argue for the claim that belief has a unique role in our mental lives. To do so I attack the recent claim that the output role of belief, its ability to influence action and other attitudes, is shared by other cognitive attitudes such as (propositional) imagination, acceptance and supposition. My argument comes in various steps. First, I defend the claim that *only* belief (and desire) can motivate action. Second, even assuming that other cognitive attitudes can motivate action, that is not enough to show that their consequences are the same as belief's. Belief is an 'unqualified' cognitive attitude whereas all these other attitudes *must* come with some kind of constraint on their use that limits the type of inferences or reasoning patterns they can enter into. This has two main upshots: for belief the possibility to be defined in terms of its unique role in inference and reasoning is still open. And we do not need to refer to belief's relation with truth to individuate belief uniquely. As I will argue in the final chapter, it is actually *because* belief has the kind of unconstrained output role it has that we can't believe at will.

In philosophy and cognitive science, mental attitudes types (e.g. beliefs, desires) are often defined in terms of how they are connected to other mental attitudes. This relation is explained by looking at attitudes' input (or upstream) and output (or downstream) roles, namely the sort of inputs they are sensitive to and the sort of consequences they tend to have (Fodor, 1985; Nichols and Stich, 2003). When applied to belief, it has been often argued that beliefs are individuated by their output role in causing action and the formation of other attitudes. Not long ago, in fact, it was common to individuate belief as *the* cognitive attitude *par excellence*, by being an attitude that has a content, usually a proposition, represented as being true, and combining with a conative attitude to give rise to action. Sure, there are other cognitive attitudes that populate our mind (e.g. supposition, imagination), but these are clearly distinct from belief for lacking any straightforward relation with action.

This standard picture and the possibility to individuate belief based on its motivational role have come under attack in three related ways. First, belief has progressively lost its predominant role as the attitude which produces action because, as the new story goes, belief is at times neither sufficient nor necessary to explain action. It is not necessary because often action can be explained by invoking other mental attitudes (e.g. imagination) that make sense of it just as well as belief does. And it is not even sufficient because some of the things we do just can't be explained by belief, as in the case in which we act contrary to what we believe. To illustrate, if I believe that the bridge is safe, why do I refuse to step on it? If I believe that the female candidate is just as good as the male candidate, why do I keep hiring males? And so on.

Two (and partially as a result of one), an array of new cognitive attitudes (i.e. attitudes that present their content as true) has been emerging and the category of "cognitive" attitudes has gradually become more densely populated. To offer a fully convincing account of what produces action it has appeared to many that either we need to postulate the existence of new, previously unacknowledged cognitive attitudes, such as alief (Gendler, 2008) or in-between attitudes (Schwitzgebel, 2010), or we should offer a new metaphysical picture of those well-known cognitive attitudes, and say that belief and other mental attitudes, such as imagining, are not necessarily discrete states but stand on a continuum (Schellenberg, 2013).

Three, many now believe that non-doxastic cognitive attitudes (e.g. imagination, acceptance - call them "secondary cognitive attitudes") share the *same* output role as belief and thus its downstream consequences are not enough to individuate belief. The individuating aspect of belief is, on this view, its relation with truth. This relation is

usually spelled out either in terms of some aim that is constitutive of belief (belief aims at truth, or believers aim at truth) or in terms of a norm of correctness that is part of the nature or concept of belief (see Introduction and chapter 5).

This last point has been forcefully defended by David Velleman (2000). Velleman starts off by noticing that belief is part of a large group of cognitive attitudes that all present their content as true. These attitudes, he adds, also share the ability to cause actions when combined with some suitable conative attitude. So we can't differentiate believing from, say, assuming based on belief's output role in causing action. As a result, Velleman presents the following 'Identical Output role' view:

(IOR) all cognitive attitudes share the same output role.

IOR is important because it paves the way for the argument that it is belief's *truth-directedness* (i.e. its aiming at truth) that distinguishes it from all other truth-regarding attitudes (Velleman, 2000: 252; Shah and Velleman, 2005). Velleman is not alone in this. For instance, in Gendler's own words: "I think that the most helpful way of distinguishing beliefs from other related cognitive attitudes is neither [...] nor through their dispositional connection through desire to action, but through their telos of truth." (2007: 236) Similar arguments are offered by Engel (2004), Lynch, (2009), Railton (1994), and Wedgwood (2002).

My goal here is to prove this picture wrong. First, I defend the claim that only belief (and desire) can motivate action. As a result, we don't need to invoke other or new attitudes to explain action: belief and desire will suffice. Second, even assuming that secondary cognitive attitudes can motivate action, that is not enough to show that their

consequences – as a whole – are the same as belief's.⁶⁹ This means that belief can indeed be singled out for its unique output role and independently of its relation with truth.

SETTING UP

Before starting I need to make some clarificatory remarks and explain some of the terminology I will adopt. For starters, here I talk about 'cognitive attitudes'. There is a distinction to be drawn between cognitive and conative attitudes. Although this is quite vague, for our purposes it is enough to say that conative attitudes (e.g. desire, hope, wish) present their content as to-be-made-true, whereas cognitive attitudes present their content as true.

In philosophy and cognitive science, mental attitudes types are often defined in terms of their input (or upstream) and output (or downstream) roles, namely the sort of inputs they are sensitive to and the sort of consequences they tend to have. Output roles are usually explained in terms of dispositions. By dispositions to act (or form other attitudes) I mean *ceteris paribus* motivations to act (and form other attitudes). By saying that belief is a disposition I mean that when the conditions are right, if one believes that *p* (and has a related desire) no extra motivation (or decision) is needed to make one act in accordance with the content of one's belief. In particular, if I am disposed to do *X* I am so disposed regardless of any decision pro or against doing *X*. In contrast, if an attitude does not dispose me to do *X*, then in order to be motivated to do *X* I must either decide to do it or have an additional desire to do it. (This point will become more clear once I introduced the view – held by many functionalists – that imagination does not dispose

⁶⁹ For simplicity, I will only focus on two sets of secondary attitudes, i.e. imagination and acceptance, and leave aside other cognitive attitudes such as alief.

one to act. On this view, to act on one's imaginings one needs to decide to do so or to desire to do so.)

A disposition is a *default* disposition only if it motivates by default, i.e. in normal contexts. What's more, a disposition that motivates by default motivates *automatically*. In contrast, a disposition that motivates only in *some* contexts does not motivate us automatically but it requires an additional input, namely a representation of the context as 'the right context', or as 'such and such context' (and 'context' is the practical setting in which a certain attitude was introduced). To detect dispositions we should look at them as emerging in patterns of functional roles. More clearly, as I will concede below, it may be true that *individually* imagination and belief dispose one to act and produce action in similar ways (other things being equal), but this is not enough to argue that they share the *same* functional role.

Finally, philosophers tend to draw the distinction between standing vs. occurrent belief. Some philosophers have gone so far as to claim that there are no standing beliefs and that all our beliefs are occurrent states. In contrast, following the computational view of the mind, I take standing-state beliefs to be stored representations in the belief-box, whereas I see occurrent beliefs as activations of these representations for reasoning (Fodor, 1987).

IMAGINATION & ACCEPTANCE ... WHAT ARE THEY REALLY?

The goal of this section is to convince the reader that there are things we imagine and things we accept/suppose to be true and that, despite some differences, these two attitudes share important similarities. I won't provide a full-blown account of imagination and

acceptance, but just offer some intuitive remarks that should suffice to explain why I am lumping these attitudes together.

Let me first address some understandable worries. We may be concerned that the ways in which we talk of imagination and the like are not all together unified. Ordinary talk of imagination seems to be all over the place. We say things like “I imagine we will have a good time at the party” meaning “it is likely we will have fun”. These expressions have little to do with imagination in the sense of daydreaming or fantasizing. Similarly, it is common – and I myself will do that below – to link imagination and pretense. However, I can pretend to be rich to convince you to marry me, but it is unclear to what extent I need to *imagine* to be rich to do that. On the other hand, we say things like “He imagines to be Napoleon” when someone is under the delusion to be Napoleon. But that person is *not* in any way pretending to be Napoleon. His imagining is likely the source of his delusion, as memory can be the source of one’s belief. In this sense, imagination is a faculty that can produce beliefs, mostly delusional ones. To add even more confusion: Alvin Goldman distinguishes ‘e-imagination’, i.e. enactment imagination, from imagination that has as its object a proposition and is similar to suppositions and the like (more on this in a minute). Enactment imagination is the simulation of another attitude. For instance, during a heated political discussion, I can imagine to believe that Trump will be a good US President to understand my Republican friend’s viewpoint. Here I am *simulating* to be in a certain state (e.g. believing) by running it ‘off-line’. In this sense

imagining is a process I am in. In contrast, propositional-imagination does not seem to require any simulation.⁷⁰

Given the heterogeneity of the ways in which we speak of ‘imagination’, my job here requires that I keep things manageable. I will focus on a kind of imagination that takes, as its content, a proposition. Propositional imagination thus happens when we imagine that such-and-such is the case. To imagine that the cup is full is to have an attitude with the representational content *the cup is full*. Importantly, such content is – when imagined – presented *as true*.

What’s more, I will assume that imagination and belief can share contents: one can actually simultaneously believe a proposition and also imagine it (see Leslie 1994 for a nice illustration of this phenomenon). There is a caveat here: imagination cannot have a singular thought as its content when the demonstrative is *inside* the scope of the imagining operator. I can imagine that *that* person is a policeman only if the ‘that’ refers to something outside the scope of the imaging itself. If there is no person there, then the indexical fails. So one can’t hook an indexical to an imagined object unless it is used anaphorically: ‘I believe there is a person there and imagine that person is a policeman’.⁷¹ Alternatively, one imagines that there is *a* person there and that that person is policeman.

Propositional imagination is a self-standing cognitive attitude and thus can’t be reduced to belief or any other attitude.⁷² That excludes the use of “imagination” as equivalent to “believing something to be likely”. The aim of this chapter is to show that this distinction can be drawn based on belief’s and imagining’s divergent output roles.

⁷⁰ For a lucid analysis of the difficulties of trying to account for all the roles imagination can play in terms of a single mental activity (or attitude) see Kind (2016). Also, Gendler (2011) illustrates the various taxonomic challenges posed by imagination.

⁷¹ This is why disjunctivists think that imagining and the like can’t be the same mental state as perceiving.

⁷² For a dissenting view, see Peter Langland-Hassan (2012).

Finally, the kind of imagination I am focusing on here is mostly done deliberately or is, at least, under our power. We can decide to imagine being Napoleon; and usually we can also stop doing that of our own accord.

Now, propositional imagination is very similar to a group of attitudes which includes suppositions, assumptions, acceptances, presuppositions and the like. They are similar to propositional imagination in various respects: they entertain a proposition and treat it as true; they are mental attitudes not reducible to beliefs although they can share the same content of beliefs;⁷³ and reasoning and cognitive systems treat inputs coming from acceptances in the same way as they treat inputs coming from beliefs.

For the most part philosophers agree that acceptance/supposition/assumption is a mental state that is distinct from belief,⁷⁴ and that we can resolve to accept a proposition for a variety of reasons: evidential, moral, professional, religious, and so on. Indeed, we provisionally assume a proposition for the sake of argument (Fisher (1989). Green (2000) explains that we accept propositions for the sake of drawing out their logical or material implications (Fisher, 1989), or simply to entertain a possibility (Denham, 2000). But we also do it for some practical or moral reasons. Imagine I want to save money to send my now two year-old daughter to college, so I decide to reason by assuming that it will end up costing me \$500,000. In fact, I believe it will cost less than that, but I just want to be on the safe side. Thus, I save \$2600 each month and put it on her bank account. Another case. During a conversation you say to me, “The blue car she stole was a sedan.” Even if

⁷³ Nichols et al. (1996) call acceptances ‘pretend beliefs’. To pretend-believe *p* is to feed a representation of *p* into the inference mechanism without necessarily believing it in order to produce, among other things, counterfactual reasoning.

⁷⁴ For an account of these attitudes see van Fraassen (1980), Stalnaker (1984), Bratman (1987), Rey (1988), Cohen (1992), Engel (2000), Van Leeuwen (2009; 2014) among others. The notion of acceptance includes attitudes that are so different that one may doubt they should be regarded as belonging to the same category. I will assume that they do.

I know that the stolen car is not blue, I might, for the sake of keeping the conversation going on, accept that it is. So I nod along and let you talk (Stalnaker, 2002: 718–9). Or take the following. A lawyer might accept that her clients are innocent for professional purposes. While working, her actions and behavior reflect what she accepts to be true regardless of what she actually happens to believe.

These attitudes are similar to propositional imagination in various respects. In particular, they bear a similar relation to singular thought. Thus, a lawyer can imagine, believe, suppose, or accept that her client is innocent. But she can't assume/accept/suppose that *that* client is innocent unless the indexical refers to an actual client. What's more, as for imagination, these attitudes influence action, in relation to some suitable conative attitude: If I want to save for my child's college and assume it will cost \$500,000, I will be motivated to save \$2600 each month. And so on. Let me point to one last key parallel between imagination and acceptance. This will come up again, but I want to mention it from the start. Acceptance and imagination are by their very nature compartmentalizable. This allows for the fact that one can accept/imagine *p* and that not-*p* as long as the two attitudes pertain to two different 'compartments' of the mind, in the sense that they relate to different sets of inferential relations or get activated in different situations and contexts (work vs. home). Thus one can accept/imagine contradictory propositions without facing the situation in which contradictory propositions play the role of premises in the same set of inferences and at the same time. Notably, this compartmentalization is not the result of some cognitive shortcoming of our part, but the necessary feature of having attitudes such as accepting/imagining.

So, given these similarities, I am prepared to group acceptances and imagination together and see how they differ from belief. But not everybody agrees. For instance, in a recent paper M. B. Jackson argues against what she calls the Common Nature Thesis, i.e. the thesis that imagining, supposing, and accepting are all instances of the same “mental kind”. Her argument is that imagining has a strong phenomenological component that suppositions may lack. Other philosophers have argued that imagination is a distinct category from acceptance and supposition. Gendler (2000), for instance, has pointed out that “imaginative resistance” sets imagination and other secondary cognitive attitudes apart. Take some morally charged proposition *p* that you do not believe to be true. Gendler maintains that you may feel a strong resistance to imagine that *p* is true, but you have no problem supposing or accepting for the sake of the argument that *p* holds (Gendler, 2000: 80–81). Finally, Weatherson has recently pointed out that supposition, acceptance and the like “can be coarse in a way that imagining cannot”. He seems to believe that “[w]e can suppose that Jack sold a chair without supposing that he sold an armchair or a dining chair or any particular kind of chair at all” (Weatherson, 2004: 20), whereas apparently this can’t be true of imagining.

I will not deal with these worries here as I don’t deny that there are *some* differences between imagining, supposing, accepting, and so on. The real question, though, is whether these differences are enough to set these attitudes apart. The view I am defending here is that its output role is what makes an attitude what it is. So the working assumption here will be that there *is* a category of mental states (called ‘secondary cognitive attitudes’) that share the same output role. That is, propositional imagination, supposition, and acceptance’s relation with other attitudes are roughly the same. Their

output role, though, is not the same as belief's. That said, it is not pivotal to me to establish that imagination and the like *actually* form a unitary category of attitudes. I will use that only as a working assumption to identify how secondary cognitive attitudes differ from belief, and leave for another day the question of whether there are differences among secondary cognitive attitudes themselves.

WHAT MOTIVATES ACTION

As mentioned above, Velleman puts forward the following 'Identical Output role' view:

(IOR) belief and secondary cognitive attitudes share the same output role.

Mental attitudes are commonly characterized by their functional roles in the sense that different attitudes have different functional roles associated with them. Functional roles are specified largely by defeasible dispositions to interact with other states and mechanisms. We can distinguish two sets of functional roles. There is "input/formation-role" and an "output/motivation-role". For belief the input role is, roughly, the disposition to be caused by evidence or truth-tracking mechanisms (e.g. perception, inference). The output role of belief is usually explained in terms of how belief motivates action.

On Velleman's view, secondary attitudes *also* motivate one to act in the same way as belief motivates. Hence, Velleman concludes, one can't individuate belief only by appealing to its output role. How does he get there? His first step is to point out that belief is the attitude of presenting a proposition as true.⁷⁵ When we believe something, we take it to be true. However, there are other attitudes that 'regard'⁷⁶ their contents as true : assuming a proposition means treating it as true, imagining something to be true is

⁷⁵ I intend attitudes as having propositions as their content. I don't deny there may be non-propositional attitudes, but they won't be the focus here.

⁷⁶ This is metaphorical, since attitudes don't regard their contents as true; people do.

also a way of holding it to be true, and so on. As a result, Velleman proposes to call the attitudes that entail regarding their content as true ‘acceptances’ (I prefer the term ‘cognitive’).⁷⁷ On his view, acceptances include beliefs, assumptions, suppositions, imaginations and the like. To avoid any confusion I will simply call them ‘truth-regarding/cognitive attitudes’ and set them in two groups: secondary cognitive attitudes and belief.

Now, the idea that all those attitudes entail *regarding their content as true* may seem hardly believable. Intuitively, when I imagine that I won the lottery I regard the content of my imagining as a *fiction* rather than as something true. But that’s not what Velleman believes (2000: 183). He explains:

Of course, there is a sense in which things that are merely assumed or imagined are not regarded as really true. But the relevant sense is not that they aren't regarded as true at all; it's rather that they are regarded as true but not really-regarded as true, that is, but not seriously or in earnest. What distinguishes a proposition's being believed from its being assumed or imagined is the spirit in which it is regarded as true, whether tentatively or hypothetically, as in the case of assumption; fancifully, as in the case of imagination; or seriously, as in the case of belief.

So all the attitudes mentioned above ‘regard’ their content as true in the sense that they represent it *as* true (and not as part of the content). They differ because belief represents its content as *really* true, whereas other attitudes represent it as fictionally-true or true for the sake of the argument. As a result, since attitudes other than belief can be

⁷⁷ Note that Velleman’s definition of acceptance is different from how others have defined it. Indeed, in recent years a number of writers have drawn a distinction between belief and acceptance. In his account of acceptance Cohen (1992) argues that accepting a proposition involves committing oneself to taking it as a (explicit) premise in one’s reasoning. He calls it a ‘premissing policy’. More recently, following Cohen Frankish (2004) has argued that acceptances are dispositions to premise only in a restricted sense. That is, when we accept a proposition we are disposed to use it as premise only in deliberations of a certain type. Here I will follow Velleman’s definition.

identified as regarding-as-true attitudes, then the fact that belief is a truth-regarding attitude cannot be used to individuate belief. Velleman concludes that it is belief's *truth-directedness* (i.e. its aiming at truth) that distinguishes it from all other truth-regarding attitudes.

The connection between truth and belief emerges when Velleman defines belief based on its input. It is no news that belief's functional role has an input-side: belief is disposed to be responsive to evidence. By this it is usually meant that belief tends to be *produced* in response to evidence and is *open to review* in the face of contrary evidence. That's not the case with secondary cognitive attitudes, or so it may seem. Unfortunately, defining belief based on its input role may prove harder than expected. More specifically, there may be reasons to believe that belief's sensitivity to evidence is neither *necessary* nor *sufficient* to individuate belief. As for the necessity claim, it seems intuitive to think that *some* of our beliefs are neither produced by evidence nor open to review in light of contrary evidence. An example of that are beliefs recalcitrant to any contrary evidence such as some optimistic beliefs, delusions, and religious beliefs. As a result, it seems problematic to attempt to explain the notion of belief in relations to its input conditions since belief can be insensitive to evidence. Now, one may reply that evidence-insensitive beliefs are just beliefs for which something went wrong. Note that the claim here is that beliefs are *disposed* to be evidence-sensitive, and so if a particular belief does not change in accordance to evidence it is possibly because something went awry, and the belief did not behave as it is supposed to. The problem with this is that some beliefs we have seem to show *no* disposition to respond to evidence whatsoever and so, if one insists that

evidence-sensitivity is required for being a belief, one may find it hard to call those beliefs at all.⁷⁸

As for the sufficiency claim, among secondary cognitive attitudes there is also what one may call ‘scientific acceptances’: cognitive attitudes we hold toward factual propositions that we ultimately don’t believe to be true even though they are in fact backed up by evidence.⁷⁹ For instance, some philosophers of science maintain that we don’t really believe our scientific theories but only accept them.⁸⁰ This is because we know that they will be eventually proven false or at least replaced by other theories. This sub-category of acceptance is in fact formed in response to evidence and is *open to review* in the face of contrary evidence (at least to some extent). Indeed, we tend to hold our scientific theories in response to strong indications that their content is true and typically we would not hold them without strong evidence. If so, then it becomes hard to figure out how these acceptances differ from belief in their input-dispositional role.

To rescue the intuitive idea that there is a special relation between evidence and belief, Velleman argues for a two-level structure explanation of belief: belief is created with an aim. Thus what makes a cognitive attitude a belief and not an imagining, is the fact that belief was created with the aim of truth. This aim belongs either to the agent who has the belief or to the cognitive system responsible for producing that belief. Velleman

⁷⁸ Gendler suggests that these evidence-immune beliefs are not really beliefs but aliefs. (However, see Glüer and Wikforss (2013) for interesting objections against Gendler’s proposal). Currie (2000) argues that delusions are not beliefs (but see Bortolotti 2009). Rey (1988) and Van Leeuwen (2014) focus on religious credences and argue that they are not beliefs.

⁷⁹ Although the notion of evidence is left somewhat unexplained here, I take it is possible to say that E is evidence for H if E is a sign or reliable indicator that H is true even though it can be defeated by other evidence to the contrary.

⁸⁰ van Fraassen (1980) famously proposed a distinction between acceptance of an empirical theory and believing it to be true, arguing that evidence of a theory’s empirical adequacy justifies only accepting rather than believing it.

explains the metaphor of belief aiming at truth in terms of a functional account of belief according to which the essential feature of belief is the function of the mechanism that regulates it and what this mechanism is supposed to do, namely tracking the truth. And so to be a belief is to be brought about by a mechanism whose function is to track the truth. Now, this mechanism may fail and produce a belief that is false and irresponsive to evidence. In this case, such a belief-producing mechanism has failed to fulfill its function. As he claims (2000: 254-255),

Even when a belief is prevented from responding to corrective influences, the fact remains that its regulative mechanisms are being prevented from doing what they were designed to do. A phantasy and a biased belief are alike in that they fail to track truth; but the phantasy has no tendency to track the truth at all, whereas a biased belief is diverted from truth; and something can be diverted from truth only against the background of a tendency to track it.

So in his 2000 book Velleman advances the idea that there is a telos, *constitutively* connected to belief, and that truth-directedness is the key feature of belief. That's his positive thesis.

At the same time, Velleman (2000) has forcefully objected to the possibility of sorting out belief by looking at its downstream output effects. Now, Velleman's argument encompasses three fairly controversial claims and for his argument to constitute a knockdown blow to the project of defining belief in terms of its output role, it must be able to show that *all* these claims are true. *The first claim is that secondary attitudes motivate action. The second is that they do so in the same way as belief does it. As a result: belief's output role is not unique and belief can't be individuated based on it.*

I believe that all of these claims are problematic. I will tackle them in turn. Let's start with the first. Beliefs motivate actions in the sense that they tend to 'combine' with conative attitudes (e.g. desiring) to produce actions and they satisfy those conative attitudes only when they are true. To illustrate, if I desire to drink some water and believe that there is water in the fridge then I am motivated to go to the fridge. My going to the fridge would satisfy my desire to drink only if my belief that there is water in the fridge is true.

What about how secondary cognitive attitudes influence action? There are two major views about this relation. One, say the dominant view, is committed to the following picture. Imagination produces action only indirectly. We imagine and then we have a desire to act on our imagining, and that is what produces action (Nichols and Stich, 2003). In contrast, the second view maintains that imagination produces action directly by combining with some kind of conative attitude, usually called *i-desire* (this view is defended by at least Gregory Currie, Andy Egan, Alvin Goldman, and David Velleman). These are the two competing views at the moment. As I understand it, the underlying disagreement between these two views can be put as follows. One view, the dominant one, says that imagination does not feed in the practical reasoning mechanism. The opposing view says it does.

As for imagination, there are two ways to spell out the role of *acceptances* and the like in producing action. One is to say that acceptances don't really figure in practical reasoning. What is really going on is that when we accept a proposition as true we figure out what would be the case if that proposition were true and we form a related belief. Done that, we decide whether we want to act as if that proposition were true. What guides

us in practical reasoning is not an acceptance, but a belief about how things would be if a given proposition were true. This approach is able to absorb examples like those above within the belief-desire account of action.

In contrast, the alternative proposal says that acceptance figures as a premise in reasoning and combines with a conative attitude of some kind. As briefly mentioned above, in the literature of imagination, philosophers have come up with the idea that, along with real desires, we have pretend-desires or, what they call, i-desires. That I know of, nothing of this sort has been proposed for secondary cognitive attitudes. For simplicity and by looking at the cases of acceptance mentioned above, I assume that, if acceptances and the like enter practical reasoning, they combine with normal desires to give rise to action. So, to illustrate, if I want to save for my child's college and assume it will cost \$500,000 I will be motivated to save \$2600 each month. And so on.

Velleman focuses on imagining and offers a number of examples that are supposed to show that imagining that *p* and believing that *p* are alike in disposing the subject to act as if *p*. First, he argues that the way children act when they engage in make-believe-games demonstrates that imagination has a motivational component as well. What's more, in a game of make-believe imaginings can combine with conative attitudes to produce actions in the same ways as beliefs combine with conative attitudes to give rise to actions in real life. This is Velleman's central example: a child pretending that he is an elephant. Velleman argues that such cases of pretense are cases in which imagining being an elephant disposes the child to behave *as if* he were an elephant. As a result, the child acts on the basis of his imagining that he is an elephant. So here is how things would go in this case: the child imagines that the chair in front of him is a pail of water

and forms a pretend-desire, or i-desire (which Velleman calls ‘wish’). To have a pretend-desire is to enter into a state with a first-order world-directed content whose causal role is significantly like the causal role of desire. The imagining motivates the child to move toward the pail of water and to dangle his arm between his nose and the seat of a chair.

Now, the picture Velleman is sketching by appealing to imagination and the like is at odds with a specific view of how intentional action happens. On this view, what motivates one to act is always a desire-belief combination that causes and rationalizes one’s performance (Davidson, 1963; Smith, 1994). Typically, desires and beliefs rationalize by providing a means-end rationale, and it is usually said that the agent is acting rationally because the action is based on mental attitudes that correspond to an instance of means-end reasoning. On this view, no secondary cognitive attitude motivates action.

So what happens when we imagine and act as a result of it? Can the belief-desire model make sense of this type of action? Nichols and Stich (2000) have focused on action that happens in pretend-play. As I mentioned above, they have argued that the pretending is motivated by a desire to act out one’s imaginings. According to this view, the child imagines to be an elephant, and desires to act like one. So she is motivated to engage in some elephant-like behavior given her initial desire to act like an elephant and her belief about what it means to behave like an elephant. Put it more generally, a child who acts in accordance with her imagining that p does so because she *wishes* to act more or less as she *deems* she would act if p were the case. On this picture, belief rather than imagining plays an actual motivational role.

As Nichols and Stich put it, “[p]retenders behave the way they do because they want to behave in a way that is similar to the way some character or object behaves in the possible world whose description is contained in the Possible World Box [PWB]” (2000: 128). The outcome of the PWB is a conditional belief of the form “If it were the case that p, then it would be the case that q & f, h...”. What causes the child’s pretended behavior is in part the desire to behave similarly to how one would behave if the antecedent of the conditional belief were true. If this is how pretense behaviors are motivated, then it follows that imaginings don’t really have a motivational role of their own: they are never the cognitive input into practical reasoning. It is the conditional belief that motivates the action whenever the subject desires to behave as if the antecedent of that belief were true.

Velleman’s point about imagining (and accepting more generally) stands in direct opposition to this belief-desire model because, on his picture, imaginings can *directly* motivate action when suitably paired with conative states such as desires or wishes. In opposing the belief-desire model Velleman is making *two* distinct claims. For starters, imagining can produce *intentional* actions and not just behavior. Spelling out what an intentional action is has notoriously proven very hard and here I don’t have any positive account to offer. The key point here is that imagination and all secondary cognitive attitudes more generally can offer intentional explanation of action, in the sense they have content which matters in the explanation of the action.

In this chapter, I will take the labels ‘intentional action’ and ‘explanation’ to apply fairly widely. More specifically, I will take intentional actions to include both simple activity and autonomous action, as Velleman defines them. As Velleman (2000) draws the distinction, the latter category is supposed to capture the idea that some of our actions

are the result of our choice to act based on what we take to be reasons to act. In contrast, activity is a behavior brought about *without* any actual choice. What's more, on his view, what distinguishes mere *behavior* from *activity* is that the latter but not the former is the product of our attitudes exerting their causal power *in a normal way* (Velleman, 2000:126).

Now, going back to imagination, Velleman argues that imagination does bring about and explain intentional action, which includes both activities and full-blown autonomous actions. In the latter case, when we play the game of make-believe, we act on our imaginings *because* we so choose. At other times, we simply act out of our imagining, not as a result of a choice but as a result of imagining itself exerting its causal power in a normal way.

This leads us to the second point. Velleman claims that the belief-desire model is *not* the best explanation of the actions that he claims are motivated by imaginings. Importantly, the issue here is not whether we can come up with a belief-desire explanation of those actions as Velleman concedes that such explanations are always easy to concoct. Velleman's claim is that invoking imagination is just a better way to explain those actions.

A number of writers have argued recently that the traditional belief-desire picture of intentional action must be radically revised. A source of dissatisfaction with the belief-desire model is based on the fact that it seems that such a model can't by itself explain some of the things we do. Indeed, there seems to be a number of actions that we regularly do that may be hard to explain just referring to a belief-desire psychology. As a result, recently Tamar Gendler (2008) has presented and developed the notion of – what she

thinks is - a previously unacknowledged cognitive mental state, i.e. alief. On her view, alief is not a propositional attitude but a habitual disposition to react to certain stimuli by acting in a certain way. She offers two paradigmatic examples of how alief produces intentional action and in a way that is in opposition to belief. One example goes like that. Even if we don't actually believe at all that the glass in front of us is contaminated with feces, the fact that a misleading label "feces" is applied to it is actually enough for us to refrain from drinking from it. Here is another case: when asked to walk onto the glass platform of the Grand Canyon Skywalk - even if we are positive that it is perfectly safe - we feel highly motivated not to walk out on it. An important caveat: on Gendler's view all these instances of avoidance behavior would count as intentional actions.

Gendler's main claim is that we face a dilemma. Either we give up the (input) functional account of belief or we accept that there is more to action than belief and desire. The first horn of the dilemma gains support once we notice that whatever is motivating us in avoiding drinking from the glass in the example above, is an attitude insensitive to evidence. Gendler's point is that in the examples above what motivates us is a stubbornly evidence-insensitive attitude, impossible to modify or eradicate, apparently hard-wired and triggered by only certain types of stimuli. As a result, it is hard to believe such an attitude could be belief if we are to keep our traditional notion of belief. There is a further worry. If we were to say that such an attitude was indeed a belief, we would have to accept that, at least in those situations, we were entertaining two contradictory beliefs at the same time. Although it may be possible to argue that this is *not* a psychological impossibility (but just a result of our irrational nature), it is not a position everybody may be comfortable with. So to save our intuitive or traditional picture

of belief, we must accept that what is motivating us is not a belief but a sui generis mental state that she calls ‘alief’. However, if we grant Gendler that there are aliefs, then we fall in the second horn of the dilemma and we have to give up our belief-desire model and the idea that one needs belief to be motivated to act. Gendler recommends that we embrace the second horn of the dilemma.

The literature on alief and imagination is by now quite vast, and I don’t aim at offering an answer to many of the issues that have been brought up there. However, I generally worry about the explanatory gain of explaining action in terms of imagination, acceptance or other attitudes such as alief. It seems there is a perfectly legitimate way to make sense of what is going on in the Grand Canyon above: I believe that the glass platform of the Grand Canyon Skywalk is perfectly safe and I am disposed to act on that. However, I can’t stand heights and so I just do not like to walk on the platform. You can say I have irrational fears, or fears that aren’t cognitively construed, but this is hardly a new phenomenon. So it escapes me what is the explanatory gain we get from postulating the existence of the new mental phenomenon of alief, when emotions would just do the trick.

Going back to Velleman’s point about acceptance and the like, when I accept or suppose something for the sake of an argument, for instance, it is arguably *not* the proposition accepted to be true that motivates me to act or reason in a certain way. What motivates me is the fact that I believe I have strategic reasons for not contradicting you at this stage of the argument. Indeed, I want to show you that, even granting you the premises, your conclusion is wrong. The acceptance itself is not what pushes me to do

anything, it seems. It is my belief that accepting a proposition will foster some goal I have that does the job of motivating.

But what about imagination? Velleman argues against the belief-desire explanation of what happens in various cases of children's pretense.⁸¹ His main point seems to be that pretending is *acting out* of one's imagination, whereas the belief-desires model leaves the pretending child outside the pretense. The belief-desire view, says Velleman, offers a very depressing picture of the child's game as if the child, in her pretense, was spending most of her time strategizing about what it means to be an elephant rather than *acting* as one. But even assuming Velleman is right about this, I wonder whether he is confusing two distinct senses of 'imagining' here. Recall that here I am focusing on propositional imagination, as this is what Velleman seems to have in mind when he talks about imagination. However, as pointed out above there is also a second kind of imagination, e-imagination. Imagination in this sense is a faculty that simulates belief and desires. This faculty is used to make various calculations or understand others' viewpoints, for instance. It can also be the source of the child's pretending to be an elephant when playing a game of make-believe. These belief/desire-simulated states produce choices or decision. So in *this* sense imagining does produce action but only because imagining means simulating one's belief/desire states (Goldman, 2006). This clearly contrasts with Nichols and Stich's approach to how imagination gives rise to action, and solves Velleman's worry that their approach was excluding imagination from action-production. But this is not the sense of imagination Velleman is concerned with. And it is not instrumental to Velleman's threat to the belief-desire model

⁸¹ Others have argued against this model. See, for instance, Currie and Ravenscroft (2002).

since e-imagination motivates because it re-enacts the belief/desire pair. But on Velleman's view imagination is a different mental state than belief and not merely a way to re-enact a belief. He clearly states that the state of imagination that motivates the pretending children is similar to a supposition, and presents its content as true. This is why, he maintains, it is able to produce an action. But it is unclear that this is the case, as I just suggested. The type of imagination at play in pretended play could be described as a faculty that allows us to re-enact beliefs and desires to produce actions. If so, Velleman's attack is off target.

HOW ATTITUDES MOTIVATE ACTION: REPLY TO VELLEMAN (PART 1)

In this second part of the chapter, I will grant Velleman that what motivates pretense is usually not a belief but an instance of imagining. In other words, for the sake of the argument I am prepared to accept that imaginings *do* motivate intentional action. I will also grant him that the same extends to other belief-like attitudes, acceptances, suppositions and the like. Importantly, though, that by itself is not enough to establish that these cognitive attitudes have the *same* output role as belief. In fact, I don't believe that they do. My position is that belief is an 'unqualified' cognitive attitude whereas all secondary cognitive attitudes *must* come with some kind of constraint on their use that limits the type of inferences or reasoning patterns they can enter into. A qualified attitude is an attitude that needs, to enter reasoning, an attitude that is unqualified.

Before laying out my proposal, let me flesh out a possible way to reply to Velleman's claim that imagination motivates action *in the same way as belief*. The proposal is that belief and imagination share a similar motivational role but imagining is

context-bound in a way belief is not. And this can be extended to the other secondary cognitive attitudes. In other words, whereas belief motivates in all situations, secondary attitudes' motivational role is restricted to specific contexts and is activated *only* when such contexts are represented in some way. This proposal has the merit of making sense of some of the intuitions that back up Velleman's argument, while discounting his final conclusion. Regardless of its strong plausibility, I worry Velleman may be in a position to reject this line of argument. As the final step, I will present my argument for why *pace* Velleman belief's output role is in fact unique.

The claim that imagination motivates action in the same way as belief has struck many as frankly implausible. Even if I imagine that I am a bird who wants to fly, many would say, I am definitively *not* motivated to jump out of my window and fly out. That idea simply does not cross my mind! As a reply, Velleman has pointed out that we should look at what happens in games of make-believe and other cases of actions putatively produced by imagining. But if Velleman wants to have a chance of getting his story off the ground, he needs to make sense of the following important phenomenological point:

Explanandum: in many cases merely imagining, supposing, accepting something to be the case does not seem to motivate one to do anything (even when paired with the relevant i-desire)

This can be rephrased by saying that, even if we actively imagine such and such, that bit of imagination (paired with an i-desire of some sort) does not seem to show up as a possible source of action (unless perhaps something else happens).

A proposal has been recently brought forward that tries to make sense of the role of imagination (and other cognitive attitudes) in producing action. This proposal is able to make sense of Explanandum while maintaining that imagination does produce action.

The proposal is simple. It grants Velleman that there are indeed a lot of similarities between imagination/acceptance and belief with respect to action, but it also points to the intuitive idea that imagination, assumption and the like can motivate only in the contexts in which they were introduced. Let me try to be a little more specific. By ‘context’ I mean the situation in which the acceptance, imagination and so on was formed. To illustrate, ‘a game of pretense’ is the context in which I imagine I am an elephant and that the chair is a pail of water; ‘work’ is the context for accepting that my client is innocent; ‘my current paper’ is the context for my granting Velleman’s that imagination motivates. And so on. Importantly, though, context does not necessarily (only) indicate a time-place relation but also a set of *inferential moves*. Thus, context may refer to a subset of the (logic and material) inferential relations a proposition may give rise to. As a result, if I accept a proposition for the sake of an argument, the argument in question does not encompass *all* the propositions that follow from what I accept, but only a *subset* of those propositions. Similarly, if I imagine I am an elephant I may imagine I have a trunk but other things may just fall outside the context of the fiction or game, e.g. that I can trump or that I weigh 7,000 kg, and so on.

For instance, O’Brien (2005) has pointed out that, granting there may be some similarities in the way the child reacts if she imagines she is an elephant and the way she would act if she believed that she were an elephant, there are also some striking differences. Believing you are an elephant presumably means pursuing a elephant-like

life, i.e. living outdoor in the Savanna, trying to communicate with other elephants and consuming “elephant edibles” (Noordhof, 2001: 253). No such thing happens when the child imagines she is an elephant. Thus, there are some things we are disposed to do when believing that are not shared by imagination. What’s more, just imagining something does not motivate us accordingly. If I am the office, even imagining I am the Queen/King of England does not motivate me do anything with that. This is because I don’t see my office as the right setting for me to act on my imagining. Differently, when I am playing with my daughter, imagining may motivate me to act in various ways. Imagination is, thus, restricted to a context: it motivates us only in some specific situations and with respect to a limited set of possible actions.

So here is the same idea put in different words. When I imagine that *p*, I *treat* *p* as true and I am disposed to act as if *p* were true *only* if I insert an extra premise that represents the context as being such and such. Hence, imagining that *p* is not enough to motivate me to do anything or, put it differently, imagining that *p* is not enough to be disposed to treat *p* as a premise in practical reasoning. Imagination per se does not access the practical reasoning mechanism. It requires an extra-input: a representation of the context as fictional.

When I am watching or reading *Othello*, I imagine that Othello will kill Desdemona, and I desire to avoid that. However, this imagination + desire combination does not motivate me (to pretend) to stop Othello from murdering Desdemona. I am not motivated to pretend to do anything in this context. Why? Well, a possible suggestion goes as follows: although I am reading fiction, I am myself not *in* it, but I am simply a spectator of it. As a result, imagination does not motivate me to do anything. In contrast,

when I am playing a game of ‘cops and robbers’, I am in the context of enacting that particular fiction. That means I have

- A) the belief that such a game consists of such and such actions, moves, inferential patterns, and so on
- B) the belief I am in that specific game

When I read a work of fiction, in contrast, I have A but not B. Say that, instead of being a simple spectator, I am an actress acting in a theatrical representation of *Othello*. In that case, one may say, I am actually participating in the (fictional) story, I am making it, I am part of it. I know I am not Desdemona but it is possible that (good) acting requires me to *imagine* being her. When acting, I also have the belief that I am enacting the fictional story *Othello*, and that I am in the context of that theatrical representation. In *that* case, the belief about the context I am in, paired with my imaging of being Desdemona, may actually motivate me to act as if I were Desdemona.

But if this view is right, then it is bad news for Velleman. For two reasons. One, belief is not bound to context. It is not that we believe one thing at home and its negation at work. So belief and those other cognitive attitudes can be set apart just by looking at how they behave with respect to context. They’ll share the same motivational role, but only in a limited, narrow sense: whereas belief’s functional role is a *default* disposition, imagination and acceptance don’t motivate by default but only in relation to a context.

As a result, we have now an **asymmetric relation** between belief and secondary cognitive states (Van Leeuwen, 2009). That is, there is an *asymmetrical relation* between imagining, acceptance, assumption on one side and belief on the other, such that imagining and the like need belief to be able to motivate action, whereas belief does not

need imagining to motivate action. This is because, as we saw, imagining's ability to motivate depends on the fact that there is another attitude which represents the context as the right one. Thus, my imagining to be a Queen motivates me when playing with my daughter but not when I am at home doing laundry: only in the context of the game does the imagining enter my reasoning as a possible promoter of action. In addition, to be able to motivate, imagining belief to set the context (Van Leeuwen, 2009). To reiterate then, the view proposed is that the output role of belief must be different from the output role and downstream consequences of secondary attitudes because there is an asymmetrical relation between the two categories: secondary attitudes need belief to be able to motivate, whereas the opposite is not the case.

As a result of this, if Velleman accepts this view, he so-to-speak wins the battle but loses the war. That is, in adopting this view he has a way to show that there is in fact a clear parallel between belief and secondary cognitive attitudes, but this parallel holds only because there is a larger, more significant disanalogy between these two categories.

VELLEMAN'S REPLY

Velleman separates the rational import attitudes have on other attitudes and on actions from their motivational role. This allows him to argue that secondary cognitive attitudes can motivate beyond the context in which their influence would be deemed rational (or at least reasonable). On Velleman's view, the fact that imagining rarely produces action is *no* evidence that imagining is not so disposed, but is evidence of the fact that there must be something that often *countervails* imagination's motivational power (Velleman, 2000: 272). So Velleman may happily grant his critics that imagination

seems context-dependent while denying that it actually is so. He has an error-theory for why things seem this way. It is just that in many situations there are *countervailing forces* –we are largely unaware of – that stop us from acting on our imagination. But that does not mean that, absent those forces, we would not act on what we imagine to be the case.

To be more specific, Velleman individuates three elements that may play a role in stopping imagining from producing action. First, he mentions that we are often *inhibited* from acting on our imaginings. More specifically, as adults we learn to inhibit the motivational push of what we imagine, and this inhibition explains why adults are not very good at playing games of pretense. More importantly, invoking this inhibition can be a way to explain why we often don't act on our imagining.⁸² There is a second element that often prevents us from acting out of our imaginings. Usually, when we imagine that p we also believe that p is false. As a result, this contrary belief exerts its own motivational power that is presumably contrary to the one exerted by what we imagine.

Clearly, the fact that beliefs can act as counterbalancing forces is not a small concession: saying that beliefs can counterbalance imagining is tantamount to admitting that there is something beliefs can do that imaginings can't. However, at this point this reply is insufficient. It is not clear that only beliefs play the role of counterbalancing forces. Velleman never says that and I doubt he believes it. And even if true, that does not show that belief and imagining are not the same with respect to their motivational role. It only shows that they differ *somehow*.

⁸² I find Velleman's talk of 'inhibition' slightly unsatisfactory. A reason for that is that he says very little on how inhibition is supposed to work. He also does not say anything on whether such inhibitory force exerts its power over all acceptances or only over imaginings. For the time being, I will set this issue aside, but also plan to come back to it when I offer my positive argument.

Finally, Velleman could argue that the reason why imagination and belief seem to have different upshots has nothing to do with beliefs and imagination *per se* but with a *third* element that makes it the case that we either act on what we believe or on what we imagine. It would be quite challenging to try to offer a complete account of Velleman's theory of action here, but let me point to some key elements of it (which I have already hinted at above). On his view, our intentional *autonomous* actions are the result of a choice, which is an additional capacity that goes beyond the motivational power of whatever attitude may cause the action. This capacity *reinforces* the motivation force exerted by some specific attitudes by also blocking any countervailing attitude. The notion of choice at play here is the same as a 'practical cognition', and explained in terms of a second-order motive that influences first-order motivates (Velleman, 2000: chapter 8). That means that even when there are two opposing attitudes exercising their motivating force, we usually subscribe only to one of those motivating forces by putting additional force on it. The attitude we subscribe to, Velleman maintains, is the one that – in our light – makes sense and makes our actions intelligible (Velleman, 2000: 96).

It would take me too far afield to try to critically engage with Velleman's account of action here, and so I think it is only fair for me to try to understand the implications of his view with respect to the motivational role of imagining. On that note, can Velleman's view about action explain what is going on in the two different situations in which I imagine to be Queen? Possibly. Note that it is plausible that in both scenarios what I do may be the result of a choice. If Velleman is right about choice, then in both cases I am acting based on what attitude I choose to be motivated by. When I am playing with my daughter, I choose my imagining as a source of action and act accordingly. When I am

doing laundry: since acting on my imagining would make very little sense, I act on what I believe. Crucially, though, that does *not* mean that my imagining is not able to motivate me in both cases. It simply means that I made a choice to not be so motivated.

Does Velleman have any positive evidence to support his view that imagining and other similar attitudes motivate by default? Remember we are granting him that secondary attitudes can motivate action. Based on that assumption, yes, I believe he might produce two sets of evidence to support his view. **First**, Velleman offers examples in which imagining motivates outside imaginative contexts and without the help of belief. That breaks the alleged asymmetric relation between imagining and believing. Here is one. As I am walking down the street, I start imagining that someone has stolen the wallet that is in my pocket. Instinctively, I reach for my pocket to check that my wallet is still there. According to the view sketched above imagination can motivate us only *if* some belief represents the context as the ‘practical setting’ of pretense or make-believe. But clearly this does not happen in this case: I check my pocket because in fact I don’t believe this is a situation of pretense⁸³. Or similarly: imagine that I am watching a soccer match on TV and start yelling at the referee. Perhaps, my behavior in this case does not count as a full-blown intentional action⁸⁴, but it is surely close enough to intentional action to allow for explanation by means of cognitive states with representational contents. And so, arguably, what is going on here is that I *desire* that the referee would change his decision to assign a penalty and *imagine* that he can hear me and so I start shouting at him without, perhaps, even realizing it. If so, then my imagining is able to

⁸³ Other examples of this kind have been discussed by Gendler (2008) as cases of action motivated by ‘aliefs’.

⁸⁴ In fact for Velleman a full-blown intentional action is only an action that results from a choice, what he calls ‘autonomous action’. As I mentioned above, this is different from mere activity which is still intentional but not autonomous.

motivate me to act as if its content were true, even in contexts that have very little to do with pretense. My disappointment with his decision, my wish to change his mind and my imagining cause me to act – but the action takes place in a real setting, not a setting of make-believe, i.e. not a setting that I conceive of as ‘pretense’.

Velleman is not alone in thinking that there are actions like these, called “a-rational actions”, actions that seem pointless as they do not fulfill any specific goal, and don’t seem to satisfy any desire. Typical examples of these actions include mussing up the hair of one’s child, scratching the eyes in a photograph of someone, or banging one’s fist on the table. Hursthouse (1999) believes these actions to be expressive of an emotion. Velleman (2000: 269) points out that the explanation of these actions involves some kind of imaginings too. Be as it may, the point here is that it seems that imagining (or some other attitude that is not belief) can motivate *beyond* pretense and make-believe contexts and without the need to represent context in any way.⁸⁵ On this view, the reason why imagining in fact rarely causes action across contexts and, when it does it, it is only for a short moment, is because of some opposing force (e.g. conscious choice, inhibition, a countervailing belief) that stops or weakens imagining’s defeasible motivational push.

A-rational actions aside, there is a **second** element that puts some pressure on the idea that what takes belief and other cognitive attitudes apart is context-dependency. Some secondary cognitive attitudes are reasonably used across the board and need no belief to motivate. In fact, examples of *default* secondary cognitive attitudes abounds, although they are seldom recognized as such:

⁸⁵ I don’t mean to imply here that Velleman’s diagnosis of what is going in these cases is not controversial. For a skeptical look at his view see Raz’s point that, in cases like those I mentioned above, it is far from clear that the agent is producing an intentional action at all. And even if one were to concede that those are intentional actions, these are still borderline cases of intentional actions.

PHYSICS: Gina has only a very superficial knowledge of physics as she rarely encounters situations that require a thorough understanding of it. She holds that the general theory of relativity is the correct theory of gravitation. Even so, in everyday situations she assumes or accepts the much simpler, Newtonian theory of gravitation. She knows it to be false, and thus doesn't believe it. And yet this is what she commonly assumes to be true across contexts.

CORNERSTONES PROPOSITIONS: "I am not dreaming"; "I am not a brain in the vat". These are propositions that are not contents of beliefs but of non-epistemically warranted acceptances. These are, in other words, not things we believe but that we *accept* to be true (Wright, 2004).

MATH FICTIONALISM. When we normally engage in mathematical discourse we presuppose that mathematic entities exist (Yablo, 2002a,b; 2005).

What makes these attitudes secondary cognitive attitudes? They are attitudes towards propositions not believed to be true, and they can function as premises in reasoning, guide and motivate action. Interestingly, though, although some secondary cognitive attitudes are attached to specific contexts (e.g. work, argument), the examples above show that some are operative by default and its members can enter our reasoning *without* needing any context-setting representation. Thus, Gina accepts Newtonian physics across the board and usually uses it whenever she has to do some calculation. She does not have to reflect on it, it simply what she does. For her Newtonian physics is not supposed to be restricted to some situations, but permeates whatever she does: it is her default motivating attitude.⁸⁶ Why? Mainly for practical reasons as it requires much

⁸⁶ Dub (2015:11) has offered similar examples and noted that it is unclear why we could not imagine that some of our suppositions or acceptances actually take over our entire life and are never really "discharged".

easier calculation than the general theory of relativity, which she holds to be the correct one. CORNERSTONES PROPOSITIONS is also a clear example of an acceptance used by default, if any. Wright (2004) has quite forcefully argued that propositions such as “I am not dreaming”, “I am not a brain in the vat”, are not things we believe, even though we may *take* ourselves to believe them. Something similar can be said for mathematical entities, at least according to Yablo. What these examples show is that perhaps Velleman is right in thinking that secondary attitudes motivate by default, without the need of any context-setting representation and thus without belief. When this is not the case, it is because other forces stop them. But these forces may not be always active, as we just saw. Notice that this is also true for belief: sometimes we don’t act on some of the beliefs we have because other forces countervail them. If this is right, then it is unclear that there is any asymmetrical relation between secondary attitudes and belief.

BELIEF AS OUR UNQUALIFIED ATTITUDE: A REPLY TO VELLEMAN (PART 2)

In this final section, I will show that belief’s output role is not shared by other cognitive attitudes such as imagination and acceptance, not even when used by default. Before I start, let me point out that here I concede that the motivational role of belief, narrowly construed, may be shared by secondary cognitive attitudes. That is, I grant that imagination and the like may well motivate action in the same way beliefs does. And yet, I conclude here, that does not mean that they share the same output role *in general*.

Importantly, the force of my argument hangs also on whether it is able to show that belief has a unique output role without presupposing the idea that belief is directed at truth. Recall that Velleman’s starting point was that what makes belief distinguishable

from secondary attitudes is belief's relation with truth. According to Velleman (2000) that meant that belief is necessarily formed by a mechanism whose function is to track the truth. Velleman's point is that belief's truth-directedness is the condition of possibility for belief's role in reasoning. In contrast, I will now show that no appeal to the notion of truth is needed to explain belief's role in the mind.

THE REGRESS STOPPER

By their very nature, secondary attitudes are *contextualizable* and thus compartmentalizable. So it is a *condition of possibility* for these attitudes to exist that there is an attitude that is *not* contextualizable and thus non-compartmentalizable. This is what I mean. It is possible (and legitimately so) to both accept a proposition in one context and its negation in a different one. It is possible to imagine one thing and its opposite, at least in different contexts or fictions. So it is possible to imagine that Superman is bald in one fiction or story and that he is not in a different one. Similarly, one needs to acknowledge that it is possible to accept, suppose, and assume different things at the same time in different contexts and that we often do so. Nothing prevents that psychologically: adults seem perfectly able to keep in mind multiple different possible scenarios. And that also seems a perfectly rational thing to do. Importantly, though, we can accept contradictory propositions only if these belong to different contexts. Put it differently: we can't employ contradictory propositions (of the same type of attitude) as premises of the same reasoning patterns at the same time.

Importantly, we should not misunderstand this point and think that the fact that secondary attitudes shift from context to context is what makes them 'contextualizable'.

That is, it is not my position that the reason why secondary attitudes vary with respect to context is that in one context one has (usually pragmatic) reasons to accept, imagine, or suppose a proposition, and in a different context one may lose those reasons and acquire new ones, and thus come to accept, imagine or suppose some opposite proposition. This, I claim, happens with belief as well; that is, (reasonable) beliefs change according to the (epistemic) reasons one has. To illustrate: If I have sufficient evidence that *p* is true, I believe that *p*. Were that evidence undermined by some other consideration, I would stop to believe that, and if I acquire new (contrary) evidence I may end up believing the opposite of what I initially subscribed to. So if context-dependence ultimately means sensitivity to reasons, then belief is just as much context-dependent as acceptance is. But that's not my point.

What makes secondary attitudes contextualizable is that I can, *at the same time* and in relation to different contexts and patterns of inferential relations, imagine, suppose, and accept that *p* and that not *p*. I can, in other words, keep both prospectives in mind at the same time without confusing myself or feeling the pressure to eliminate one of them. That this is possible is a key feature of secondary attitudes as a class of attitudes. That means that not only I can imagine that *p* and suppose that not-*p* at the same time; it also means that I can imagine that *p* and imagine that not *p* at the same time. I can hold two contradictory pieces of imagination at the same time.

Secondary cognitive attitudes are 'qualified' in a further sense. I may accept a proposition for the sake of the argument I am having with a friend, but the set of propositions I may end up accepting based on that initial acceptance is constrained by the reasons for which I form that initial acceptance in the first place. Schematically, if I

accept that p and proposition 'f' is entailed by 'p', it seems rational to accept that f only if 'f' is somehow relevant to my accepting p in the first place. That means that the set of the inferential relations in which 'p' is set is not necessarily identical to the propositions one is disposed to rationally accept having accepted 'p' in the first place. So it is always in principle possible to accept a proposition without accepting its entailment. This even more true for imagining, where one can restrict the number of propositions imagined to a small subset of the implications of what one imagines. Again, this is what makes acceptance and secondary cognitive attitudes more generally 'contextualizable', i.e. applicable to a subset of all possible inferential relations and situations. As I said, this feature is, I believe, part of the nature of these secondary attitudes.

The fact that secondary attitudes are always potentially dependent on contexts and confinable to a subset of inferential relations makes it the case that there must be an attitude that is, as a whole, not subject to restrictions (modulo some specific cases I will tackle below). This attitude is the condition of possibility of secondary attitudes, because it is able to draw the limits for them, quarantine them and set them apart while being in itself limitless. This is our primary cognitive attitude.

The phenomenon of quarantining can account for the fact that usually what we imagine, suppose, or accept does not influence action outside appropriate contexts (what I called 'Explanandum' above). Quarantining also prevents imagination and the like from contaminating our beliefs: if I imagine of an object that it is a glass of water, I will *not* come to believe it to contain a wet, transparent and thirst- quenching liquid. Importantly, as I mentioned above, that does *not* mean that quarantining always prevents us from forming a belief based on acceptances and the like. If I know that a scientific theory has

good predictive power, I can assume it to be true and derive predictions and thus beliefs from it. Quarantining can allow that. In contrast, contagion happens when this is true: secondary attitudes influence action and belief *outside the context in which they were introduced*. To illustrate, if I imagine that a chair is a pail of water in the context of a game of pretense, *that* imagining should not influence what I actually believe, and so I should not come to believe that the chair contains a wet, transparent and thirst-quenching liquid. And if I do that, then that's an example of contagion.

Finally, quarantining prevents between-contexts permeability, that is, any carrying overs from one fictitious context to other contexts. For instance, imagine a lawyer who accepts that his client is telling the truth about that night but at some point - while working on her case - also assumes that he is lying, and that he was not where he says he was. She does that to anticipate possible moves of her opponent in court, and thus come up with a good defensive strategy. To be able to do that, she has to keep in mind the implications of these nested assumptions while making sure they don't overlap with other assumptions she makes. Although this trick probably requires some effort, it is plausible that she has the cognitive resources to do it. If so, then there has to be a way to temporarily quarantine one set of assumptions from both her beliefs and other possible sets of assumptions that she may be holding at the same time.

There is a similar common phenomenon: we often prevent producing the expectation that real-world objects and events have the characteristics they are supposed to have in the context of games of make-believe or in the context of supposition. And evidence shows that this quarantining ability is present even from an early age. As Gendler points out, children who are playing a game of make-believe "readily deny that

pretend cookies are edible, and show serious dismay if an experimenter tries to bite one with the apparent intent to consume it. They [...] do not they expect that if the red blocks represent sandwiches in one game, and bars of soap in another, that the sandwiches will somehow be soapy or the soap-bars somehow edible.” (Gendler, 2010: 241) What’s more, in normally functioning agents we don’t see cases of massive contagion.⁸⁷ We all experience some cases of cognitive contagion, perhaps more than we would expect, but I think it is fair to say that quarantining holds it in check for the most part. And likely so, since an extreme tendency towards contagion tends to be associated with systemic forms of psychopathology such as schizophrenia.

So let’s go back to my initial argument. To fulfill its function, our quarantining mechanism needs to be able to represent the context in which a given acceptance was introduced, the inferential relations that compose such context, and so on. That is, to do its job and prevent contagion our quarantining mechanism needs, *at a bare minimum*, to hold a representation of the context in which a specific acceptance was introduced. Thus, to quarantine an imagination *I*, the quarantining mechanism has to represent the context in which *I* was introduced. Without such a representation, quarantining is not possible.⁸⁸ The context-representing attitude that is employed by quarantining must not itself be open to quarantining, otherwise a regress would potentially ensue. So to avoid regress secondary cognitive attitudes *cannot* have the function to ground quarantining. There must a primary attitude in charge of that. This attitude is not contextualizable, and its influence can in principle be extended to all the inferential relations related to its content.

⁸⁷ We have more cases of emotional contagion. See Gendler (2003), for a discussion of the relevant literature.

⁸⁸ Contexts can be represented by propositions such as “this is where I assume that...”, “here I play the game of make believe that...”, or simply “now I am at work...”.

REASONING BEDROCK

Besides being a regress-stopper, our unqualified, primary cognitive attitude is potentially needed to ground the use of secondary attitudes in reasoning. As I will argue in the next chapter, reasoning is an inferential process guided by the aim of reaching a conclusion supported by normative reasons. There are two types of premises in reasoning: premises that play the function of being reasons and premises that do not have that role. Contextualizable attitudes are the latter. Given secondary attitudes' contextualizability, it is always possible that their ability to work as a legitimate premise is limitable to a specific set of situations and inferential relations. As we saw, it is indeed part of the nature of secondary attitudes that it is always *possible* to limit the scope of their influence by adopting an attitude of the same kind with *contrary content* or by limiting their inferential downstream effects. I may assume that p and assume that not-p without contradiction, if I do so in different contexts. The fact that this can happen without contradiction means that I will not be forced to eliminate one of the two opposing assumptions: they are both legitimate assumptions.

As a result, their status as premises is always dependent on something else that grounds their use in reasoning. That means then that their role as premises cannot be that of a reason (they do not carry justificatory value by themselves), and thus they always potentially need a further attitude to ground their role as premises. So once again, there has to be attitude that is not contextualizable and whose role as a premise is *used* as legitimate by default in the sense that we *treat* it as a legitimate premise automatically. This attitude functions, in our psychology, as a bedrock premise in reasoning, an attitude

we don't need to justify *qua* premise. And this is again the unique role our primary cognitive attitude has.

OUTPUT ROLE IN THEORETICAL REASONING

I will now address the question of how secondary attitudes behave with respect to theoretical reasoning, and show that the downstream effects of secondary attitudes must necessarily be different from the downstream effects of the primary cognitive attitude mentioned above.

I define theoretical reasoning as the kind of reasoning that employs only cognitive attitudes as premises and, as a result, produces a cognitive attitude as a conclusion. Cognitive attitudes are attitudes that represent their content as true. Acceptance, imagination, supposition and the like are all cognitive attitudes.

It is common among philosophers to maintain that the functional role of belief in theoretical reasoning involves a disposition to conform to material and logical inferential relations in accordance with requirements of formal rationality (coherence requirements). Believing that *p* means being disposed not to believe that not-*p*. Believing that if *p* then *q*, while also believing that *p* means being disposed to believe that *q*. And so on.

Do secondary attitudes tend to behave in the same way as belief? Are those disposed to accept that *p*, and that *p* implies *q*, disposed to accept that *q*? It does seem that imaginings and acceptances have properties very similar to beliefs', and behave inferentially as beliefs do. If I accept that my friend is innocent, it is plausible to say that I am also disposed to refuse any proposition that goes against that claim. If a lawyer accepts her client's explanation of where he was that night, she is also disposed to accept

what she takes to follow from it and disposed not to accept any contrary proposition (in that context). And finally, “if I imagine that my house is on fire, and imagine that if my house is on fire I’m in danger, I’ll be imagining that I’m in danger” (Sinhbabu, 2013: 160–161). Indeed, it is plausible that we are in fact disposed to (imaginatively) subscribe to the consequence of the propositions we imagine (when we are rational we do so within the boundaries of the context our imagination is placed in).

So it seems that being disposed to conform to requirements of formal coherence is a characteristic of all cognitive attitudes. This symmetry is often called ‘mirroring’. It has been studied specifically in relation to pretense and imagining but I assume it can be extended to all sorts of secondary attitudes. Mirroring says that if I imagine, pretend or assume that p, and “if I am tacitly or explicitly attending to my belief that if p then q, then (*ceteris paribus*) I will be inclined to” pretend, imagine, or assume that q (Gendler, 2006: 184). Mirroring is probably due to the fact that our inference mechanisms treat pretense (and other non-doxastic cognitive attitudes’ content) in “roughly the same way” they treat real belief’s content (Nichols and Stich, 2000: 125). So I will assume that, given their dispositions, various secondary attitudes will tend to conform to requirements of rationality, be closed under entailment, and by and large end up being coherent.

This tendency, however, will emerge only locally, i.e. only within the various contexts secondary attitudes are placed in. That means that globally, they will in fact emerge in patterns that are incoherent. This is clearly a result of their contextualizability. This is what I mean. At the local level, i.e. within the context of the limited inferential relations they are usually placed in, a set of supposed or accepted propositions will converge toward coherence (modulo irrationality, of course). The examples mentioned

above show that. However, given that across different contexts we may accept many different and at times diverging propositions, the set of *all* the propositions we accept stands in no relation of coherence. This is true also for the propositions we imagine. They are also, as a whole, incoherent. More generally, that means that secondary cognitive attitudes, as a class, are not disposed to accord with the requirements of coherence (or formal requirements of rationality).

In contrast, our primary attitude will be *globally* convergent toward coherence (modulo irrationality). This feature is the result of its not being, as a class, contextualizable and is clearly related to its functioning as a regress stopper and as a background condition for the existence of secondary cognitive attitudes. So, our background attitude will be functionally distinguishable from secondary cognitive attitudes also with respect to its output role in theoretical reasoning. Not at the local level, not at the level of individual attitudes or subset of attitudes, but globally. Hence, the primary cognitive attitude, as a whole, converges toward a coherent set; secondary attitudes don't.

IS BELIEF OUR PRIMARY COGNITIVE ATTITUDE?

I previously showed that for secondary attitudes to exist and to function as they do in our psychology, there has to be a category of attitudes that are unqualified, unconstrained, and function as the background for those secondary attitudes. It is now time to argue that this primary cognitive attitude is belief. To do that, I claim that belief is a plausible candidate for being our primary cognitive attitude because belief does not

share some of the key features of secondary attitudes. I will also discuss possible objections.

For starters, it seems to many that belief does not admit the possibility of holding contradictory beliefs. Some have argued, however, that it is not *impossible* for us to harbor contradictory beliefs. Take the following example presented in Lewis (1982): a resident of Princeton has the following three inconsistent beliefs: he believes that Nassau St. runs North-South; he believes that the railroad tracks run East-West; and finally he believes that Nassau Street and the railroad tracks run parallel to one another. What's going on here is, according to Lewis, that the agent can harbor contradictory beliefs: one is that the street and the tracks run North-South, and the other is that the street and the tracks run East-West. How is that possible? The idea is that the agent fails to bring those beliefs to bear, he simply does not see that they are in contraction because he does not activate them at the same time, in the same contexts. As a result they both get to be activated and used in reasoning, albeit in different situations. More generally, what this example illustrates is that the idea of actual human beings having a perfectly coherent system of beliefs that guides our actions and reasoning across contexts is in fact a myth. As Lewis (1982) puts it, our beliefs are compartmentalized and drive different aspects of our practical and theoretical reasoning in different contexts (Stalnaker, 1984; Egan, 2008).

When this happens, though, it is mostly the result of a cognitive failure of some kind which has to do with specific beliefs and not with the category of belief as a whole. That is, it may be possible to harbor contradictory beliefs if we fail to bring them to bear, i.e. if we don't see that they co-exist or we don't activate them at the same time. But this

is a malfunction of our cognitive system. In contrast, as we saw, there is nothing wrong with contradictory imaginations, suppositions, acceptances and the like. There is no failure in having contradictory assumptions, as long as we are able to keep them apart and relate them to their original context. So belief and secondary attitudes differ on this.

Second, belief seems to fit the description of an attitude that is not open to be voluntarily quarantined. No other attitude seems to fit the bill. This raises the following worry, though: imagine I believe my brother stole a lot of money from me, but my therapist is able to convince me that that belief is the result of a Freudian mechanism of some kind. I thus come to treat that belief as illegitimate premise and try to isolate from reasoning as much as I can. In this case, one may say, I'd use quarantining for belief as well, so what is the difference with imagination here?

First, I am not sure separating this belief from reasoning counts as quarantining at all. Quarantining mechanisms prevent various attitudes to show up as possible premises of reasoning. If something is to count as a belief, but never shows up and presents itself as possible premise in any reasoning, then it can hardly be said to be a belief at all. And even assuming that isolating one of my beliefs does count as quarantining, one notable feature of this case is that such quarantining is *not* dependent on a context. It is not that I quarantine a belief in some context and not in others. If a belief fails as a premise, it fails in all cases. So, if quarantining occurs and compartmentalization is possible for belief, it is still not the same thing that happens with imagination and the like as it does not require context-setting.

In addition, even assuming that some beliefs require quarantining, quarantining is not a condition of possibility for having beliefs in the first place. I argued above that

being open to quarantining is necessary for an attitude to be a secondary attitude. No such thing is true for belief. Indeed, we can imagine an agent who lacks the psychological mechanisms necessary for quarantining. That agent is in a position to have beliefs but no secondary attitudes. So belief can still be the category of attitudes that we treat as legitimate in reasoning, and that can function as bedrock in our reasoning system.

And finally, a belief can be quarantined because it is defective, but it is not quarantined *qua* belief. In contrast, a secondary attitude's quarantining is *not* the result of a defect in the attitude but the result of a judgment on our part that that attitude is not fit or relevant in a certain context. Based on the example above, take the following problematic piece of reasoning: "Since my brother stole my money, I will not invite him over for Christmas". The reasoning here is not bad *per se*, it is the choice of the premise that is problematic because of the epistemic status of the belief in question: I have no evidence for it. Thus, although defective, that piece of reasoning is not defective *in the same way* as reasoning from *imagining* that I am the Queen/King of England to claiming and believing that I live in Buckingham Palace. This last bit of reasoning is *absurd* because we think imagining to be the Queen/King of England should not influence beyond the context of my daydreaming.

The result of all this is that we can legitimately expect that beliefs will converge toward forming a coherent set or at least a set of beliefs that would look very different from the set of secondary attitudes an agent may have. As I pointed out above, secondary attitudes tend to form locally coherent sets, but at the global level we can expect them to be incoherent, contradictory, not closed under conjunction, etc. A different behavior has

to be expected from belief in line with the fact that belief is not contextualizable. So the downstream consequences of belief and secondary cognitive attitudes are bound to differ.

CONCLUSION

In this chapter, I tackled Velleman's claim that belief and secondary cognitive attitudes, e.g. acceptances and imagination, share the same output role. *Contra* Velleman (and others), I strongly doubt that secondary cognitive attitudes can produce action. And even if they do, most secondary cognitive attitudes' influence is limited to some specific context and situation. Most importantly, there are some inferential patterns imagination and the like cannot enter into anyway: they can't function as regress stoppers and as bedrock in reasoning. As a result, the role of these attitudes is necessarily *qualified* or restricted in some way. In addition, when involved in theoretical reasoning, each secondary attitude's behavior, as a whole, must substantially differ from belief's. Therefore, belief is functionally distinguishable from secondary cognitive attitudes also with respect to its output role in theoretical reasoning.

BEYOND RULE-FOLLOWING: REASONING WITH PURPOSE

Chapter 3

The goal of this chapter is to propose a new constitutive account of reasoning. First, I show that the widely endorsed rule-following view of reasoning is insufficient to make sense of what reasoning is. As an alternative to that view, I defend the claim that reasoning should be seen as an aim-directed activity. Finally, I argue that engaging in reasoning requires having the aim of arriving at a conclusion supported by at least sufficient reasons. This Reason View is well positioned to make sense of some key features of reasoning, and should be preferred over all its competitors.

The topic of reasoning has attracted a lot of attention lately⁸⁹. Besides being interesting in its own right, an answer to the question ‘what is reasoning?’ promises to shed some light on the type of influence we have over our mental lives while bearing on long-standing, philosophical issues such as agency and rationality. In this chapter, I will initially offer a pre-theoretical sketch of what I mean by ‘reasoning’. In doing so, I will show that we need a way to distinguish reasoning from mental processes that are not reasoning, namely inference and associative thinking. After that, I will explain my dissatisfaction with what is for many a very plausible account of reasoning, namely the rule-following view. This view, I claim, is insufficient to clearly set reasoning apart from non-reasoning processes. To amend for that, I point toward a different approach: reasoning should be seen as an activity constituted by an aim. Not any aim will work, though. Thus, I propose the Reason View of reasoning: reasoning is a process of attitude-

⁸⁹ See, for instance, Broome, 2013; Boghossian, 2014; McHugh and Way (2016, ms); Neta, 2013; Pettit, 1993; Valaris, 2014; Wedgewood, 2006.

revision that is necessarily guided by the aim of arriving at a conclusion supported by at least sufficient reasons. This approach is fully consistent with the rule-following model while avoiding its shortcoming.

A final point before starting. There are two types of reasoning: theoretical and practical, and, possibly, the difference between practical and theoretical reasoning lies solely in the conclusion-state: paradigmatically, intention for practical reasoning and belief for theoretical reasoning (Broome, 2013; Harman, 1986:2). My account of reasoning aims at offering a unifying view of reasoning, one that works for both theoretical and practical reasoning. However, to keep things manageable, in this chapter I will mostly focus on theoretical reasoning.

REASONING AND INFERRING

Recently, there has been a lively discussion concerning the nature of reasoning. Before saying what that is, though, we need to carve out a pre-theoretical account of what mental phenomenon we take reasoning to be. Reasoning – many agree – is a kind of attitude revision (Harman, 1986) that usually starts with some (full) attitude-premise and ends with dropping, forming, or withholding another (full) attitude. In this respect, reasoning is an inference.⁹⁰ Through inference one may come to have a new belief-like attitude; alternatively, one may abandon such an attitude; or it may simply be that one withholds an attitude.⁹¹ Also, for simplicity let's assume that inference is only conceptual, so for instance the transitions that go on in the perceptual systems are not

⁹⁰ Sometimes by 'inferring' we simply mean transitioning from one proposition to another (Wright, 2014), but this is not what it is usually meant by inferring in the literature.

⁹¹ The reason why I talk about belief-like attitudes is because I assume that inference and reasoning can involve not only beliefs, but also suppositions and acceptances. I will come back to this below.

inferences (Burge, 2010: 406). It is also widely held that inferring is a causal process, in which some attitude A causes one to form/drop/withhold attitude B. It is also a natural view that the causal process going on in reasoning is a *basing* process of forming or dropping attitudes for motivational reasons. Basing distinguishes inferring from purely associative thinking.

Beyond inferring *simpliciter* adult humans seem to engage in possibly more sophisticated inferential processes that we often call reasoning. To explain what I mean, I will point to some key differences between inference and reasoning by drawing the reader's attention to characteristics of reasoning that are not shared by all inferences.

Personal-level process: It is often said that reasoning is “person-level” (Boghossian, 2014: 2). First, that means that reasoning is a process attributable to the whole person rather than to a sub-mechanism of the mind (e.g. perceptual modules). But often by ‘personal-level’ theorists also mean ‘conscious’. Thus, in defining reasoning, the psychologist Haidt (2001: 818) states that “[t]he reasoner searches for relevant evidence, weighs evidence, coordinates evidence with theories, and reaches a decision [...]. Some of these steps may be performed unconsciously [...], but a key part of the definition of reasoning is that it has steps, at least a few of which are performed consciously.”

In general, reasoning, but not necessarily inference, is a mental process in which (at least some of the) premises and conclusions, the attitudes we operate from and arrive at, are *conscious* (Pettit, 1993: 223-224). (By ‘conscious’ I mean, roughly, that they are available for report). Although the inference carried out in reasoning is usually unconscious, we are aware of the starting and arrival points of that inference. Again, psychologists Mercier and Sperber (2011: 57) explain that “[r]easoning, as commonly

understood, refers to a very special form of inference [...], where not only is a new mental representation (or conclusion) consciously produced, but the previously held representations (or premises) that warrant it are also consciously entertained.” In ‘simple’ inference, as I call non-reasoning inferential processes, we form an attitude or arrive at a conclusion through a causal process, but we are often unaware of the premises and, at times, also the conclusion remains unconscious. Here is an example. At the end of a long dinner party, we leave with the strong impression that the atmosphere at the party was somewhat tense; but we can’t really tell why: the process that got us to form that belief is based on subtle cues we detected only unconsciously. Many are happy to call this inference, but not reasoning (Johnson-Laird 2008: 60–72; Valaris, 2016). Note that these unconscious inferential processes can be *based* on reasons, i.e. considerations in light of which an agent arrives at a certain conclusion, just as much as reasoning.⁹² So reasoning is a subclass of basing relations. More specifically, reasoning is *conscious* basing.

Activity: Reasoning is an activity. By this is often meant that reasoning is something we do and take active part in. In addition, theorists seem to agree that reasoning is “attention hogging and effortful” (Boghossian, 2014: 2; Haidt, 2001: 818). Non-reasoning inferential processes, in contrast, are something that happens to us, or at least something we have little control over (Mercier and Sperber, 2011; Pettit, 1993). Take the case of unconscious basing I mentioned above. If I find myself thinking that the atmosphere at the party was tense, I haven’t done anything to reach that conclusion, nor did I take part in the inferential process. So that inference does not seem to count as an activity.

⁹² See Evans (2013) for examples of unconscious basing.

In the literature, there have been various attempts to explain in what sense our mental lives can be active⁹³. In what follows I will offer an account of reasoning that captures the idea of reasoning being an activity. That said, it is important not to overstate the extent to which reasoning is an effortful, slow process. At times, reasoning is very slow, and captures the full attention of an agent. That's *deliberative* reasoning. Not all reasoning is deliberative, though. Deliberation seems to be a fairly high-level form of reasoning, usually sprung by the explicit questions: "What should I believe?", "Is this right for me to do?" And so on. By contrast, in (non-deliberative) reasoning we don't pose any question, but go from attitude to attitude (or lack thereof) without explicitly attending to the relation between the attitudes we started with and the conclusion we draw from them. In other words, in (non-deliberative) reasoning, we are aware of our premises, and we find ourselves working our way toward a conclusion with usually little effort and self-reflection.

Endorsing attitudes: Another element of distinction between reasoning and simple inference is that the result of reasoning is something we endorse. This point has not been fully explored in the literature, and is bound to strike the reader as more controversial than the previous two points.

Let's look at some plausible marks of endorsement. Reasoning often results in forming a propositional attitude. Attitudes such as beliefs (and intentions) are often defined as dispositions to use propositions as premises in reasoning and inference (when relevant). So endorsing an attitude means first and foremost that we have no *prima facie* objection to its use as premise in future reasoning and inference. In addition, when it

⁹³ In particular, see Gibbons (2009) and Hieronymi (2009). For worries see Setiya (2013).

comes to beliefs, it seems reasonable to say that the beliefs we endorse are the beliefs whose content we are disposed to sincerely assert in speech.

In addition, I believe that endorsing a conclusion X means that we control X and take responsibility for it. To explain this let's again look at belief. There are beliefs that are irresponsive to reasons to the point that the agent feels she has no control over them, and is reluctant to take responsibility for them. Examples of these are paranoid, alienated (Hunter, 2011), dissonant (Borgoni, 2015) or phobic beliefs (Egan and Elga, 2005). We can use Huddleston's (2013) convenient label and call them "naughty beliefs", i.e. evidence-immune beliefs we have, but often openly reject. (I assume that the category of "naughty attitudes", i.e. attitudes that are irresponsive to reasons, can be extended to other attitudes beyond belief⁹⁴). In contrast, there are attitudes we actually have control over, and happily take responsibility for. And note that the idea of endorsed beliefs is hardly new. de Sousa's "assent" (1971), Dennett's "opinion" (1978), Sperber's "reflective belief" (2000, 1996), and Frankish's "superbelief" (2004) seem all to suggest that we support some of our doxastic states with particular force or conviction, and that these are the result of some kind of high level reasoning. These are the beliefs we see as reflecting our rational agency (Moran, 2001; Boyle, 2011). These are the beliefs we endorse.

True, reasoning does not always conclude with the formation of an attitude: in reasoning, we may end up dropping or withholding an attitude as well. But this poses no worry for my view. We can endorse our dropping/withholding of an attitude just as much

⁹⁴ As for desires, current discussion of un-endorsed desires stems from Harry Frankfurt.

as we can endorse an attitude. In these cases – as in the case of attitudes we endorse – we feel we have control over what happens in our minds.⁹⁵

Now, I believe that reasoning is the *paradigmatic* way in which we produce attitudes (or lack thereof) we endorse, not inference. Our inferential processes can give rise to conclusions we don't endorse; indeed, these processes can go wrong in important ways, and we can be perfectly aware of that. And so we can find ourselves believing something we don't subscribe to. Imagine that your therapist convinces you that your conscious belief that you should hire the male job candidate over the female one is inferentially based on your unconscious belief that women are intrinsically less qualified than men. And yet you have difficulties taking control over it: you just can't stop believing that you should hire the male job candidate over the female one. This is a belief you now don't endorse because you know it to be poorly based. Still, the inferential process that grounds it is not disrupted by that. In contrast, the idea I am pushing for here is that there is no such thing as reasoning toward a conclusion we don't endorse: if we inferentially arrive at a conclusion, and don't endorse it because we believe it is falling in some ways, then *that* inferential process does *not* count as reasoning.⁹⁶ It is, of course, possible to reason poorly. But imagine a reasoner who reaches a conclusion through reasoning and becomes aware that that reasoning pattern was badly handled. Such a recognition on her part counts already as a way of reasoning out of that conclusion. If she can't do that, then there are reasons to suspect that she was not reasoning in the first place (as argued by Haidt, 2001 for moral judgments).

⁹⁵ Although commonly used in philosophy, the notion of 'control' is admittedly very obscure and controversial; for now I trust the reader has an intuitive grasp of it.

⁹⁶ As a result, I don't think there is such thing as akratic reasoning. There are akratic attitudes (or lack thereof), but the idea of akratic reasoning seems to me a contradiction in terms.

Before ending this section, let me mention an obvious worry the reader may have at this point. What about hypothetical/ exploratory reasoning or reasoning with *reduction ad absurdum*? That seems to be clear case of reasoning in which we do *not* endorse our conclusion. As a reply, let me point out, first, that these are only limited cases of reasoning. When we think of paradigmatic cases of reasoning, we do not think of hypothetical reasoning. So it should not come as a surprise that our conception of reasoning does not perfectly fit those cases, as they are at the fringes. In addition, note that the end point of hypothetical reasoning is not a belief, but a hypothesis. Similarly, reasoning *per absurdum* means starting with propositions one accepts to be true for the sake of the argument and ending with another proposition that is again taken to be true for the sake of the argument. The end point of our reasoning, i.e. this attitude of accepting or taking to be true, is something we endorse, albeit only within the limited context of our argument or hypothetical reasoning (see chapter 2 for an argument on this). Note that acceptance is not necessarily responsive only to epistemic reasons as belief. Indeed, we accept propositions for the sake of an argument or to show the absurdity/ falsity of some premises. So it is not the case that we don't endorse the conclusion of our, e.g. *reduction ad absurdum*, reasoning. We do, but for pragmatic reasons.⁹⁷

Justification-transmission process: Inference and reasoning plausibly differ in their epistemic roles: whereas inference can transmit only epistemic “entitlement”,

⁹⁷ “After a long philosophical discussion, she reasoned herself into a corner, but could not accept the conclusion that time does not exist”. Isn't this a counterexample to the idea that in reasoning we always reach a conclusion we endorse? Hardly. Note that the sentence above is very ambiguous. It is not even clear that the reasoner there formed some kind of attitude as the last step of her reasoning process. She does not believe, or accept, or withhold judgment concerning the conclusion that time does not exist. However, as stated at the beginning, in the literature of reasoning, it is usually said that reasoning can't just end with some proposition but produces some kind of stance vis-à-vis that entertained proposition. Otherwise it does not count as reasoning, but simply as an *attempt* at reasoning.

reasoning is a process that can transmit epistemic justification (Burge, 2003). Plausibly, this is due to the fact that reasoning requires *some* reflective awareness concerning the justificatory status of the premises vis-à-vis the conclusion. This idea squares nicely with the inferential internalist claim that justification is transferred from premises to conclusion only if the subject is in a position to know or justifiably believe, by reflection alone, that her premises provide her with a reason for the conclusion (e.g. Bonjour, 1985). One way to explain this is to say that to be justified in arriving at a conclusion as a result of reasoning one needs to exclude blame. Being justified – in this sense of justification – is in part a matter of being blameless. Hence, to transmit justification reasoning – being a reflective activity – requires that the agent justifiably take her premises to justify her conclusion. Note that the ‘taking’ here can take various forms. Some forms of internalism require that in reasoning one actually justifiably believes that the premises are supported by the conclusion. Other forms of internalism only require that one is in a position to come to form that belief. Regardless of these differences, however, the important point here is that reasoning seems to be able to transmit a (sophisticated form of) warrant in part in virtue of its being an inferential process that requires *some* reflective self-awareness. An analysis of reasoning should be able to account for this notion of ‘taking’ which, on this view, is key to transmit justification.⁹⁸

⁹⁸ A final word about the reasoning vs. inference distinction I am trying to draw here. A way to capture it is in terms of the system 1 vs. system 2 distinction (Haidt, 2001; Boghossian, 2014: 2). However, I believe that doing so would be of little help. Some psychologists have recently put forward the idea that we “reason” by using a suit of mechanisms that produce responses in a fast, automatic and fairly effortless way. These mechanisms – called System 1 – are often said to use associative or intuitive processes to deliver their output (Kahneman, 2011). In contrast, System 2 processes are much slower, more cognitive demanding thought-processes, in which we consciously weigh the options we have. System 2 is akin to what I called reasoning (although, as mentioned above, I don’t think that reasoning is all-conscious, nor do I think that it always all that slow and cognitive demanding – but I won’t press this point). But unconscious inference hardly fits in this System1 vs. System 2 picture. First, inference can’t be part a system based on

THE RULE-FOLLOWING MODEL OF REASONING

Reasoning is often described as a rule-following process. On this view, thought-transitions count as reasoning only if guided by rules. The two main recent proponents of the rule-following view are Paul Boghossian and John Broome, so I will consider their views in turn. Although the rule-following strategy has enjoyed a lot of support, and many find it extremely compelling⁹⁹, I will criticize it as insufficient to make sense of reasoning. Importantly, I don't aim at arguing that the rule-following approach is mistaken, or that no successful rule-following account of reasoning can be developed. I will show, however, that since the two existing strongest rule-following proposals fall short in important ways, we should look for a new theory of reasoning.

Boghossian. Paul Boghossian has forcefully defended the claim that reasoning is a matter of following a rule. So in reasoning attitude A produces attitude B because we are guided by rules that somehow mandate that transition (what *types* of rules we will say in a moment). His account of reasoning emerges out of what he takes to be an important feature of reasoning, namely the taking condition:

(Taking Condition): reasoning necessarily involves taking the accepted truth of your premises to support your conclusion, and drawing such a conclusion because of that fact.¹⁰⁰

associative thought-processes. So inference does not fit in System 1. At the same time, unconscious inference hardly fits the standard description of System 2, which is mostly conscious and produced with some effort. So for these reasons I find this System 1-2 analogy somewhat unhelpful for my purposes, and I ask the reader to put it on one side.

⁹⁹ See, for instance, Boghossian (2014), Broome (2013), Ichikawa and Jarvis (2013), Wedgwood (2006), Wright (2014).

¹⁰⁰ The idea of the taking condition is taken from Frege's (1979: 3) following statement: "To make a judgment because we are cognisant of other truths as providing a justification for it is known as inferring."

In Boghossian's words that means that "[a] transition from some beliefs to a conclusion counts as [reasoning] only if the thinker takes his conclusion to be supported by the presumed truth of those other beliefs." (2014: 4) Boghossian believes that this attitude of 'taking' needs to play a causal role so that the agent draws her conclusion because of that taking attitude. Hence, the taking condition seems to shed some light on the distinction between the causal process that goes on in reasoning vs. deviant causal chains. It is also helpful to draw a distinction between reasoning and associative thinking. As Boghossian (2014: 4) points out, "A habitual depressive's judging 'I am having so much fun' may routinely cause and explain his judging 'Yet there is so much suffering in the world' [...]." However, the direct causal route between the two attitudes (belief "I am having so much fun", and belief "there is so much suffering in the world") does not make this a case of reasoning. What's missing, asks Boghossian? Well, he believes that the taking condition can at least be one of the missing elements in that picture. Finally, it is worth noticing that the taking condition captures in one scoop some of the key features of reasoning mentioned above. In particular, by imposing a degree of self-reflection and awareness on reasoning, the taking condition seems well positioned to explain why reasoning is something we – as agents – do, and not simply in inferential process that happens sub-personally (Boghossian, 2014). And it can also be helpful to explain how reasoning can transmit justification.

Following Boghossian, many theorists agree that the taking condition is tracking something substantial about reasoning (Neta, 2013; Sperber and Mercier, 2011: 57;

Tucker, 2012; Valaris, 2016)¹⁰¹. I personally believe that the focus of our analysis of reasoning should not be on the taking condition, but on the aim we have when reasoning. That said, for now I will grant that the taking condition really is at the essence of reasoning, and come back to this issue again when I offer my positive view.

Before raising my worry about Boghossian's rule following strategy, I would like to make the following proposal which – I believe – is completely in the spirit of his account. The taking condition, as it is stated above, is too weak to capture something substantial about reasoning. Here is what I mean. The taking condition, whether one believes it is an actual feature of reasoning or not, has been mostly understood in terms of “taking the premises to be reasons”, and in this context by “reason” theorists usually mean normative reasons, namely considerations that count in favor of some response.¹⁰² Notoriously, though, there are different kinds of normative reasons: sufficient, conclusive, pro-tanto and even *prima facie*. *Prima facie* reasons are considerations that *appear* to be reasons, but may actually not be reasons at all; so I doubt the taking condition concerns them. Similarly, taking one's premises to offer some (pro-tanto) reason for a certain conclusion is not enough to capture the spirit of the taking condition. To see this imagine someone inferring p from r while saying that r is only *a* reason to conclude p, but perhaps there are better reasons for not concluding that p. Could the taking condition accept that? I doubt it, and it is excluded by the fact that Boghossian (2014: 15) himself explains the taking condition as taking the premises to *justify* one's conclusion. Thus, it seems that the taking condition, as defined above at least, is much

¹⁰¹ Wright (2014) and McHugh and Way (forthcoming) have argued against the taking condition. Hlobil (2014) has raised worries about the taking condition.

¹⁰² Broome is an exception to this, as he holds that no normative taking is needed to explain reasoning (see fn. below).

too weak to explain what is going on in reasoning. If it is true that we take the premises to be reasons for our conclusion, we need to take those premises to form at least *sufficient* reasons to reach that conclusion. And, even more strongly, in deductive reasoning, we take our premises to give us conclusive reasons to reach our conclusion (Boghossian, 2014: 5). So I think that we should redefine the taking condition as follows:

(Taking Condition) reasoning necessarily involves taking your premises to (at least) sufficiently support your conclusion, and drawing such a conclusion because of that fact.

Does Boghossian (2014) manage to make sense of the Taking Condition? Boghossian (2014) himself takes the “taking” in the taking condition to be problematic. It seems intuitive to say that the taking condition requires a doxastic reading. Accordingly, when reasoning we have a second-order belief that our conclusion is supported by the premises. As Boghossian puts it, “[t]he first thought that is likely to occur to one is that the Taking Condition requires that a thinker have a meta-belief about the relation between his premise judgments and his conclusion, a belief to the effect that his premise judgments supply him with a justification for believing his conclusion.” (2014: 6) However, such an explanation leads us into troubles. It is a natural thing to say that this meta-belief should play the role of a premise in bringing about the conclusion of one’s reasoning. This, many agree, raises the worry that now we would need a further premise that says that all these premises taken together support the conclusion. This clearly pushes us into a regress.

To avoid this, Boghossian (2014) has favored an alternative approach, namely that the taking condition can be captured in terms of rule-following. However, any rule-

following proposal struggles with the problem of explaining what rule-following is. More specifically, if applying a rule requires some reasoning, then in order to follow a rule in reasoning, one has to go through a bit of reasoning; but to do so, we need to follow a rule and thus a regress ensues (Boghossian 2012).¹⁰³ To avoid the regress, Boghossian settles for the following position: we should take the notion of “following a rule” to be a primitive notion.

Importantly for our purposes, Boghossian thinks that this proposal can explain the taking condition: if someone follows the rule to do F in circumstance X, then she takes X to be a reason to do F in the sense that the obtaining of X is treated as a reason to do F. Applied to epistemic reasoning, this means that if someone follows a rule in reasoning such as, roughly, “if (believed) p then (believed) q”, she takes the truth of p to be a reason to believe that q. Moreover, such a person believes that q because of this. Hence, understanding reasoning as involving following a rule seems to be able to accommodate the taking condition, at least according to Boghossian.

Boghossian’s view is appealing as it promises to explain many of the features of reasoning we have been hoping to make sense of. Unfortunately, I believe Boghossian’s rule-following view fails where it hopes to succeed the most. That is, his view proves too weak to account for the taking condition, which – as the reader may recall – was his main motivation for offering the rule-following approach in the first place.

According to Boghossian (2008), in reasoning we follow rules, such as:

[MP]: If you are rationally permitted to believe both that p and that ‘If p, then q’, then, you are *prima facie* rationally permitted to believe that q.

¹⁰³ Another solution is that following a rule is a dispositional state that does not itself require any reasoning. Alternatively: following a rule is a feature of a system of the mind we have no access to. Boghossian (2014) finds both these proposals problematic.

[I] For appropriate Fs and Gs, if you have observed n (for some sufficiently large n) Fs and they have all been Gs, then you are prima facie rationally permitted to believe that all Fs are Gs.

Boghossian's claim that rule-following is a primitive notion is bound to strike many as unsatisfactory. Still, here we don't need to concern ourselves with the question of how agents follow a rule like [MP] or [I], and how, for instance, they bring about the transition from believing p and believing $p \rightarrow q$ to believing q . Instead, assuming [MP] and [I] are indeed rules we follow in reasoning (however we do it), we need to ask whether following [MP]/[I] can guarantee that we satisfy the taking condition.

First, as John Broome has pointed out in his criticism of Boghossian's proposal, I can follow a rule without thinking that it is a good rule, or a rule worth following. So I may be growing increasingly skeptical of [MP], and still decide to apply such a rule to some propositions p and q . Understandably, I would *not* take my premises as reasons for my conclusion, in any plausible reading of "take". Second, as mentioned above, the taking condition does not simply boil down to the fact that we take our conclusion to be supported by or follow from the premises, or that our premises function as reasons for our conclusion. If the taking condition tracks something substantial about reasoning, it needs to be much stronger than that. We need to take our premises as *sufficient* support for our conclusion. This is because that 'taking' needs to somehow figure in the explanation of why we form the conclusion we form; and we don't reason toward a conclusion if we merely think that all we have is a pro-tanto reason for it. This casts doubts that following [I] in reasoning can explain the taking condition. Suppose that, while following [I], you also believe that you have independent reason to believe that it is in fact impossible that

all Fs are Gs. Now, that consideration undermines the support that your observations can give to the conclusion, and you know it. So you can follow rule [I] *without* thinking that your premises are *sufficient* reasons for our conclusion. Hence, following a rule such as [I] does not capture the taking condition.

To address this, one could insert a “*ceteris paribus*” clause in [I] to make sure that following a rule like [I] gets us the taking condition. *Ceteris paribus* clauses are problematic as they threaten to conflate the distinction between correct and incorrect performance (Earman and Roberts, 1999), and should be avoided, if possible. More importantly, this proposal threatens to make [I] unhelpful for Boghossian’s purposes because what follows under the ‘*ceteris paribus*’ clause is any condition that allows the agent, who follows [I], to take her premises as sufficient reasons for her conclusion. But this turns out to be the taking condition. i.e. the condition the rule-following story was meant to explain! With the ‘*ceteris paribus*’ clause, claiming that in reasoning we follow [I] becomes explanatory idle. This problem generalizes. If Boghossian is to say that reasoning means following (what we take to be) a good rule – whatever that is – he needs to tell us what a good rule is. He seems to be forced to say that that is a rule that makes the premises of our reasoning be sufficient reasons for its conclusion. But that is viciously circular: the rule-following proposal was fashioned to explain the taking condition not vice-versa. So his proposal is insufficient for that.

Broome. Perhaps there is hope for the rule-following strategy: John Broome has recently put forward a proposal which is somewhat different than Boghossian’s. As Broome (2013: 232) puts it, “[i]t is an essential feature of reasoning that the operation is rule-governed. In reasoning you follow – are guided by – a rule.” On his view, following

a rule is, first, a disposition to conform to it. Second, he rightly points out that, “[i]n order to follow a rule, you do not need to know explicitly what the rule is. You follow the rule, but you may do so in the way in which you often follow rules of grammar. You may compose grammatical sentences without knowing explicitly what grammatical rules you follow in doing so.” (Broome, 2013: 232)

However, as in the grammar case, when you reason from premises to a conclusion you *see* your conclusion as right.¹⁰⁴ Broome explains ‘seeming right’ as a counterfactual disposition. Processes that are guided by a rule are open to revision and checking. He says, “[h]aving the attitude of seeming right involves a disposition to stop having this attitude in particular circumstances, specifically if you were to check the act and it were no longer to seem right.” (Broome, 2012: 22) On this view, seeming right is the mark of rule-following.

On Broome’s account, the fact that reasoning is rule-following process has two important implications. First, Broome maintains that a rule following process is necessarily an active process. So by virtue of being a rule-following process, reasoning is thereby an activity. So the rule-following account can explain one key element of reasoning, namely its being an activity. Second, the fact that reasoning is rule-following avoids the problem of deviant causal chains which, Broome seems to believe, is a problem an account of reasoning needs to address. Causal relations are subject to deviant causal chain, as I mentioned above. Only those causal relations among attitudes that are guided by a rule, Broome says, can count as reasoning.

¹⁰⁴ Tucker (2012) also adopts this idea of intellectual seeming.

Broome's solution may look more attractive than Boghossian's because it promises to explain key features of reasoning without getting entangled with the difficulties of the taking condition. Yet, I believe that Broome is throwing the baby out with the bathwater.¹⁰⁵ Boghossian's taking condition was, I believe, addressing something important about reasoning that Broome's account leaves unexplained, namely the fact that the result of reasoning are attitudes (or lack thereof) we endorse. For Broome (2013: 237) "[s]eeming right does not entail an implicit belief whose content is normative [...]. The 'right' in 'seeming right' is only relative to the rule; it is not normative. You can follow a rule without believing it is permissible to do so." Assume I follow a rule R in my reasoning and when I reach conclusion C, I take C to be correct or right vis-à-vis R. Say, however, I take R to be a wrong or stupid rule to follow, and thus I am also inclined to believe that C is not a good attitude to have. On Broome's account then I can reason toward an attitude I don't endorse. And this is problematic, I believe.

Above I pointed out that what makes reasoning different from other inferential processes is that in reasoning we end up with a conclusion we endorse. (Again, 'endorsing' means, roughly, that we are willing to use it in further reasoning and take responsibility for it.) I believe endorsement is a key feature of reasoning, but let's put it in question for the sake of the argument. If not through reasoning, how do we reach attitudes we endorse? One possibility is to say that having an attitude *automatically* means endorsing it. But besides being quite a commitment to take, it contradicts Broome's point that we can reach conclusions based on reasoning whose rules we don't endorse at all. So we can have an attitude that is the result of us following a rule we

¹⁰⁵ See Boghossian (2016), McHugh and Way (2015), Pettit (2016), and Valaris (2016) for additional problems with Broome's proposal.

don't subscribe to. So Broome is committed to saying that having an attitude is not the same as endorsing it. So how do we explain endorsement? A seemingly plausible solution would be to maintain that endorsing an attitude is having a second-order belief (or seeming) that such an attitude is good, or right or justified. But a regress is looming at large. If to endorse an attitude, I need some second-order attitude, then how do I endorse that further attitude? How can that second-order attitude function as an endorsement if it is not itself endorsed? So it seems natural to say that to endorse an attitude we don't need any second-order attitude. What else? Perhaps endorsing an attitude means having no disposition to change it. The problem with this is that, on Broome's account, that is true for *all* cases of rule-following. As I pointed out above, all rule-following requires that you see your conclusion as right *with respect to the rule*, which for him means that you are not disposed to change that conclusion, otherwise - on his view - that does not count as rule-following to begin with.

As a result, I believe it is natural to suppose that we endorse an attitude when we reached it through reasoning. I am not saying this is the only way, but it seems to me to be the *paradigmatic* way in which we do that. As we saw, Broome's account of reasoning is compatible with the idea that we can take the rule we follow as incorrect. But if reasoning is what brings us to have attitudes we endorse, Broome is forced to say that, when we reach an attitude through reasoning, something else is needed, something beyond rule-following that makes the reached conclusion into an attitude I endorse. And this seems to be the normative belief he is refusing to include in reasoning.¹⁰⁶ Hence, I

¹⁰⁶ In his book on reasoning, Broome introduces a further necessary condition for reasoning which, however, applies only to theoretical reasoning. He believes that to reason from belief A to belief B one needs to have an implicit linking belief. Apparently, this point seems to go in the same direction of Boghossian's taking condition. Note, however, that this is not a normative belief: it is an agent's belief that,

believe that, as stated, Broome's solution either becomes the same as Boghossian's or avoids the problems faced by Boghossian, but at the expense of failing to make sense of what I take to be an important feature of reasoning.¹⁰⁷

Against rule-following. Finally, let me mention some worries that threaten the rule-following strategy *as a whole*. First, I worry about the *scope* of the rule-following strategy. Boghossian and Broome both seem to believe that the fact that reasoning is rule-following at least partially solves the problem of deviant causal chains and basing. In addition, they seem to believe that it can also explain why reasoning is a first-person activity. But they can't have it both ways! The problem emerges because the deviant causal chain issue is faced also by inferential causal chains in general, and as such seems to require a similar explanation across the board. If the rule-following strategy is meant to tackle that issue, then it means that inference – not just reasoning – is a rule-following process. If so, then appealing to rule-following can't explain why reasoning is a first-person activity, and so we need an account for that too. Thus, rule-following is at best incomplete as an account of reasoning: it may capture why reasoning is an inference, but

given some premises, a conclusion will follow because of the rule she is following, independently of whether the agent takes that rule to be correct or not. So this does not help him with the problem mentioned above.

¹⁰⁷ A supporter of Broome's rule-following strategy could try to push back and say that all we need is to introduce a further clause: reasoning is following a rule one believes to be good, a rule that that other things been equal delivers a conclusion supported by the premises. This solution, however, is at risk of pushing us back into the regress problem. This belief is meant to explain why in reasoning we reach conclusions we endorse; as a result, it seems that this belief needs to play a direct role in producing that conclusion. But this raises the regress-worry mentioned above. Alternatively, one could try to argue that such a supporting belief needs to somehow 'accompany' the reasoning process without being a causal force into it. This proposal needs to be spelled out more, though. In what follows, I will offer an account of reasoning that goes into a similar direction, in the sense that reasoning is an inferential processes guided by a specific aim. So those who don't want to give up the rule-following approach could keep it within the framework of the view I propose below.

not why it is an activity. Alternatively, if rule-following serves to explain the active part of reasoning, then it can't be the solution for deviant-causal chain issues.

Second, the rule-following strategy is committed to anti-particularism. Moral particularism is the claim that moral reasoning does not consist in the application of moral principles to specific cases. The rule-following view of reasoning necessarily opposes this view. There is also a version of particularism for the theoretical realm which can be easily dismissed. Although it seems that we have a set of deductive rules at our disposal, this is notoriously a problem for induction. It has proven very hard to come up with a universal logic of induction (Norton, 2010), and it is unclear that, when reasoning inductively, we really apply any general rule. So I worry that the rule-following picture would work only for deductive reasoning. In contrast, the view I propose below is fully compatible with particularism of all stripes.

Finally, even if the rule-following story can address all these worries, it does not capture the point of reasoning. To use Dummett's famous example, the point of playing chess is to defeat your opponent. We could define chess by listing its rules and saying that playing chess is tantamount to following those rules. That may be right, but misses the whole point. Similarly, if reasoning is indeed rule-following, we need an account of reasoning that make sense of that while explaining what the point of reasoning is. Clearly, the point of reasoning is not to follow rules *per se*, as that would be pointless. *If* it is true that we follow rules in reasoning, we do it with some other goal in mind, as I will now argue.

REASONING WITH AN AIM: CORRECTNESS

Reasoning is an inferential process with an aim. The aim of reasoning is an aim *of the agent*, a goal the agent necessarily has when reasoning. An aim is a desire to obtain a certain result (Sosa, 2010). The aim one has is mostly shown in the *dispositions* one has. In this sense, having an aim or goal means being motivated by certain considerations and not others, and being disposed to prevent certain occurrences to happen and change things around if the aim is not satisfied. Having an aim X does not mean necessarily *intending* to do X; it is enough to have a certain disposition to do something if it turns out I am failing to achieve X (Alvarez 2010). Also, the aim of reasoning does not have to be necessarily consciously entertained by the reasoner; however, a normally competent reasoner will not explicitly (and sincerely) deny that this is her aim when reasoning.

Also, note an aim-directed activity may involve a lot of processes I have no direct control over. Take the activity of growing roses in my garden. There is surely something I do first, i.e. planting them and watering them. But other than that, there is not much I am directly involved into, as the plants grow by themselves. However, my *aim* of growing them makes it the case that if something goes wrong (e.g. too little rain) there will be action on my part (e.g. more watering). Notably, we say that *I am growing* roses in my garden even though I am only attending to their growth. Similarly, when reasoning we are guided by a specific aim. And yet reasoning is also an automatic process, as I can't directly control the causal (inferential) relations among attitudes. However, what makes reasoning *reasoning* is the fact that we are supervising those causal relations: if those processes fail to deliver a specific goal, there are things we are disposed to do; we are disposed to change the outcome by focusing on new thoughts, by looking for different

inferential connections, and so on. If things go well, instead, we let it take its course and don't intervene.

But what is the aim of reasoning? A very natural proposal is that reasoning aims at – is guided by the aim of – forming correct attitudes and dropping incorrect ones. 'Correctness' here is a normative property of attitudes such that, roughly, they fit the world right given the type of attitudes they are. This normative property can be rendered in terms of what one ought to believe, fear, desire or what is good or fitting to believe, fear or desire.¹⁰⁸

The idea that reasoning has the aim of correctness is not new. Shah (2003) makes the point that reasoning or deliberation aims at forming an attitude that is correct with respect to the constitutive correctness norm of that attitude. The paradigmatic example of a correctness norm is belief: belief is correct if and only if true. Perhaps this analysis can be extended to other attitudes. Intentions are correct if and only if good, or permissible. Fear is correct if and only if its object is dangerous. And so on. Along similar lines, Moran (2001) argues that – when taking a deliberative stance toward our attitudes – we are really asking a question and inquiring about the *world*. Similarly to Shah, Moran's view seems to be that our goal in reasoning is guided by the correctness norm that is constitutive of the attitude we are forming. More recently, McHugh and Way (2016a) have offered a new version of the correctness view by developing an account of good reasoning. They see correctness as the same as fittingness. As they put it, "[t]he core idea is that attitudes are associated with standards for their objects. An attitude is fitting when its object meets its standard." (2016: 13) Again, the paradigmatic example is that a belief

¹⁰⁸ A synonymous of correctness is 'fittingness'. The concept of fittingness is often employed to offer a fitting-attitudes account of value: O is good when it is fitting to value O. A popular recent suggestion is that fittingness should be understood in terms of sufficient reasons. I will leave this issue aside here.

is fitting if and only if true; an intention is fitting iff permissible. On their view, reasoning aims at issuing fitting attitudes.

The Correctness view is highly intuitive: it does seem that when we reason we want to get at the world, and we want to get things right. Yet, I have some worries about the correctness strategy in general. One worry is that the correctness proposal does not clarify *why* reasoning would always aim at forming attitudes that are *correct*. It is in fact possible to imagine someone arguing that there is more to attitudes than correctness. Indeed, one may think that attitudes can be valuable beyond correctness, namely they may harbor moral and practical values. This of course does not show that the aim of reasoning is not correctness, but it leaves open why we would *always* care about correctness when there may be other more important values to pursue. On the correctness view then, it seems that in reasoning we may end up with conclusions that are correct but that we still don't see as rational, as we may think that are reasons for having or not having an attitude independent of correctness. This squares poorly with the idea that in reasoning we form conclusions we endorse and which we see as reflection of our rationality.

A somewhat related point is that, at least for some philosophers, it is possible that when forming attitudes such as belief we may be *partially* influenced by pragmatic goals (see chapter 5 for why). Unfortunately, the goal of correctness either excludes that possibility (in reasoning we aim only at correctness!) or, if it allows for other goals to enter the reasoning process, then it can't explain why the goal of correctness is seen as the most important one, or the strongest one. Neither of these alternatives seems

particularly appealing, and I am not sure the correctness view has a clear way out of this dilemma.

Furthermore, I worry that the correctness-aim may not be *necessary* for reasoning since we use reasoning also to form attitudes that have no (clear) correctness conditions. I am thinking of suppositions, acceptances in context and even bits of imagination. There is no clear sense in which an acceptance is correct or incorrect (given the sense of correctness we are dealing with here). There is no sense in which imaginations are correct or incorrect. So we would need an explanation for what happens in those cases. I am not saying it cannot be done, but it requires some further investigation. More worryingly, this extends to suspension of judgment as well, as there are no correctness conditions for that one either (Wedgwood, 2002: 272). And suspension of judgment seems to be a genuine attitude just as much as belief and disbelief (Friedman, 2013).

Relatedly, some subscribe to the idea that the standard of correctness for belief is shared by other attitudes, such as guesses and conjectures (McHugh, 2011: 385). This raises the following dilemma. On the one hand, one may reasonably say that that in reasoning we do not aim at forming true conjectures. If we did, we would do our best to ensure their correctness. But we don't: we make conjecture for all sorts of reasons with little evidence to back up their truth. So the aim of reasoning does not seem to work for those attitudes either. On the other hand, if we insist that in reasoning we do aim at forming true conjectures, then we need to explain how the aim of truth differs in belief vs. conjecture. That makes the correctness view explanatory insufficient. Again a dilemma the correctness view needs to address.

A related worry is that this fundamental goal leaves issues of rationality totally unexplained, and is therefore insufficient to explain what is going on in reasoning. When applied to the correctness of belief, truth would be the goal that should guide us in reasoning and deliberation. Unfortunately, it fails to “determine how one is to balance the value of having a correct belief about *p* against the disvalue of having an incorrect belief about *p*; so it cannot determine when it is rational to believe *p* and when it is rational to suspend judgment about *p*” (Wedgwood, 2002: 274). Hence, if truth were our fundamental goal when reasoning theoretically, that would provide little guidance and we would be at loss about what to do most of the time (Engel, 2013).¹⁰⁹

So, although quite intuitive, the correctness view leaves us with a few issues to solve. I will come back to the correctness view in the next chapter. But now it is time to present a different view of the goal we have in reasoning.

THE REASON VIEW OF REASONING

As an alternative to the correctness view, here is the view I will now defend: **reasoning is an inferential process in which we have the aim of arriving at a conclusion supported by (at least) sufficient reasons.**^{110 111}

¹⁰⁹ A way to block this is to defend the view that a correct belief has to be knowledge.

¹¹⁰ In a similar light, Boghossian (2014:5) points out that “no causal process counts as [reasoning], unless it consists in an attempt to arrive at a belief by figuring out what, in some suitably broad sense, is supported by other things one believes. In the relevant sense, reasoning is [...] something that we do with an aim—that of figuring out what follows or is supported by other things one believes.” An approach sympathetic to the Reason View can also be found in Dancy (2004). McHugh and Way (2016) also discuss the Reason view. But to my knowledge this version of the Reason View has not been proposed by anyone before.

¹¹¹ Talking about sufficient reasons may be problematic when referred to suspension of judgment and dropping of attitudes. I think these worries can be addressed by referring to the idea that when reasoning we want our conclusion to *correctly respond* to the reasons we have. If intended broadly enough, that means we can respond to the lack of sufficient reasons for an attitude by dropping it. So, for instance, if I lack

The aim of reaching a conclusion supported by at least sufficient reasons is constitutive of reasoning in the sense that it is what makes reasoning what it is: reasoning *is* having an inferential thought processes guided by this aim.¹¹² The conclusion of reasoning consists in dropping, withholding, reaffirming or forming an attitude. Doing one of those things in a way that is supported by sufficient reasons and correctly responds to (the balance of) all the reasons one has, is the constitutive aim of reasoning. Insofar as the reader accepts the idea that correctly responding to reasons is tantamount to being rational, my claim here is that reasoning aims at rationality, or at least at *this* sense of rationality.¹¹³

The view I am offering here fits nicely with the claim that the notion of ‘reason’ is fundamental (see Parfit 2011; Scanlon 1998), and not dependent on our goals.¹¹⁴ That said, however, I will not take an explicit stand on this issue, and assume that the notion of reason at play here is compatible with different approaches.

Normative reasons are things that count in favor of (or against)¹¹⁵ a certain response. By “response”, I mean the holding, withholding or dropping of an attitude (rather than propositions). It is also a natural position to have that reasons support responses by contributing a kind of epistemic *weight* to them. And ‘sufficient normative

sufficient reasons to hold a belief that p, while lacking sufficient reasons to disbelief that p, I respond correctly to my (lacking of) sufficient reason only if I suspend judgment on p.

¹¹² McHugh and Way (ms) and Whiting (2014) also offer views of reasoning as an aim-directed activity. The Reason View as presented here is deeply influenced by their views.

¹¹³ We are rational, Joseph Raz has suggested, insofar as we exercise the capacities that enable us, in general, to respond correctly to reasons. So when reasoning, I claim, we aim at fulfilling that capacity, at being rational in this sense. However, even if the reader is not sympathetic to the idea that rationality is responding correctly to reasons, it is important to clarify that my argument here does not depend on it in any way.

¹¹⁴ This does not exclude that reasons may be naturalistically grounded, though. I will remain agnostic on this and other issues of the ontology of normative reasons here.

¹¹⁵ I assume that reasons against A are reasons in favour of not-A.

reasons' are reasons that, on balance, sufficiently weigh in favor of a response. Roughly, an agent S has sufficient pragmatic reason to ϕ only when S has a reason to ϕ , and no stronger reason not to ϕ . In contrast¹¹⁶, having sufficient evidence to form a belief that p involves arguably more than having stronger evidence for p than for not-p; it requires some high degree of subjective probability that p is true.

What kind of reasons my conclusion should be responsive to when reasoning? Any reason there is, or the reasons *I* have? This point is important when we think that there may be reasons for doing, believing, intending something we may be totally unaware of and that are not my reasons. Views differ on this issue, but fortunately my proposal is compatible with all of them. To keep things manageable let me focus on the two main approaches on this topic.

Many distinguish the following two things: the fact that there are reasons to do or believe something and the fact that I have those reasons, that those reasons are my reasons and are available to me. Those who endorse this view tend to maintain that reasons are true propositions or facts, and that to have those reasons one needs to stand in some relations to those facts or true propositions (Lord, 2010). If we endorse this view, then the aim of reasoning is to get to a conclusion that correctly responds to the reasons we have. These, however, are potentially only a subset of the reasons that there are for an agent to reach that conclusion. This may seem unsatisfactory, but it is not. The reason why it may *seem* unsatisfactory is that it seems that when we reason we want a conclusion that is supported and justified by the balance of *all* the reasons. Our prospective may be limited and the reasons we have, even when sufficient, may

¹¹⁶ Some authors hold that one can have an *insufficient* reason for a belief even if no opposing reason is present (Raz, 2009).

constitute only a small subsets of the reasons we could potential get access to were we in a better epistemic position. In contrast, I believe that asking that this be our aim is asking too much. There are reasons we may never get to make ours and that are beyond our reach. And there are cases in which we have to make a choice with the limited information we have. In those cases, it is unclear that, when we reason, we can aim at anything else than being supported by the reasons we have. Why? Well, only *available* reasons can guide us in deliberation and reasoning, whereas reasons we have no possible access to can't guide us at all. So it would be indeed odd to insist that, when reasoning, we aim at arriving at a conclusion that is supported (also) by reasons we don't have within reach, and that can't function as our base when reasoning.

On a different view of reasons, theorists are more inclined to draw a distinction between the objective reason relation and the subjective reason relation: an objective reason is a reason there is *for* someone to do or believe something; a subjective reason is a reason someone *has*. Objective reasons are factive. Subjective reasons are mental states and may not be factive (Schroeder, 2008). A supporter of this view can adopt my approach to reasoning and agree that reasoning aims at arriving at a conclusion supported at least by sufficient *subjective* reasons. Again, this may appear quite unsatisfactory. When we reason, one may think, we don't aim at arriving at a conclusion based on *false* premises: we want to get at a conclusion based on intuitively good, i.e. true, premises. Although this is true, it should be noted that the agent can't differentiate between objective and subjective reasons. From the inside, subjective reasons appear objective; so

for an agent aiming at having a conclusion supported by objective reasons is the same as aiming at a conclusion that is supported by subjective reasons.¹¹⁷

In conclusion, I believe that the plausibility of my view of reasoning is preserved on various accounts of normative reason. It is now time to argue for the Reason View by showing how it can deal with various issues concerning reasoning.

DEFENDING THE REASON VIEW

The idea that reasoning has an aim constitutive of it – whatever that aims is – is not far-fetched. At least intuitively, it seems that when reasoning we have a direction, we are trying to get to a conclusion. So saying that reasoning is guided by an aim does not sound implausible to begin with. Similarly, it is plausible that agents aim at reaching conclusions supported by sufficient reasons: agents want to do, believe, and intend what is supported by reasons. Why? Well, first, we want attitudes we can defend and argue for if questioned. Reasoning is thus the way in which we form attitudes for which we can give reasons, and don't need to retract if challenged. This may be key in securing a status as reliable epistemic and communicative agents (Brandom, 1994).

In addition, it is quite plausible that in reasoning we want to reach attitudes that are intelligible to ourselves and to others. As our attitudes are part of our identity and self, we want to have attitudes that are not only correct but also rational and that form a somewhat coherent, intelligible psychological unity. Again, reasoning seems to be the most obvious way to reach that goal.

¹¹⁷ Also, some theorists have defended the claim that reasons are *only* either factive psychological attitudes (e.g. true beliefs, knowledge, evidence) of the agent or facts that the agent has access to. On this view, no reason is not-had by the agent.

But beyond being quite intuitive, the Reason View can make sense of important features of reasoning.

Personal Level. Reasoning is *personal level* and reasoning, but not necessarily inference, is a mental process in which premises and conclusions, at least some of the attitudes we operate from and arrive at, are *conscious*. The Reason View squares nicely with this idea: an inferential process can be guided by the aim of getting conclusion supported by reasons only if at least part of the process (the premises, at least) is consciously available to the agent. Otherwise, how would she be able to monitor that the process is leading to a conclusion supported by reasons, if she has no clue about what is going on, or no clue about what motivated that conclusion?

Activity. Reasoning, contrary to simple inference, is an *activity*. It's aim-directedness makes it analogous to other personal-level *activities*. Take an activity such as gardening. Arguably, there is no gardening if you stop aiming at making your plants grow. You can decide to stop watering or trimming them, and let them grow as they want. However, the moment you do that, you are not gardening your plants anymore. Hence, the activity of gardening itself has a constitutive aim. Similarly, reasoning has a constitutive aim, a goal that it cannot be undermined by any other aim one may have, if one still counts as reasoning. Importantly, this aim determines the kind of control we exercise in reasoning (Hieronymi, 2009). Reasoning, like gardening, involves a lot of automatic processing that is beyond our direct influence. Still, we do exercise partial control on the inferences that constitute reasoning by overseeing the causal processes, and being active in changing things around if something goes awry, and threatens the aim of reasoning. Thus, reasoning's aim-directedness puts reasoning very close to other

activities that share that feature, and makes sense of the idea that reasoning is itself an activity.

In addition, there is a tradition in philosophy that distinguishes between a passive and an active role we have with respect of our attitudes: activity is, paradigmatically, responding correctly to reasons. On this view, reasoning is an active process because, by being guided by the aim of arriving at a conclusion supported by reasons, reasoning is how we exercise a reflective control over our mental attitudes. Thus, reflective control over attitudes is the result or the expression of our (rational) agency (Boyle, 2009, 2011; Moran, 2001). By being expression of our agency, reasoning is thus something we do, not simply something that happens to us.

Endorsing attitudes: The Reason View makes sense of the idea that in reasoning we produce attitudes or conclusions we endorse. Endorsing an attitude does not necessarily mean having a second-order belief that that attitude is justified or rational: it is enough for an attitude to be endorsed that it is the result of reasoning. On my view, the reason why reasoning forms attitudes we endorse is because reasoning *is* the process *aimed* at forming attitudes supported by reasons - so arriving at the conclusion of reasoning means accepting something *qua* supported by reasons. We do not need to believe we have reasons for it: the fact that we reasoned to it is already expression of our endorsement. Thus, my proposal explains why reasoning to an attitude is tantamount to endorsing it *without* any need for meta-cognitive beliefs to take place in the reasoning process (Pettit, 1993).

Note that nothing I have said requires that reasoning is in fact a rational process, or that the attitudes we endorse through reasoning are in fact rational. There is mounting

evidence that reasoning may actually take us away from rationality and justification. In addition, recent studies in psychology seem to indicate that some of our attitudes are formed in a way that is mostly not transparent to us. These attitudes pop up in our minds and we have little idea how they got there. Reasoning, on this view, is there to rationalize the attitudes we have. For instance, in his landmark paper, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," Jonathan Haidt considers the mechanism by which we arrive at moral judgments, asking whether our moral judgments are made intuitively or as a result of reasoning. He claims that moral judgments are made by reliance upon unconscious intuition and that reasoning is merely *ex post facto*, applied only after the fact to justify one's intuition and subsequent moral judgment. These results are extended also to other types of judgments and indicate that our attempts at justification of attitudes and choices sometimes collapse into forms of confabulation (Nisbett & Wilson, 1977). This confabulatory process forms new attitudes that are held basically because of their ability to rationalize attitudes we formed in different ways (Mercier & Sperber, 2011).

What I have said about reasoning is compatible and in line with these recent findings. First, my proposal is not about what reasoning achieves, but about the status of the attitudes produced by reasoning. If – as I believe – the status of those attitudes is the one of endorsement, then it does not matter whether that endorsement is in fact well-placed. Even if it turns out we are not all that good at reasoning correctly, the important issue here is what we are *trying* to do when we reason, not whether we achieve it. In addition, my view is compatible with the possibility that in reasoning the causal process of attitude-formation may happen backwards: from the conclusion to the premises that

are meant to support it. On my view, such a process is supervised by the aim of getting attitudes supported by reasons, and that squares nicely with the studies in human confabulation. Even if, as a result of reasoning, we end up confabulating about the reasons we have for having a particular attitude, it still must be the case that we somehow *aim* at forming a conclusion supported by normative reasons, otherwise the confabulation would not take place at all. So this actually supports my view, and the idea that reasoning is guided by the aim of having attitudes backed up by reasons.¹¹⁸

The Taking Condition: Boghossian and others defend the taking condition, namely that reasoning necessarily involves taking your premises to (at least sufficiently) support your conclusion and drawing such a conclusion because of that fact. As I mentioned, this boils down to the claim that in reasoning we take our premises as normative reasons for the conclusion. The Reason View intuitively explains the taking condition, and thus could be congenial to those who accept that condition of reasoning. When reasoning, we have the aim of reaching a conclusion supported by at least sufficient reasons. Since the premises we adopt are our means to reach a well-supported conclusion, it is no surprise that we aim to use premises that are reasons. That does not mean we have any explicit attitude toward our premises in reasoning; more specifically, we don't necessarily believe that our premises support our conclusion. But in reasoning we do aim at a conclusion supported by reasons, and the suggestion here is the 'taking' of

¹¹⁸ It possible – as suggested by Mercier & Sperber (2011) – that the biological function of reasoning is to form arguments. On their view, the biological function of reasoning is to argue and offer reasons for attitudes we already have. This explains, on their view, why we are so prone to biases such as the confirmation bias. My analysis, however, is about what the agent is actually trying to achieve in reasoning, and not about the biological function of reasoning. As stated before, the aim of reasoning is an aim agents have. Although an agent's aim in activity X does not have to be necessarily consciously entertained, a normally competent agent will not explicitly (and sincerely) deny that this is her aim when engaging in X. Relatedly, agent's aim motivate us to do certain things, and see certain conditions as reasons for doing those things (Alvarez 2010: ch. 3). These conditions do not apply to biological functions.

the taking condition is captured in terms of an aim rather than a belief. Having an aim or goal means being motivated to do certain things and not others, prevent certain occurrences from happening, change things around if the aim is not satisfied, and so on. Accordingly, if in reasoning we see that our premises do not sufficiently support our conclusion, there are things we will be disposed to do to change that. In other words, suspecting that our premises fail to support our conclusion will *motivate* us to change it. And we will also be disposed to see that as a reason to change it. This is what aims do. Unless we suspect that something is wrong, no belief is needed. The aim of reasoning explains why, in that case, the fact that we sail along and do not intervene is testament of our *taking* our premises to be reasons for our conclusion.

This is true modulo some exceptions that are usually acknowledged in the literature on the taking condition (Boghossian, 2015:43; Valaris, 2014, fn. 3). That is, in reasoning we don't *always* take our premises as reasons for our conclusion. A clear example of that is *reductio ad absurdum* in which the premises are propositions we assume to be true for the sake of the argument we are making and don't take to be reasons. This extends to all the case of reasoning in which we use, as our starting points, propositions we accept to be true in some contexts but not others, things we suppose, or even imagine to be true (Velleman, 2000). In all these cases, we hardly see our premises as reasons. This is compatible with the Reason View. It is not that in hypothetical reasoning and the like we don't aim at arriving at a conclusion supported by reasons. Of course, we do. It is just that in those cases the reasons that support our conclusion are not going to show up as our premises.

A final point. By saying that reasoning requires some sort of metacognition, I am implicitly stating that animals and babies don't reason. They have inferential processes all right, but they do not count as reasoning. I take this to be hardly controversial. That said, we should also notice that the metacognitive capacities that are required for reasoning may actually be very minimal. That is, taking a premise to support a conclusion may only require a fairly basic understanding of what counts as a reason. For such a basic understanding all that is needed is sensitivity to the normative standard of attitudes. Interestingly, even young children may have that sensitivity. If a child is able to assess the validity of a claim or belief based on its causes, we may be in a position to grant her an implicit understanding of reasons. More concretely, if a child reliably spots normatively appropriate vs. inappropriate beliefs by accepting or refusing them, this may be an indication she knows – at least implicitly – what norm governs belief and what counts as a (good) reason for believing (Birch, Vauthier & Bloom, 2008; Koenig, Clément & Harris, 2004; Rakoczy and Tomasello, 2009). If this suggestion is on the right track, then perhaps we can venture the following hypothesis: 3 and 4-year old children do have an implicit grasp of what a reason for belief is. And so they *may* be able to engage in a form of inferring that requires this kind of implicit grasp.

Inference: Reasoning is a type of inferential process. Inference that happens in reasoning is a causal process among attitudes. In understanding the inferential process that goes on in reasoning, not any causal process will serve our purposes, however; what we need is an account of inference among attitudes that can avoid deviant causal chains, and distinguish inference from associative thinking. I have a suggestion that could work in that direction and is compatible with recent views on basing (Goldman, 1979: 346;

Turri, 2011; Wedgwood, 2006): A causal relation between attitudes is an inference only when it is the manifestation of an agent's competence to produce good inferential patterns.

Let me be more specific. Inferences are subpersonal processes, and by this I mean that we have no access to them. That said, whatever mechanism brings about our inferences, it can be seen as what allows instantiating the agent's¹¹⁹ cognitive competence to either implement good rules or establish causal patterns that conform to good inferential patterns. I don't have a way to fully explain what good patterns are here; however, it is not unreasonable to suppose that inference is good when it produces good reasoning (for those creatures who can reason). That does not mean you need to be able to reason to make a good inference. However, the notion of good inference becomes explanatorily dependent on that of reasoning (see next chapter for more).¹²⁰

Now, the aim of reasoning is what determines what good vs. bad reasoning amounts to. It is also natural to suppose that good reasoning sometimes falls into *patterns* (McHugh and Way, 2016). Intuitively, if reasoning has a constitutive point or aim, a good pattern of reasoning (or "principle of revision" – Harman 1986) is a pattern that can, other things being equal, be used to achieve that aim. On the view I propose reasoning aims at a conclusion that is supported by reasons. Thus, correct patterns of reasoning are those patterns which normally allow the formation of a conclusion that is actually supported by the premises (if these are fit to be reasons, and let's agree that for an attitude or a content of an attitude to be a reason it has to be the case that the content is true or the attitude is factive.). To illustrate, Modus Ponens is a good attitude-forming pattern

¹¹⁹ I remain agnostic on whether this includes non-human animals. It seems intuitive that it includes children. But I will put this issue aside.

¹²⁰ This introduces a circularity in the definition of reasoning, but not a vicious one.

because it allows to produce attitudes conclusively supported by reasons when the premises are true. In contrast, affirming the consequent is bad deductive reasoning because it does not transmit conclusive support from premises to conclusion, even when those premises are true beliefs, and thus fit to be reasons. So my proposal is that a causal relation among attitudes is not an inference, but is a deviant causal chain or an instance of associative thinking, when it is not a manifestation of the agent's competence to implement/conform to good inferential patterns. This claim, however, needs to be substantially refined. In particular, to make this work I need to say more about how this account of inference is different from my proposal about reasoning. Unfortunately, this will be the task for another time.

REASONS VS. CORRECTNESS: WHEN IS REASONING GOOD?

Chapter 4

What is reasoning? When is reasoning good? In this chapter I will argue again for the view of reasoning I called the Reason View: reasoning is the process of attitude revision that aims at a conclusion supported by at least sufficient reasons. Here I argue for the Reason View by expanding on some of the arguments presented in the previous chapter. In particular, I will show that the Reason View can make sense of what we mean by ‘good reasoning’, and that the Reason View is more plausible than its immediate competitor, i.e. the Correctness View: the view that says that reasoning aims at forming attitudes that are correct.

REASONING AND ITS CONSTITUTIVE AIM¹²¹

Take the following set of propositions:

It is raining
If it is raining, I will get wet
Therefore, I will get wet

When is the transition from the first proposition to the last an instance of reasoning? The reader will recall that I started with a pre-theoretical understanding of ‘reasoning’. Thus, following others in the literature, I take reasoning to be a mental process in which we operate on contents of mental attitudes (Broome, 2013). Contents, I assume, are propositions such as ‘It is raining’. They are the contents of belief-like and intention-like mental attitudes. The output of reasoning also has to do with the content of our attitudes. More specifically, the output of one’s reasoning may be that one comes to have a new belief-like and intention-like attitude; or the output of one’s reasoning may be

¹²¹ For other views of reasoning see, for instance, Broome, 2013; Boghossian, 2014; Neta, 2013; Wedgewood, 2006.

that one *abandons or withholds* such an attitude; or it may simply be that one *reaffirms* an attitude one already has (Wedgwood, 2006).

There are two types of reasoning: theoretical and practical. It is likely that the difference between practical and theoretical reasoning lies solely in the conclusion state (Broome, 2013; Harman, 1986:2). Practical reasoning is the kind of reasoning whose conclusion state is an intention-like state (or the abandonment of it)¹²², and theoretical reasoning is the kind that has beliefs-like states (or the abandonment of it) as its conclusion. My account of reasoning aims at offering a unifying view of reasoning, one that works for both theoretical and practical reasoning. The reason why I talk about intention-like and belief-like states is because I assume that reasoning can involve not only belief and intention, but also cognitive and conative attitudes such as assumptions and suppositions. That said, for simplicity here I will mainly focus on reasoning that involves beliefs and intentions, and make reference to other attitudes when it becomes relevant.

Reasoning is a kind of attitude revision (Harman, 1986), but there are kinds of attitude-revision processes that are not reasoning. Our attitudes might change through simple association, or they may change through some inferential process we have no access to. Note that these unconscious inferential processes are *based* on reasons, i.e. the considerations in light of which an agent arrives at a certain conclusion, just as much as reasoning. However, I contend that not all cases of inferential basing are instances of reasoning. Here I am interested only in a specific kind of inferential activity and as a result I see the term ‘reasoning’ as applying narrower than the term ‘inferring’. Why?

¹²² Some philosophers think that practical reasoning concludes with an action, but I doubt this makes much of a difference for my analysis here.

Well, first, reasoning is a “person-level” process in which premises and conclusions, the attitudes we operate from and arrive at, are *conscious* (Boghossian, 2014: 2; Pettit, 1993). And by ‘conscious’ I mean, roughly, that they are available for report. In contrast, inference and basing can at times be unconscious (Evans, 2013). What’s more, reasoning is an activity, in the sense that reasoning is “attention hogging and effortful” (Boghossian, 2014: 2; Haidt, 2001: 818). Relatedly, reasoning is not something that simply happens to us and we have no control over, but something we do and take part in. In contrast, we seem to have little control over unconscious basing and unconscious inferential processes in general. Finally, as explained previously, another element of distinction between reasoning and inference is that the result of reasoning is something we necessarily endorse. Roughly, endorsing a conclusion X means that we control X and take responsibility for it. I think there is strong evidence that some of our attitudes are irresponsive to reasons to the point that the agent feels she has no control over them, and is reluctant to take responsibility for them. In contrast, there seems to be attitudes we actually have control over. These are the beliefs or attitudes we see as reflecting our rational agency (Moran, 2001; Boyle, 2009). These are, in other words, the attitudes we endorse. I contend that the standard way to endorse them is to form them through reasoning. In contrast, when we arrive at an attitude through unconscious inference, we may not necessarily stand behind it. And this marks a difference between the two processes (but for more see chapter 3).

I argued before that we should explain reasoning as a process of attitudes-revision guided by a constitutive aim. The aim of reasoning is an aim of the agent, a goal the agent necessarily has when reasoning. An aim is a desire or motivation to obtain a certain

result. And a desire is – at a bare minimum – a disposition to do certain things. That is, having an aim or goal means being motivated to do certain things and not others, prevent certain occurrences to happen, change things around if the aim is not satisfied, and so on.

As stated in the previous chapter, saying that reasoning is guided by an aim does not sound implausible.¹²³ Second, the claim that reasoning has a constitutive aim makes sense of the fact that, intuitively, reasoning is an activity; non-reasoning processes, in contrast, simply happen to us (Broome, 2013). Another important point: the idea of reasoning being guided by an aim seems helpful to make sense of the highly plausible claim that reasoning is a thought process in which the conclusion we arrive at is a conclusion we *endorse*, i.e. not simply something we happen to reach, but a conclusion we stand behind. The idea that reasoning has an aim can make sense of this because the attitudes we reach through reasoning are attitudes that (seem to us to) fulfill the aim we have when reasoning; they fulfill our goal and thus we endorse them (with some caveat - more on this below).

Now I will also show that reasoning's aim can offer a plausible account of correct or good reasoning. Plausibly enough, reasoning is constituted by (causal-inferential) *transitions* among attitudes. In particular, in reasoning premise-attitudes cause a conclusion-attitude (or lack thereof). Some of these patterns or transitions, such as those which constitute affirming the consequent, or the gambler's fallacy, lead to bad reasoning. Modus ponens and modus tollens are, in contrast, good patterns. The aim view is able to make sense of this distinction. On this view, a good *pattern* is what *normally* allows reasoning to fulfill its aim, when the starting points of the transition, i.e. the

¹²³ Also McHugh and Way (2016) argue that reasoning has a constitutive aim. See also Whiting (2014).

premises of the reasoning, are good (McHugh and Way, 2016). In addition, this promises to shed some light also on what constitutes good basing too. Indeed, as I mentioned above, in the inferential transitions that constitute reasoning, we base our conclusion off of the premises, i.e. the reasons in light of which we come to that conclusion. A good basing-relation is thus a transition that instantiates a good pattern.

Now a question naturally arises: what is the aim of reasoning? And when is a transition among attitudes a good transition? In what follows, I will first examine a very intuitive proposal: reasoning aims at forming correct attitudes. After criticizing that view and its solution to the problem of good reasoning, I will move on to defend the Reason view: reasoning is the process of attitudes-revision that aims at a conclusion supported by at least sufficient reasons.

THE CORRECTNESS VIEW

In the previous chapter, I introduced the correctness view: reasoning aims at forming correct attitudes and dropping incorrect ones. ‘Correctness’ here is a normative property of attitudes such that, roughly, they fit the world right given the type of attitude they are. Recently, McHugh and Way (2016a) have offered a new version of the correctness view by developing an account of good reasoning. They see correctness as the same as fittingness. As they put it, “[t]he core idea is that attitudes are associated with standards for their objects. An attitude is fitting when its object meets its standard.” (2016: 13) A belief is fitting if and only if true; an intention is fitting iff permissible. Based on this, they develop a view of good reasoning which can be roughly summarized as follows: “if we start from fitting attitudes, we can expect good reasoning to lead us to

further fitting attitudes, other things equal.” (2016: 13) They also make the point that good reasoning falls into patterns, and that we can understand what good reasoning is in terms of good patterns of reasoning, i.e. patterns that on their view are fitting-preserving.

The Correctness view is highly intuitive. It does seem that when we reason we want to get at the world, and we want to get things right. And yet I have some worries about the correctness strategy in general, and McHugh and Way’s account of good reasoning, in particular. First, let me expand on some of the worries I raised in the previous chapter. One worry is that the correctness-aim may not be *necessary* for reasoning since we use reasoning also to form attitudes that have no (clear) correctness conditions. I am thinking of suppositions, acceptances in context and even bits of imagination. Relatedly, many subscribe to the idea that the standard of correctness for belief is shared by other attitudes, such as guesses and conjectures (McHugh, 2011: 385). But it just seems false that in reasoning we aim at forming true guesses. If we did, we would do our best to ensure their correctness. But we don’t: we make guesses and conjectures for all sort of reasons that have little to do with their truth. So the aim of reasoning does not work for those attitudes either.

Also, the correctness proposal does not clarify *why* reasoning would always aim at forming attitudes that are *correct*. As I mentioned before, it is in fact possible to imagine someone arguing that there is more to attitudes than correctness. Indeed, one may think that attitudes can be rational based on moral and practical values. This of course does not show that the aim of reasoning is not correctness, but it leaves open the question why we would *always* care about correctness when there may be other more important values to pursue.

One reply to this is to deny that there is anything beyond correctness: attitudes can't be practically good or bad. This view may work for belief, but I think it would be a stretch to say that moral norms or values don't apply to attitudes in general. A second reply is that in reasoning we *can't* aim at those other values. Our psychology is such that we are never motivated by practical or moral considerations when forming (or dropping) an attitude through reasoning. And we can't try to get something impossible to obtain. On this view, though, the aim of reasoning becomes a triviality as it is just the result of the fact that one can't believe, intend, desire, or wish at will: correctness is the only game in town for us. Unfortunately, this seems to substitute one mystery for another: why can't we form attitudes at will?

A better strategy is to say that correctness is a constitutive norm for attitudes, and that when we try to form those attitudes we can't but aim at satisfying that norm (Shah, 2003). This point has been made about belief: the constitutive norm of belief is that it is correct to believe a proposition if and only if true. The proposal is that, when forming beliefs, we are necessarily bound by this norm. Note, however, that no other constitutive norm seems to work that way. The constitutive norm of assertion, for instance, says that one ought to assert if and only if one knows. But when deliberating about assertion, one can form an assertion even when one is well aware one lacks the required knowledge for it.¹²⁴

Finally, the correctness view could try to appeal to the idea that reasoning is an activity of attitude-revisions and - this is the key point - attitudes are functional kinds: their constitutive function is to be correct (McHugh and Way, 2016b). In addition, one

¹²⁴ This point is widely recognized in the literature, and explicitly made by Steglich-Petersen (2006) and McHugh (2011), so I won't dwell on it any longer. See Introduction for more.

may try to argue that one can't try to build an object with a constitutive function *unless* one is trying to build an object that satisfies that function (Shah, 2003; Korsgaard, 2008: 113). Here is an example. Good houses provide shelters. The aim of house-building is thus to provide a shelter, and one can't build a house unless one is trying to build a shelter. But this strikes me as very controversial. It is true that when I build a house, I am not completely indifferent to the standards of good (shelter-) house-building. This is because, if the house I build is too shabby, then it will likely fail to be a house at all (e.g. it won't have a roof, it will fall apart). But this is different from saying that, when I build a house, I can't aim for a house that does not fulfill its function. It could well be that my only aim is to build a house that looks very nice from the outside to make my neighbor jealous, and that I care very little about whether it actually provides shelter or is a good house. It is unclear why, because of that, mine would not count as an instance of house-building.

As a reply, one may point out that objects get their constitutive function only (or mainly) from their creator (be this natural selection or human intention). So a house built without the intention of building a shelter is an object which does not have the *function* of being a shelter. Since having that function is constitutive of being a house, an object does not belong to the kind 'house' unless produced with that goal in mind.

I am skeptical this will work for mental attitudes, though. Even assuming that attitudes do have correctness as their constitutive function, this means that all the processes that lead to attitude-formation must have the function of forming correct attitudes. More specifically, any inferential process that leads to attitude revision should have that function. So one is bound to ask: what is the difference between *those* processes

and reasoning? Indeed, it would seem that, on this view, *any* process of attitude revision with the aim or function of forming correct attitudes would count as reasoning. But that seems wrong. As argued above, reasoning is distinct from other processes of attitude revision. The correctness approach, in contrast, blurs this distinction.

To block this worry, the correctness theorist will say that the aim of reasoning is a personal aim. When it comes to subpersonal mechanisms that form attitudes, they have an aim in the sense that they have a biological *function*, i.e. to create attitudes that are correct. But this is far from obviously correct. My mind can create attitudes, e.g. beliefs, through associative thinking or as a result of emotions, i.e. via processes that do *not* have the biological function of forming correct beliefs. Would they not count as beliefs simply because created through association? This is a quite strong claim to make, I believe, and a view that avoids such a strong commitment is to be preferred. So I am not sure the correctness view is as feasible as it originally appears to be.

Things are also problematic when we consider how the correctness view deals with good reasoning. As mentioned above, McHugh and Way's recent paper on good reasoning clearly spells out how good reasoning can be explained in terms of correctness or fittingness preservation. As they put it, "The move from $P_1 \dots P_n$ to C is a good pattern of reasoning iff, and because, other things equal, if $P_1 \dots P_n$ are fitting, C is fitting too" (2016: 13). This is a very intuitive proposal. Take modus ponens: it is plausible that the reason why Modus Ponens constitutes, intuitively, a good pattern of reasoning is because it gets you from correct/fitting premises to a correct/fitting conclusion. Similarly for pragmatic reasoning: if it is correct to intend to A , then it is permissible to A . And if it is permissible to A and M is a necessary means to A , then it is also permissible to M . Thus

it is correct to intend to M. Thus the view can make sense why means-end reasoning is a good pattern (McHugh and Way, 2016: 13).

Although I find their view of good reasoning quite intuitive, I have worries. My first worry emerges out of the point made earlier that guesses and conjectures' standard of correctness are the same as belief's. If this is right, on the correctness view of good reasoning the following would count as a good pattern of reasoning:

(1) Guess (p), Guess (if p then q) > Belief (q)

Recall that on their view a pattern of reasoning is good when other things equal it gets you from fitting premises to a fitting conclusion. Applied to the case above, it is true that if my guess that p is true, and my guess that if p then q is also true, my belief that q will be true too. In this case, that is, correct or fitting guesses get us correct or fitting beliefs, but the reasoning above is intuitively far from good. More generally, it seems that, at least for theoretical deductive reasoning, their view is unable to distinguish between valid logical transitions among well-formed formulas and good reasoning. This leads to problematic patterns such as (1).

Relatedly, we have the following intuitively good pattern of reasoning:

(1-A) Belief (q), Belief (if p then q) > Guess/Conjecture (p)

There seems to be nothing wrong with (1-A), and yet it does *not* instantiate a fittingness-preserving pattern. Notoriously, affirming the consequent is not a valid logical inference, but this does not mean that (1-A) is not a good pattern of reasoning, contrary to what the Correctness view anticipates.¹²⁵

¹²⁵ McHugh and Way will point out that “good patterns of reasoning are not fittingness-preserving only in virtue of the premise- or conclusion-responses. They preserve fittingness (other things equal) in virtue of a relationship between the premise- responses and the conclusion-response.” (2016: 16) But this does not rule out (1): in (1) it is because my premises are true that my belief is true.

Another source of concern I'd like to raise is that McHugh and Way's understanding of good reasoning seems to categorize the following as good reasoning:

(2) Belief (p), Belief (if p then q) > Belief (q) [However, I am aware I have good reasons to believe that one of the premises of my modus ponens is false.]

On McHugh and Way's view, in (2) I am reasoning well because I am instantiating a good pattern. But I think this is a mistake: we should keep good reasoning separate from good patterns. Good patterns have to do with inferential transitions among attitudes. Transitions can be unconscious and thus not be part of reasoning processes. In contrast, as I mentioned above, reasoning goes beyond inference, as it is an inferential process guided by an aim. In (2) above, it seems intuitive to say that I am making the right transitions: the inference is good. At the same time, though, in (2) I am not reasoning well, as I am not weighing all my reasons correctly. The view I am about to offer can make sense of the intuitive idea that there is a distinction to be drawn between good inferential transitions and good reasoning. And this is why I believe it should be ultimately preferred over the Correctness view.

THE REASON VIEW¹²⁶

Reasoning, I claimed, is guided by a constitutive aim or goal. This aim is what makes reasoning a unique type of inferential thought process. The view I support says that **the aim of reasoning is arriving at a conclusion supported by at least sufficient reasons**. The conclusion of reasoning consists in dropping, reaffirming, forming or

¹²⁶ This view has been adopted in passing by a number of authors such as Boghossian (2014), Dancy (2004), Pollock (1987). I argue for this view in chapter 3.

suspending (avoiding forming) an attitude. Doing one of those things in a way that is supported at least by sufficient reasons is the constitutive aim of reasoning.

By “reason” here I mean normative reasons. Normative reasons are things that count in favor of (or against) a certain response. By “response”, I mean the holding, withholding or dropping of an attitude. Also, I will assume a *weighing approach* according to which reasons support responses by contributing a kind of epistemic *weight* to them.¹²⁷

Reasons have different strengths and speak in favor or against responses in different ways. As a result, in deliberation they tend to be weighed against one another. Pro-tanto reasons are reasons that have genuine weight in favor of some response. Pro-tanto reasons may be outweighed by other opposing reasons, though¹²⁸. I will assume that an agent S has sufficient pragmatic reason to ϕ only when S has a reason to ϕ , and no stronger reason not to ϕ . In contrast, I assume that having sufficient evidence to form a belief that p involves more than having stronger evidence for p than for not-p; it requires some high degree of credence that p is true. I will also assume that an agent S has decisive reason to ϕ only when her reasons to ϕ are stronger than her reasons not to ϕ .¹²⁹ Finally, conclusive reasons are reasons that mandate a response. So typically it is said that if S has overall conclusive reason to ϕ , then her ϕ ing is normatively required. Also, typically it is said that if S has overall sufficient reason to ϕ , then her ϕ ing is justified, warranted or normatively permitted.

¹²⁷ I assume that reasons are always reasons *for* someone and that they favor types of response-attitudes, not propositions.

¹²⁸ In contrast, *prima facie* reasons are considerations that appear to be reasons, but may actually not be reasons at all.

¹²⁹ This way of defining sufficient and decisive reasons is controversial. However, that does not matter for my purpose here.

Now, I doubt that in reasoning we *always* aim at reaching conclusions supported by conclusive reasons. Deductive reasoning is definitively governed by that aim, but not all reasoning is that way: we can correctly form an intention to do something based on reasoning, even if that action is not mandated by conclusive reasons but only permitted by sufficient reasons. The aim of reasoning should thus account for those kinds of cases as well. I also doubt that the reasons at play in reasoning can be only pro-tanto, as I doubt that when we reason we usually aim at a conclusion that is supported by reasons that could be outweighed by other reasons we have. In contrast, what we aim for in reasoning is a conclusion supported at least by *sufficient* reasons.

What kind of reasons my conclusion should be responsive to when reasoning? Any reason there is, or the reasons *I* have? I have already tackled this question in the previous chapter, but let's expand on this more, as views differ on this issue.

Many distinguish the following two things: the fact that there are reasons to do or believe something and the fact that I have those reasons or that those reasons are available to me. Those who endorse this distinction tend to hold that reasons are true propositions or facts, and that to have those reasons one needs to stand in some relations to those facts or true propositions, or at least that those reasons are available to me. (I will assume, somewhat vaguely, that a reason is available to me when it is a true proposition or fact perceptually within my reach and/or easily inferable given what I know¹³⁰.)

If we accept this view, then the aim of reasoning is to get to a conclusion that correctly responds to the reasons that are available to us. The reasons available to us, however, are potentially only a subset of the reasons that there are for us to reach a

¹³⁰ For something very similar see Kiesewetter (ms) where he explains available reasons as evidence-based reasons.

certain conclusion. We may (wrongly) think that, when we reason about some issue, we want to reach a conclusion that is supported by the balance of *all* the reasons. Unfortunately, though, there are reasons we may never get to make ours and that are beyond our reach. In those cases, it is unclear that, when we reason, we can aim at anything else than being supported by the reasons we have. Why? Well, only *available* reasons can guide us in deliberation and reasoning, whereas reasons we have no access to can't guide us at all. So it would be indeed odd to insist that, when reasoning, we aim at arriving at a conclusion that is supported (also) by reasons we don't have, and that can't function as our base when reasoning.

If the reader is still not convinced that the aim of reasoning should be restricted to the reasons available to us, let me offer two examples in support of my claim. In the famous "Jill case" (Jackson, 1991: 462–463), Jill is a doctor who faces a difficult choice. She wants to cure a patient with a mild skin problem. She can choose among 3 different drugs: Drug A is very likely to relieve the patient's symptoms but will not completely cure him. Either one of drugs B and C will completely cure the skin condition whereas the other will kill the patient, and there is no way that she can tell which of the two will kill him (unbeknownst to Jill, C is the cure for the patient). Were Jill to deliberate what to do, how can she go about doing that? She clearly can't access all the reasons there are to make that choice, and she knows that. The only thing she can do is to use what she has, and aim at making the choice supported by the reasons available to her. Even if she *knows* the conclusion she will reach is only half-good because of the limited information she has, that's the best she can do in that situation. There is no other way.

Take the following somewhat similar example. I know that the coin that was in

front of me a few minutes ago has now been tossed. I know it came out either head or tail but I don't know which. In this case, when reasoning about what to believe concerning the status of the coin tossed, I can't aim at reaching a conclusion supported by reasons I have no access to. What I do in this case is to suspend judgment because, given the evidence I have, I am not allowed to form a belief at all. Again, the only thing I can do is to go with the reasons available to me. So this example supports my claim that, if one wants to keep the distinction between the reasons one has and the reasons that there are for someone, the aim of reasoning can only concern the former and not the latter.

Not everybody, however, is comfortable with that distinction. In the literature on reasons, there is a position that is at times called 'prospectivism'. Roughly, the idea is that normative reasons are *only* either factive psychological attitudes (e.g. true beliefs, knowledge, evidence) of the agent or facts or true propositions that the agent has easy access to. On this view, no reason is not-had by the agent, or not available to her.¹³¹ This prospectivist position squares nicely with the cases in which our epistemic position prevents us from accessing information relevant to make a decision. Accordingly, on this view reasoning aims at arriving at a conclusion supported by at least sufficient reasons, and reasons here are only those considerations already available to the agent.

There is a third option: some theorists adopt a view that says that there are two reason-relations, i.e. the objective reason relation and the subjective reason relation. An objective reason is a reason there is *for* someone to do or believe something; a subjective reason is a reason someone *has*. Objective reasons are factive. Subjective reasons are mental states and may be non-factive (e.g. Parfit 2001, 2011; Schroeder, 2008; Way,

¹³¹ This is often put in terms of 'ought' rather than reason. Prospectivists about 'ought' such as Ross believe that 'ought' depends on the agent's beliefs, or evidence available to her.

2009). For example, Derek Parfit (2011:35) claims that “when we have beliefs whose truth would give us a reason to act in some way, we have [...] an apparent reason to act in this way.” The key here is that proponents of the idea of subjective/apparent reasons seem to hold that claims about subjective/apparent reasons are *genuinely normative* claims. Parfit (2011) maintains that “normativity involves reasons or apparent reasons”. As in the following famous case: the fact that the glass contains poison rather than gin is an *objective* normative reason for Bernard *not* to drink it, even if he is not aware of this fact. In contrast, his belief that the glass contains gin is a subjective reason he has to drink. The two types of reasons come apart.

Subjective/ apparent normative reasons are reasons that appear to the agent to be objective reasons or that would be objective reasons were they true¹³². As a result, the introduction of subjective/apparent normative reasons is supposed to capture the idea that in the famous gin case there seem to be, as Schroeder puts it, all the “earmarks” that Bernard has a reason to drink. For starters, we *expect* Bernard to drink from the glass given his beliefs and his desire to drink gin. What’s more, we would reasonably criticize Bernard if he were not moved to drink from the glass: he doesn’t know that it contains gasoline, so why is he not drinking it? (Schroeder, 2008: 61). So the non-factivists are holding on to the idea that Bernard has reasons to drink, even though there are reasons for him not to drink.¹³³

¹³² There are actually at least two different interpretations of what subjective/ apparent reasons are. The *de dicto* view maintains that R is an apparent normative reason for agent A to F iff it appears to agent A that R is an objective normative reason to F. This view is held mainly by Scanlon (1998) and Kolodny (2005). Others, such as Parfit (2001, 2011) and Way (2009), argue for a *de re* view: R is an apparent reason for A to F when R’s truth would give A an objective reason to F, and it appears to A that R is true.

¹³³ This distinction is orthogonal to the distinction between normative and motivating reasons. The point of making such a distinction is not that Bernard has a motivating reason to drink from the glass. Rather, the

I won't try to offer an assessment of the subjective/apparent reason view here. To me the most natural view about normative reasons is that they are facts or true propositions. So for simplicity here I will adopt the view that in reasoning we aim at forming a conclusion supported by facts or true propositions available to the agent. As a result, the **Reason view** I am defending here says that **in reasoning we aim at a conclusion on balance sufficiently supported by the facts or true propositions available to us.**

GOOD REASONING AND GOOD INFERENTIAL PATTERNS

Now that I have explained the main tenets of the Reason view, I will argue that it can perform better than the Correctness view on a number of important issues, especially concerning good reasoning and good patterns of reasoning.

First, the Reason view is very plausible. That is, it is plausible that agents aim at reaching conclusions supported by sufficient reasons. If an agent were to reason from premise P to conclusion C while also saying that P offers no (sufficient) support to C, we would doubt this agent is reasoning at all. The point of reasoning seems indeed to offer support for a conclusion, and disavowing this aim is tantamount to stop reasoning.

Second, the Reason view does better than the Correctness view in explaining why the conclusion of our reasoning is a conclusion we necessarily endorse. Arguably, not all the goals we have are goals we *want* to have (or that we endorse). We may be engaging in an activity (e.g. preparing a bomb) because forced to do so (e.g. by terrorists). This is not an aim we wish we had, so its conclusion is not obviously something we endorse. As

point is that believing that the glass is full of gin makes it *reasonable* for Bernard to drink from it, even if he wasn't actually motivated by that.

pointed out above, the Correctness View owes us a story for why, if the aim of reasoning is arriving at a correct attitude, this is an aim we necessarily endorse. Indeed, there may be situations in which we think we morally or prudentially should, e.g., believe something our evidence indicates to be false.

In contrast, the Reason view has no problem with that. On the Reason view, the aim of reasoning is not just any old aim: it is the aim of reaching a conclusion supported by (at least) sufficient reasons. The reason why reasoning forms attitudes we endorse is because reasoning *is* the process *aimed* at forming attitudes supported by reasons - so arriving at the conclusion C of a reasoning process means endorsing C *because* one necessarily sees C as rational/supported by reasons. Thus, it is the fact that we reasoned to C that expresses our endorsement of C.

Third, contrary to the Correctness view, the Reason view has no problem accommodating reasoning toward attitudes that have no correctness conditions (e.g. suppositions, acceptances), since these are attitudes held for reasons just as any other attitude. Similarly, the Reason view tells us why reasoning toward a belief requires more support than forming a guess: the amount of reasons needed to support a guess is different than the amount of reasons required by beliefs.

Fourth, the view I am offering here fits nicely with the view that the notion of a 'reason' is primitive or indefinable (Parfit 2011; Scanlon 1998). But it is also compatible with the idea that reasons are reducible to oughts. This means that the Reason View is not necessarily committed to a specific normative framework.

Finally, the Reason view offers a plausible account of *good* reasoning. Roughly, the Reason view will say that good reasoning is reasoning which correctly calculates the

weight of the available reasons. Thus, good reasoning is, in general, reasoning in which an agent is able to assign the appropriate weight to the reasons available to her, and produce the appropriate response on balance. The goodness of reasoning is defeasible: even if your calculation is right, you may believe something false. To see what I mean take the gin case again. On the Reason view, Bernard's reasoning is good; that is, he does not have sufficient reasons to drink from the glass because he falsely believes that the glass contains gin; however, given the evidence he has, he seems to be doing everything right in terms of calculating the weight of the information at his disposal. Or take the Jill case. Jill's reasoning to give the patient drug A is good. In fact, Jill has sufficient reasons to prescribe drug A given the evidence available to her, even if drug C is what would in fact cure the patient's illness.

There is a further layer of assessment for reasoning: reasoning happens by means of inferential *transitions* among attitudes, from premises to a conclusion. So one important question is, when are these transitions good?

Two clarifications before starting . First, reasoning's premises are attitudes with a propositional content, but not just any propositional attitude works well as a premise: perception and belief can work as premises, guesses cannot. In addition, a premise can support a conclusion by being a reason one has for that conclusion, and that requires (at least) that the premise is an attitude of right kind (e.g. perception, belief), and that its content is true.¹³⁴ As I mentioned above, reasons are factive, so a premise supports a conclusion only if true. In the next chapter, I will say more about belief and its role in reasoning.

¹³⁴ Some require that the attitude in question is sufficiently justified, or amounts to knowledge. For simplicity, I will put these views aside.

I believe that an inferential transition is good when it instantiates a good *pattern*. A good inferential pattern is a transition of attitudes that, other things being equal, goes from some premises to a conclusion and, if those are true, the conclusion is sufficiently supported by those premises.

Importantly, instantiating a good *pattern*, making a good inference, is *not* the exact same as reasoning well or correctly. So, for instance, a transition among beliefs that instantiates a Modus Ponens is a good pattern of reasoning, but adopting it does not mean one is reasoning correctly. As mentioned above, if one has good reasons to believe that one of the premises of her reasoning is false, correct reasoning requires that she does not adopt those premises.

Let us see what that means in detail by offering a definition of good patterns.

Good Pattern of deductive reasoning:

The move from $P_1 \dots P_n$ to C is a good pattern of deductive reasoning iff the theoretical conclusion C is conclusively supported by $P_1 \dots P_n$, when $P_1 \dots P_n$ are fit to be normative reasons we have (i.e. they are true beliefs).

Intuitively, if reasoning has a constitutive point or aim, a good pattern of reasoning or transition is a pattern that can be used to achieve that aim. On the view I propose, deductive reasoning aims at a conclusion that is supported by conclusive reasons. Thus, a good pattern of deductive reasoning is any pattern of reasoning that will yield a conclusion that is supported by conclusive reasons, when one in fact has those reasons and uses them as premises in reasoning. Hence, good patterns of deductive reasoning are those patterns which allow the formation of a conclusion that is actually conclusively supported by the premises if these are fit to be reasons.

Importantly, as Harman (1986: 5) has pointed out, logic principles do not equal good patterns of reasoning. So we cannot simply adopt rules of logic and assume that they constitute good deductive-reasoning patterns. So for instance, beliefs that are jointly inconsistent imply any proposition. But it is not good reasoning to infer any proposition whatsoever from those inconsistent beliefs. I believe my view accommodates Harman's point.¹³⁵ Inconsistent beliefs (or propositions) do not support anything as they cannot be both reasons for believing. So reasoning from them is not good reasoning. In contrast, syllogisms constitute patterns of good reasoning: Clara believes that Socrates is a human being. She also believes that all humans are mortals. Those propositions can be on balance conclusive reasons for her to believe that Socrates was mortal. Now, this is a correct pattern of reasoning because, given the premises "Socrates is a human being" and "humans are mortals", it produces a conclusion *prima facie* justified by those premises (It is only *prima facie* justified because the premises may be false, and thus not reasons for the conclusion.) Similarly, Modus Ponens is a good pattern because it is a transition that allows to produce attitudes conclusively supported by reasons, when the starting point is good. In contrast, affirming the consequent is a bad pattern of deductive reasoning because it does not transmit justification from the premises to the conclusion, even when those are fit to be reasons.

Good pattern of inductive reasoning:

The move from $P_1 \dots P_n$ to C is a good pattern of inductive reasoning iff theoretical conclusion C is *other things being equal* sufficiently supported by $P_1 \dots P_n$, when $P_1 \dots P_n$ are fit to be normative reasons we have.

¹³⁵ My proposal should in principle be attractive also to those who see material relations as material inferences, and rescind good reasoning from logically sound reasoning (Brandom, 1994).

Thus, enumerative induction is a correct pattern of reasoning when it produces conclusions sufficiently supported by the premises, if other things are equal and the starting point is good. Alternatively, we incur in fallacies like hasty generalization: the conclusion is poorly supported by the premises because – even if true – these are *not* sufficient reasons for the conclusion. This is because basing one’s belief (e.g. all students in this college are nice) on a too small of a sample (e.g. the students in my class are nice) is a pattern of reasoning that yields a conclusion poorly supported by the premises. Indeed, here the pattern that connects premises and conclusion does not establish sufficient support for the conclusion. In contrast, basing one’s conclusion off of a large randomized sample of cases makes that conclusion defeasibly well supported. So enumerative induction with a large randomized sample of cases usually leads to good reasoning because it is a pattern that yields conclusions that are defeasibly well supported, if the premises are fit to be reasons.

It is worth drawing the reader’s attention to the fact that my view is compatible with the claim that patterns of good inductive reasoning are defeasible. It can be a good pattern of inductive reasoning to move from some premise to a conclusion even if it is not good inductive reasoning to make that move *considering* some further premise-responses. For instance, it might be a good transition in inductive reasoning to move from the belief that someone said that it will rain tomorrow to the belief that it will rain tomorrow, even if it is not good inductive reasoning to make this move given the additional belief that the source of testimony is unreliable (McHugh and Way, 2016).

Good pattern of practical reasoning:

The move from $P_1 \dots P_n$ to C is a good pattern of practical reasoning iff practical conclusion C is *other things being equal* sufficiently supported by $P_1 \dots P_n$, when $P_1 \dots P_n$ are fit to be normative reasons we have.

Patterns of reasoning such as means-ends reasoning are then good patterns of practical reasoning because – when premises are true or well supported – they produce a response which is (defeasibly) supported by reasons. Indeed, if you intend to make some tea and boiling the kettle is a means to achieving that end, you have reasons to boil the kettle. By adopting means-end reasoning one is then able to reach a conclusion supported by reasons. Now, of course, if it turns out that drinking tea makes you sick, it appears your intention to boil the kettle is *not* well supported after all. But, as mentioned above, the goodness of the pattern of reasoning is independent of the goodness of the input provided: one can conform to a good pattern of reasoning while starting from bad (i.e. false, defeated, undercut) premises. As a result, my proposal is also compatible with Bernard Williams' (1979) famous example that I already mentioned above: you intend to drink gin and believe that there is gin in your glass. It is a good pattern of reasoning to move from these attitudes to intending to take a sip from your glass, even if in fact there is petrol in your glass. My proposal can make sense of this intuition: the transition from premises to conclusion instantiates a good pattern of reasoning because, if the premises were reasons he had, the conclusion would have been well supported. The fact that this was not the case has little to do with the goodness of the transitions happening in reasoning.

OBJECTIONS

Reasoning with Wrong Kinds of Reasons: For the Reason view to be able to explain good vs. bad reasoning, it has to commit to the controversial distinction between right and wrong kind of reasons. The reason for this is that, if it does not do so, such a view would be forced to say that, for instance, I desire to believe that p, therefore I believe that p is good reasoning. And this is intuitively unacceptable (McHugh and Way, 2016: 11).

Contrary to this objection, I don't think that the Reason view is forced to accept the distinction between right vs. wrong reasons. Indeed, I believe it can happily stay neutral on this issue. Why? The short answer is that "I desire to believe that p, therefore I believe it" is not good reasoning because it is not a transition we can actually bring about in reasoning. Admittedly, justifying this claim requires a long argument; but here is a sketch of it: in reasoning we can't aim at forming a belief with disregard for its being an appropriate default premise to be used in further reasoning. This is because belief's functional role *is* to be a default premise in reasoning, and in reasoning we can't try to form a belief with disregard of its role. And to be able to be *appropriately* used as a default premise, a belief needs to be epistemically sufficiently supported. So in reasoning we must aim at forming a belief epistemically sufficiently supported. This is why "I desire to believe that p, therefore I believe it" is not reasoning at all: it is simply not a transition that constitutes reasoning.

Admittedly, this is all very quick, but let me try to explain the rationale for this idea. The aim we have when we form beliefs through reasoning is constrained by a more

fundamental aim we have when forming attitudes in general, namely that those attitudes need to be supported by sufficient reasons. Now, belief has a special role vis-à-vis reasoning: it functions as a default premise. As a result, to use a belief appropriately in reasoning one needs sufficient evidence for it, evidence that indicates that the belief is true, and that it can be a reason one has. This is why the way we can form beliefs in reasoning is constrained by the more general aim of reasoning and thus limited to the support offered by epistemic reasons.¹³⁶

The Reason view says that means-end reasoning is a good inferential pattern because, if the premises are fit to be reasons, it will produce attitudes supported by sufficient reasons, other things being equal. This seems to rest on the assumption that intentions can be reasons. This view, however, is quite controversial (Broome, 2001: 99–100).

Contrary to what it may seem, I am not committed to that view, at least not directly. I am willing to adopt a form of normative cognitivism about intentions: intending to F means believing you ought to F or that you have sufficient reason to F. So, for instance, Schroeder (2009: 237) says that “intending entails believing that you ought”, Scanlon (2007: 87–88) maintains that “insofar as a subject has [...] intentions, it must see these as responsive to its assessment of the reasons for these states”, and Korsgaard (1997: 245–246), “you must take the act of your own will to be normative for you”. When applied to means-ends reasoning we have that, e.g., if I intend to have eggs for breakfast I see myself as having sufficient reasons for that. If it is true that I have

¹³⁶ I offer a full argument for these claims in the next chapter. For a somewhat similar argument, that could be adopted here to make the same point, see Whiting (2014).

sufficient reasons to have eggs for breakfast and that I won't have eggs for breakfast unless I go to the supermarket, then I also have sufficient reasons to intend to go to the supermarket.

Gilbert Harman, Kieran Setiya, and David Velleman defend a view that goes in a similar direction: intending to F means believing one will F. At the heart of this, there is Anscombe's point that we characteristically express our intention to F by asserting "I am going to F" (1957: 1). So the fact that I am going to have eggs for breakfast and that I won't have eggs for breakfast unless I go to the supermarket, gives me reasons to believe that I will go to the supermarket. Believing that I will go the supermarket is intending to go there, so the fact that I am going to have eggs for breakfast and that I won't have eggs for breakfast unless I go to the supermarket, supports that intention. This move *de facto* transforms means-end relations into a *modus ponens* transition, which explains why means-end transitions are good.¹³⁷

¹³⁷ Although these views have raised concerns, it is worth keeping in mind that they have been independently and forcefully argued for in the literature, so my adopting either of them would not constitute an *ad hoc* move on my part.

WHAT GUIDES BELIEF

Chapter 5

“I do not think it adds anything to say that truth is a goal, of science or anything else. We do not aim at truth, but at honest justification” (Davidson 1999:461).

Truth is what guides us when forming beliefs, and this is the result of the very nature of belief. Or so many believe. In this final chapter, I raise concerns for both of these claims, and propose a different view: when forming beliefs we try to form beliefs supported by sufficient epistemic reasons, and this is the result of the general aim we have when reasoning toward attitudes and the unique role belief has in securing that goal.

Bernard Williams famously said that “belief aims at truth” (1973: 136). On a common teleologist reading of this expression, the aim of truth is constitutive of belief. That means that belief’s truth-directedness is part of the nature of belief and it is a feature that distinguishes beliefs from other kinds of attitudes.¹³⁸ Also belief aims at truth because belief is an attitude necessarily regulated by some truth directed mechanism.¹³⁹

One key argument in support of the teleologist view is that it promises to explain a puzzling phenomenon about belief (hereafter the Explanandum): the fact, as it is often put, that we can’t believe at will, or that we have limited power over our beliefs. In this chapter, I first show that at least one recent prominent teleologist’s attempt to explain the limited control we have over beliefs fails, and this calls for a better explanation of the Explanandum. Then I offer my own view: what guides belief-formation is the goal of

¹³⁸ Lynch, (2009), Railton (1994), Sosa (2008), Velleman (2000) all subscribe to this claim. Please see the Introduction for an explanation of the teleologist approach to the nature of belief.

¹³⁹ On the norm-based reading of belief’s truth-directedness, belief aims at truth because it is the attitude governed by the truth norm. On this view, the expression “belief aims at truth” does not mean that believers literally aim at having true beliefs or that belief itself has the goal of truth, but rather that truth is a norm that beliefs are necessarily governed by (Wedgwood 2002: 267). Usually, this view is coupled with attempts to ground epistemic normativity into the constitutive norm of belief. I will put this aside here, but see the Introduction for more.

having a belief only if it is sufficiently supported by evidence. This aim of belief derives from a more general aim we have when forming attitudes through reasoning (as I argued in chapter 3 and 4) and from the role of belief in reasoning (the topic of chapter 1 and 2).

WHAT WE ARE AFTER

Imagine a demon offers you eternal happiness, if you believe something you have absolutely no evidence for but only evidence to the contrary. Such a belief is nothing really important and has very little value for you. And you are fully aware of that. Unfortunately, no matter how much you'd like to, it seems you really can't form that belief just because you want to. This is, in a nutshell, the Explanandum, the phenomenon I will try to make sense of in this chapter.

We will get a clearer understanding of this phenomenon later, but before doing that let me first point out a few things. First, a number of different and competing definitions have been offered to try to capture this puzzling mental phenomenon. In fact, I am calling it "the Explanandum" to try to remain neutral vis-à-vis all the different ways people have labeled it (e.g. doxastic involuntarism, transparency, weak/strong exclusivity). In addition, it is worth keeping in mind that the discussion surrounding the Explanandum is mostly based on intuitions of what we seem able or not able to do when it comes to forming beliefs. So the reader should be aware that, as people's intuitions may differ, some disagreement on the phenomenon we are trying to pin down here is to be expected. Finally, one may think that the issue at stake is somewhat outlandish. The example of the demon may reinforce this idea: why should we care about the fact that we can't accept the demon's offer? However, real life examples abound in which we would

be better off having a false belief. For instance, we may be better off believing we will survive a terrible illness because this may increase our chances of survival. If our ability to influence our beliefs is limited, as I believe it is, then this has important consequences for our lives. What's more, the issue here has some important philosophical upshots. For instance, the Explanandum seems to point to something deep about belief's primary relation with truth, namely that truth appears to be our strongest and perhaps only concern when we deliberate and reason toward a belief. Some philosophers have gone as far as to say that this supports the view that truth is the only normative standard for correct belief.

Here I will simply focus on the phenomenon at hand and ask, what is the Explanandum, and what is it that we *can't* do when it comes to belief? Luckily, a proposal recently offered will help us get a better sense of the phenomenon we are investigating. A number of philosophers seem to subscribe to the following:

Strong Explanandum¹⁴⁰: in deliberation/reasoning¹⁴¹ non-epistemic considerations cannot motivate us, qua reasons, to directly form or withhold a belief.

¹⁴⁰ See, in particular, Kelly (2002), and Steglich-Petersen (2006). The formulation I am employing here is largely taken from McHugh's (2013) formulation of (strong) exclusivity. Also, in a series of papers, Shah (2001; 2003; 2006; and Shah and Velleman, 2005) has investigated a phenomenon that he calls transparency (see Introduction for more). Transparency takes place in doxastic deliberation, namely when we ask ourselves what to believe. Initially, Shah (2003) presents the phenomenon of transparency as a *phenomenological* fact about how we doxastically deliberate. He points out that in doxastic deliberation the question *whether to believe that p* is transparent to the question *whether p is true*. However, in his papers Shah points also to a weaker, slightly different notion of transparency. He refers to it as the "psychological phenomenon of transparency" (2006: 483), i.e. the "psychological truth that we cannot believe for non-evidential reasons" (2006: 481). That means that, "only truth-regarding considerations move an agent in such [doxastic] deliberation" (2003: 468). This is the phenomenon I am focusing here. Finally, Parfit (2001) and others make the distinction between object/state reasons. Roughly, a reason for an attitude is object-given if it is a consideration that bears on the attitude's object, while it is state-given if it is a consideration that bears on having the attitude itself. On this view, considerations that bear on the attitude's object can motivate one to form or suspend an attitude in deliberation, where this is not true for state-given reasons.

At times, we come to form or drop a belief as a direct result of reasoning, i.e. as the last step of our reasoning process (Harman, 1986:2), and many seem to believe that in those cases non-epistemic considerations can't move us. More specifically, if the Strong Explanandum is true, all considerations that have nothing to do with *the truth of p* cannot motivate us in deliberation or reasoning about whether to believe that p. And that includes considerations regarding the possible epistemic benefit of having some false beliefs. First, it is, of course, possible that non-epistemic considerations (e.g. "it would be pragmatically beneficial to believe that ...") may cause the formation of a belief in non-deliberative contexts. What is special about the motivational role of epistemic considerations, though, is that in *deliberation* they can function as the *reasons* for which one forms a belief. It is notoriously hard to define "reasons for which". Let's say that R is a reason for which X ϕ s only if R is capable of *consciously* and *directly* motivating R to ϕ when X deliberates on whether to ϕ and uses R as a premise in such deliberation (as explained by Shah, 2006: 485). Thus, the view here is that pragmatic considerations can play a role in motivating you to form an *intention* to engage in all the necessary actions that could make it the case that you end up forming a belief; but in deliberation there is no direct motivational link between consciously entertaining a pragmatic consideration as a premise and the formation of a belief.

In addition, doxastic deliberation doesn't always result in forming an actual belief but can, at times, conclude in suspension of judgment (or withholding of belief). I take

¹⁴¹ Deliberation seems to be a fairly high-level form of reasoning, usually sprung by the explicit questions: "What should I believe?", "Is this right for me to accept?". I'd claim that we don't often deliberate about beliefs; we mostly simply reason with and toward our beliefs. In (non-deliberative) reasoning we don't pose any question, but go from belief to belief without explicitly attending to the relation between those attitudes and with usually little effort and self-reflection. I will offer my account of reasoning below.

withholding of belief (or suspension of judgment) that p to be the absence of the belief that p and the belief that not-p. Non-epistemic considerations, according to the formulation above, are not able to motivate us even to *withhold* beliefs as a result of deliberation. That means that, when you ask yourself whether to believe that p, and you reach the conclusion that you have strong evidence that p is true, non-epistemic considerations cannot make it the case that you suspend judgment about p.

Theorists usually refer to intuitions which, they claim, strongly indicate that any consideration that appears to us as having nothing to do with the truth of a given proposition 'p' is *incapable* of playing the role of motivating us to form/withhold the belief that p as a result of deliberation. It is vital to stress that it is not simply that non-epistemic considerations *don't* motivate us in doxastic deliberation. The point is that non-epistemic considerations *cannot* motivate us in doxastic deliberation. Here is an example often invoked. We generally believe that having optimistic expectations about one's abilities and the future in general can, for the most part, boost people's performance in various fields. An athlete, for instance, may realize that if she believes she will win the next game, she is more likely to be successful than if she holds realistic expectations. We usually agree that such practical considerations, however, are motivationally impotent with respect to her deliberating about whether to have that belief. Coming to the conclusion that it would be beneficial to have optimistic beliefs cannot motivate her to form those beliefs as a result of deliberation. Another classical example of this phenomenon is the Pascal's Wager. The Pascal's Wager has arguably demonstrated that it is prudentially advisable to believe in God. However, it seems that prudential considerations cannot, by themselves, directly motivate us to form a particular belief.

Indeed, noticing that it is prudentially advisable to believe that God exists will not directly motivate you to form the belief that God exists. At best, it is able to motivate you to form an intention to do what could bring about the belief that God exists. More specifically, it could motivate you to trick yourself into believing that God exists. And similar examples and intuitions are invoked to argue that suspension of judgments cannot be motivated by non-epistemic reasons.

Notwithstanding the fairly strong support these examples seem to have, some theorists find this analysis too strong. They claim that at times sufficient evidence does not compel us, and thus non-epistemic considerations are able to motivate us to refrain from forming a belief, even when one thinks one already has sufficient reasons for believing it. So, for instance, a judge on a death penalty case may be motivated to suspend judgment, even after reviewing evidence that he deems sufficient to judge the suspect to be guilty. The reason for doing so may be based on the consideration that other evidence may be coming in before he has to make the final decision, and so it is more prudent to keep an open mind vis-à-vis any new incoming proof. As a result, he suspends judgment on that.

Similarly, based on pragmatic considerations, we may decide to suspend judgment on the truth of some proposition *p* and – as a result – keep inquiring on whether *p* is true, even when faced with sufficient evidence for the truth of *p*. Thus one may use pragmatic considerations (e.g. if I believe my husband is guilty that will ruin my life) to motivate oneself to keep inquiring on some important issue (e.g. did he really commit that crime?) even when faced with sufficient reasons to believe.

What's more, some have worried that Exclusivity is too strong because it rules out cases like the following: "You are playing tennis. You are much stronger than your opponent, so you regard the possibility of defeat as far-fetched. If you believed that you would win, your confidence would make victory even more likely" (McHugh, 2012: 17). The suggestion here is that you can form a belief in reasoning based on the non-epistemic consideration that doing so will raise your chances of winning the match. In this case, it seems that non-epistemic considerations can influence your belief, qua reasons and in reasoning, because you already take yourself to have sufficient evidence for believing that you will win in the first place.¹⁴² Similarly, take two competing scientific theories which enjoy almost identical explanatory power and are well supported by the evidence. You have no clear epistemic reason to favor one over the other. However, whereas the first theory is very simple and elegant, the other one is extremely complicated and involves a good deal of cumbersome calculations. We might say that, in this case, we ground our belief in the first theory on non-epistemic reasons (e.g. avoiding complicated calculations). If right, this would show that pragmatic (non epistemic) considerations can motivate us in reasoning.

Anyway, the jury is still out on who is right on this issue, and I suspect opposing intuitions will keep clashing. Luckily, though, we don't need a verdict to continue our investigation. For simplicity, I will assume there is some *minimal explanatory level* any view offered has to meet to be considered feasible. We can call it 'Weak Explanandum' as it leaves open the possibility that pragmatic considerations may have some motivational force when it comes to suspending judgment. It also allows pragmatic

¹⁴² But see Archer (2015) for arguments against this possibility.

considerations to play a motivating role in forming the belief that *p* when one already has sufficient evidence that *p* is true:

Weak Explanandum: in reasoning considerations not concerning the truth of *p* cannot motivate, qua reasons, to directly believe that *p* if an agent takes herself¹⁴³ to lack sufficient evidence that *p* is true.¹⁴⁴ In contrast, in reasoning epistemic considerations concerning *p* can motivate, qua reasons, an agent to directly believe that *p* based on those considerations, unless she takes those to provide insufficient evidence that *p* is true.

The gist of this is that epistemic considerations (concerning the truth of proposition *p*) are sufficient to motivate an agent to believe as a result of reasoning, unless she has worries that those considerations do not provide enough evidence that *p* is in fact true. In contrast, in reasoning considerations that have nothing to do with the truth of *p* are themselves never sufficient to motivate one to believe that *p*.¹⁴⁵ Weak Explanandum is thus what we are after, the phenomenon we will try to make sense of in the rest of the chapter.

THE TELEOLOGICAL VIEW: BELIEVERS WANT THE TRUTH!

Here is a very popular way to explain the Explanandum: believers aim at having true beliefs and avoid false beliefs, and this is why they are predominantly motivated by epistemic considerations when reasoning. It is not that believers aim at truth in all cases of belief-formation; indeed, they don't aim at truth when unconscious processes produce

¹⁴³ Merely having or lacking evidence that *p* is false/true is insufficient to prevent pragmatic considerations from motivating. It seems that more is required, i.e. the recognition that one lacks that evidence.

¹⁴⁴ This is taken from McHugh (2013). Below I will explain what 'sufficient' stands for here.

¹⁴⁵ For possible counterexamples to even this weak version of the Explanandum see, for instance, James (1896), Sharadin (2016) and Worsnip (ms).

their beliefs. However, they cannot but aim at true beliefs when engaging in the practice of reasoning.

Indeed, the idea that believers want to have beliefs that are true has struck many as very intuitive to the point that this is often taken to be a truism. In addition, such a proposal seems to be theoretically quite fruitful: ascribing an aim to belief may help to explain the source and authority of epistemic norms. Aims (at least when obtainable) determine rules conducive to achieving them. If the aim of belief is truth, then epistemic norms are what will get us there. Since this aim is allegedly constitutive of belief, all beliefs and believers are subject to these norms as well.

Although intuitive, this view collides with another powerful intuition: sometimes believers think they would be better off believing something totally false. Indeed, if facing the choice between having a trivial false belief and saving humanity from destruction, I am quite sure it would be extremely callous on our part to choose the latter over the former! So we need more to explain why, contrary to our inclinations, we don't seem to have much of a choice when it comes to forming beliefs.

There are two strategies to address this. One is to say that believers want true and only true beliefs because truth is the sole constitutive normative standard for belief. When forming beliefs we are necessarily bound by this truth norm.¹⁴⁶ Note, however, that to explain the Explanandum this view requires subscribing to a strong form of motivational internalism in which accepting the truth norm means being necessarily motivated by it. And this seems quite a commitment to take, which has turned many away from this proposal (McHugh, 2013).

¹⁴⁶ See, for instance, Wedgwood (2002), Boghossian (2003), Shah (2003), Gibbard (2005), Engel (2013), Shah & Velleman (2005).

The second, more promising view is that by *nature* belief is truth-directed/regulated, and that constrains believers' goals in forming beliefs as well. This is the *teleological view* of belief.

As already mentioned in chapter 2, this view has David Velleman (2000) among its strongest proponents. Velleman's view starts off by noticing that belief is part of a large group of attitudes – called acceptances (i.e. belief, supposition, imagining) – that all present their content as true. However, they do it for different reasons. Belief regards its content as true with the aim of it really being true. In contrast, imagining presents its content as true with the aim of, e.g., having fun. And supposing presents its content as true with the goal of, e.g., showing that it is false.

So what sets belief apart is that it aims at truth, which, for Velleman, means that belief is a *truth-regulated acceptance*: believing that p means accepting that p is true, and such a state is either produced by an intention to accept only the truth of p or it is *de facto* regulated by a cognitive mechanism with the function of tracking the truth.

As a theory of the nature of belief, we should expect that the teleological account will be able to make sense of the Explanandum (either weak or strong). Unfortunately, this is where Velleman's proposal fails. Here is why. There are two possible readings of belief's truth-directedness: one is that for an acceptance to count as a belief it needs to be entirely/mostly truth-regulated (strong reading). Alternatively, we may say that for an acceptance to count as a belief it needs to be only *minimally* truth-regulated: it is possible that non-epistemic factors may predominantly influence the formation of a belief (weak reading). This, however, raises the following dilemma (taken from Shah, 2003). If we adopt the strong reading, we have a nice way to make sense of the Explanandum. What

we desire to believe, for instance, simply cannot (strongly) influence our beliefs, and this explains why in reasoning non-epistemic considerations cannot be – at least not predominantly – reasons for which we form a belief (assuming that ‘reasons for which’ are causes). Unfortunately, however, if we embrace this strong view, we end up with the very unintuitive proposal that our attitudes are never subliminally influenced by our desires and emotions. But this is too strong: we know that some of people’s beliefs are predominantly influenced by non-epistemic factors, but we still want to call them beliefs. If, on the other hand, we adopt the weak reading, then we allow for the possibility that our beliefs may be predominantly influenced by non-epistemic factors; but then we lose sight of the Explanandum. Although our beliefs are at times strongly influenced by our desires, in reasoning these cannot motivate us, qua reasons, to form a belief for which we do not have sufficient evidence. They cannot be predominant, at least. The (weak) aim of belief does not seem able to explain why this is the case.

But there is hope for the teleologist: Steglich-Petersen (2006) has recently offered a way to diffuse this objection. The teleologist’s dilemma is avoidable, he says. Steglich-Petersen is in agreement with Velleman that belief is an acceptance that aims at truth. First, belief **aims₁** [subpersonally] at truth in the sense that if belief is at least weakly truth-regulated, it is influenced, formed, caused by some truth-tracking mechanism. Second, belief **aims₂** at truth in the sense that, in forming a belief, the believer has the goal to form an acceptance that is *strongly* regulated for truth. This second sense of belief’s truth-directedness is what can explain the Explanandum.

But why should we think that belief aims₂ at truth in that way? Let me explain the problem faced by Steglich-Petersen. For a belief, to aim₁ at truth it is enough that it is

weakly regulated for truth. So to form an attitude that aims₁ at truth it is enough to aim at forming an attitude that is weakly regulated for truth. But to be weakly regulated for truth it is enough that one's belief is minimally influenced by some truth-tracking mechanism. But if to form a belief through reasoning, one can simply aim at forming a belief *weakly* regulated for truth, it is possible that one is motivated by non-epistemic considerations to form a belief for which one has insufficient evidence.

How can Steglich-Petersen solve this? To explain why belief aims₂ at truth Steglich-Petersen invokes the "common-sense concept of belief" which, on his view, is different from the empirical concept of belief as only the latter is open to the idea that belief is weakly regulated for truth. As he (2006: 511) explains this,

[A]lthough there is a sense in which all it takes for something to be a belief is that it involves weak truth-regulation, the common-sense concept of belief involved when intending to form a belief [...] involves strong truth-regulation simply because in that context, what one normally means by 'intending to form a belief' is intending to accept a proposition if and only if that proposition is true.

That is: In any intentional or personal-level belief-formation process – which by and large encompasses doxastic deliberation and reasoning – we apply the concept of belief and this contains the idea of strong truth-regulation and making sure one's belief is true.

This point becomes clearer once we turn our attention to Steglich-Petersen's parallel between the concepts "believing" and "concealing". He points out that these concepts both refer to aimed-directed activities: belief's point is to get to truth and concealing aims at making sure something is not found. Now, Steglich-Petersen maintains that, on a closer look, it appears that there exist two slightly different concepts

of “concealing”. Whereas in our intention to conceal some chocolate we are guided by the idea that our action will count as concealing only if it is carried out very well and it is likely to succeed, in reality “not all the instances of what we would be willing to call ‘concealings’ are particularly well carried out” (Steglich-Petersen. 2006: 512). If I *intend* to hide some chocolate, put it in the cupboard and then unwillingly leave the doors of the cupboard open, my action seems to count as an instance of ‘concealing the chocolate’, regardless of the fact that it clearly fails at that. That is, Steglich-Petersen’s point is that it seems that when we describe the act of intentionally concealing something, we use some strong notion of what counts as concealment; when we describe an act of concealment per se, we are less strict.

Steglich-Petersen takes this true of belief as well, but the parallel fails. Here is why. Steglich-Petersen (2006: 512) says, “all it takes for a bit of behaviour to count as ‘concealment of x from Y’ is that it is weakly regulated for ensuring that Y will not find x.” But is that always so? Again imagine I move the chocolate from the kitchen table to the cupboard, but then leave the doors of the cupboard open; absent any intention to hide the chocolate in the first place, does my action really count as concealment? I doubt it: unless one is trying to hide the chocolate and fails, weak concealment is no concealment; it is just moving the chocolate from one place to another. That means that an action is an instance of concealing only if it has some strong regulation toward fulfilling the aim of concealing, namely a regulation either coming from my *intention* to hide the chocolate or from an action that in fact results in hiding the chocolate. So the intention to conceal must be an intention to *successfully* conceal because, if the concealing is poorly conducted, it is the intention itself that makes my action into an instance of concealment.

That is not true of belief. Recall, the teleologist wants to say that a weakly regulated belief is a belief nonetheless, even absent any intention to make it a successful belief. That means: belief can aim at truth in a weak sense, whereas concealing aims at hiding only in a strong sense. Here is the problem for the teleologist: when we intend to form a belief, we aim at bringing a new belief into existence. If the only thing required for belief is weak truth-regulation, I should be able to intentionally bring about a belief that is only weakly truth-regulated. Contrary to what Steglich-Petersen claims, the analogy with concealing is unhelpful because, in the case of concealing, if I intend to bring about a weakly regulated concealment I am not concealing at all, because there is no such thing.

Steglich-Petersen's strategy was to claim that, as for the concept of concealment, the concept of belief just entails that beliefs strongly aim at truth. But as I showed, whereas there is an explanation for why this is so in the case of concealment (i.e. there is no such thing as weakly regulated concealment), we are left with no explanation for why this happens with belief (given that belief can be weakly regulated). The teleologist can't explain the Explanandum, and so we need a better view.

REASONING'S CONSTITUTIVE AIM

Since the teleologist has no clear way to account for the Explanandum, the teleologist's claim that something in the nature of belief determines its special relation with truth seems to lose ground. The view I am about to propose is in clear opposition to the teleologist's narrative about truth and belief. In chapter 2, I defended a definition of belief independent of the claim that belief aims at truth. I will briefly rehearse that

argument here, and then argue that the Explanandum is the result of the goal we have when forming *beliefs* through reasoning. This aim is the result of the more general aim we have when reasoning and of the unique role belief has in reasoning. The reader should keep in mind, though, that nothing I am about to say is meant to inform the debate on whether there is a truth norm of belief or not. That normative issue is simply beyond the scope of this chapter.

Let me start by rehearsing the view of reasoning presented in chapter 3 and 4. Reasoning, I claim, is a special kind of inference. It is an inference guided by a constitutive aim or goal.¹⁴⁷ More specifically, *reasoning is the inferential process guided by the aim of arriving at a conclusion supported by at least sufficient reasons.*

The conclusion of *all* reasoning consists in dropping, reaffirming, forming or suspending (avoiding forming) an attitude. Doing one of those things in a way that is supported by sufficient reasons is the constitutive and fundamental aim of reasoning. Insofar as the reader accepts the idea that correctly responding to reasons is tantamount to being rational, my claim here is that reasoning aims at rationality, or at least at *this* sense of rationality.¹⁴⁸ On my view, the aim of reasoning is an aim *of the agent*, a goal the agent necessarily has when reasoning. An aim is a desire to obtain a certain result (Sosa, 2010). The aim one has is mostly shown in the *dispositions* one has. In this sense, having an aim or goal means being motivated by certain considerations and not others, and disposed to prevent certain occurrences to happen, and change things around if the aim is not

¹⁴⁷ For other views of reasoning see Broome, 2013; Boghossian, 2014; McHugh and Way (2016); Neta, 2013; Pettit, 1993; Wedgewood, 2006.

¹⁴⁸ We are rational, Joseph Raz has suggested, insofar as we exercise the capacities that enable us, in general, to respond correctly to reason. So when reasoning, I claim, we aim at fulfilling that capacity, at being rational in this sense. Also, it is common in epistemology to use ‘rational’ and ‘justified’ interchangeably.

satisfied. Also, the aim of reasoning does not have to be necessarily consciously entertained by the reasoner; however, a normally competent reasoner will not explicitly (and sincerely) deny that this is her aim when reasoning. Finally, note that it is possible that there are multiple aims in reasoning. However, the constitutive aim is the one aim that is always present.

Normative reasons are things that count in favor of (or against)¹⁴⁹ a certain response. By “response”, I mean the holding, withholding or dropping of an attitude. Sufficient normative reasons are considerations that ‘sufficiently’ count in favor of a given response given all the information one possesses and that are available to her. (I will assume, somewhat vaguely, that a reason is available when it is a true proposition or fact perceptually within my reach and/or easily inferable given what I know). And, an agent S has sufficient reason to ϕ only when S has a reason to ϕ , and no stronger reason not to ϕ ; but I will have more to say about sufficient reasons later.

As already pointed out at length above, there is a lively discussion in philosophy on the nature of normative reasons. In particular, people debate on whether or not reasons are factive (facts, true propositions, true mental states). Although many find the factive view to be quite reasonable, some have questioned it arguing that rationality consists in responding correctly to one’s subjective reasons (e.g. Parfit 2001, 2011; Schroeder, 2008; Way, 2009). That is, some theorists maintain that there are two reason-relations, i.e. the objective reason relation and the subjective reason relation. An objective reason is a reason there is *for* someone to do or believe something; a subjective reason is a reason

¹⁴⁹ I assume that reasons against A are reasons in favour of not-A. However, in the case of attitudes this may be a little more complicated. So a reason against the belief that p, may be a reason in favor either suspending judgment or believing that not-p. I will leave this issue aside here as it does not have any important upshot for my argument.

someone *has*. Objective reasons are factive. Subjective reasons are mental states and may be non-factive. As mentioned previously, when it comes to my view of reasoning, though, it would be odd to say that, when reasoning, we aim at reaching a conclusion supported by reasons regardless of their truth. Subjective normative reasons are reasons that that would be objective reasons were they true. Thus the notion of subjective reason itself depends on the idea of objective reasons which is, if you will, the normative relation (whereas subjective reasons are derivative in this sense). Hence, it seems reasonable to claim that, if reasoning aims at rationality, then the notion of reason at play must be factive. That is, we don't want to reach a conclusion that is responsive to all the reasons there are in favor of it, as many of those reasons may be beyond our reach. At the same time, it seems natural to say that, for our subjective reasons, we want our conclusion to be responsive only to those that are factive. I will thus adopt this view in what follows, and assume that normative reasons must be factive.

Why think reasoning is guided by the aim of reaching conclusions supported by sufficient normative reasons? In reasoning, we don't just wonder from one premise to another while ending up with some conclusion in mind at the end. In reasoning, we have a focus, something we want to reach. And the claim that reasoning has a constitutive aim makes sense of the fact that, intuitively, reasoning is an activity (Broome, 2013). I made this point related to gardening before: there is no gardening if you stop aiming at making your plants grow. You can decide to stop watering or trimming them, and prefer to let them grow as they want. However, the moment you do that, you are not gardening your plants anymore. Hence, the activity of gardening itself has a constitutive aim. Similarly, reasoning has a constitutive aim, a goal that cannot be undermined by any other aim one

may have, if one still counts as reasoning. Its aim-directedness thus puts reasoning very close to other activities and makes sense of the idea that reasoning is itself an activity.¹⁵⁰

In reasoning we want a conclusion that is well supported. Boghossian (2014:5) presented roughly the same idea in the following way: “no causal process counts as [reasoning], unless it consists in an attempt to arrive at a belief by figuring out what, in some suitably broad sense, is supported by other things one believes. In the relevant sense, reasoning is [...] something that we do with an aim—that of figuring out what follows or is supported by other things one believes.” Indeed, if an agent were to reason from premise P to conclusion C while also saying that P offers no (sufficient) support to C, we would doubt this agent is reasoning at all. That is, denying that our premises offer support for the conclusion we just reached through reasoning undermines the very process of reasoning.

Finally, the idea of reasoning being guided by an aim seems helpful to make sense of the claim that reasoning is a kind of thought process in which the conclusion we arrive at is a conclusion we *endorse*. This will be key in this chapter: the attitudes we endorse are the attitudes we are ready to use in further reasoning. In contrast, the attitudes we have but do not take responsibility for are the attitudes we will likely refrain from using in any future deliberation. On a first pass, the idea that reasoning has an aim – any aim – seems to be able to make sense of all that: the attitudes we reach through reasoning are attitudes that (seem to us to) fulfill the aim we have when reasoning; they fulfill our goal and thus we endorse them. Although quite intuitive, this proposal raises the worry that not all the goals we have are goals we *want* to have (or that we endorse). I claim, in

¹⁵⁰ That does not mean that having an aim X requires *intending* to do X; it is enough to have a certain disposition to do something if it turns out I am failing to achieve X (Alvarez 2010; Whiting 2013).

contrast, that the aim of reasoning is not just any aim: it is the aim of reaching a rational conclusion, i.e. a conclusion supported by (at least) sufficient reasons. Now, taking something to be rational seems tantamount to endorsing it. Rationality is what is intelligible and makes sense to us,¹⁵¹ thus reaching what we take to be a rational conclusion is reaching a conclusion that we can stand behind, identify with and take responsibility for. The reason why reasoning forms attitudes we endorse is because reasoning *is* the process *aimed* at forming attitudes supported by reasons - so arriving at the conclusion of reasoning means endorsing it *because* one necessarily sees it as rational/supported by reasons. That does not mean one needs to believe that the reached conclusion is rational: it is the fact that we reasoned to it that expresses our endorsement.¹⁵² As a result, the aim of reaching a rational (i.e. sufficiently supported by reasons) conclusion easily explains why in reasoning we reach attitudes we necessarily endorse.

BELIEFS AS DEFAULT PREMISES

It is often considered a truism in philosophy that believing that *p* means being disposed to use '*p*' as a premise in one's reasoning. For instance, Williamson claims that "one believes *p* outright when one is prepared to use *p* as a premise in practical reasoning." (2000: 99). Accordingly, "[...] if a person consistently refused to rely on *p* as a premise, and rejected arguments relying on it, then it would be plausible to say that he did not really believe that *p*" (Scanlon, 2007: 91). Indeed, it seems that if one is not

¹⁵¹ Raz (2005: 8–9): "One aspect of reasons for actions is that they make choices, intentions, and actions intelligible both to the agent and to others".

¹⁵² Thus, my proposal explains why reasoning to an attitude is tantamount to endorsing it *without* any need for meta-cognitive or second-order beliefs (or cognitive attitudes in general) to take place in the reasoning process.

prepared to use the content of one's belief as premise in one's reasoning and to be guided and motivated by it, one does not really believe it. If you believe that smoking causes cancer but then fail to apply this proposition in any of your reasonings (e.g. smoking causes cancer so I shouldn't smoke), then there are good chances that you don't really believe it.

At the same time, though, that does not mean that what I believe *de facto* influences *all* my reasoning. Indeed, it is compatible with this view of belief that, at times, when reasoning, we don't let ourselves be guided by what we believe. Let me mention some relevant cases (see also chapter 1 and 2).

As in chapter 1, take Scott, a surgeon at the local hospital. He is ready to perform a kidney surgery on one of his patients. Scott has high credence that the kidney that needs to be removed is the left one. He also fully believes that, although he is not certain. Given the stakes, Scott is quick to recognize that, were he to be mistaken, the consequences would be terrible. So he decides to stop by his office and check his patient's dossier before performing the surgery. That is, when stakes are very high, we do not reason with our beliefs as premises – we use other attitudes. But belief is still our *default* premise in reasoning, namely it is what we would *normally* use to reason (Ross and Schroeder, 2014). Similarly, it is hard to deny that at times we reason with assumptions and hypothesis. These, however, are attitudes adopted as premises usually only in limited contexts. That means, believing equals being disposed to use a proposition as *default* premise, i.e. as a premise we use in day-to-day situations. Other attitudes may enter our reasoning, but only in a limited way and are at the fringes of our normal reasoning occurrences.

To be clear, that does not mean that in normal contexts we *always* reason based on what we believe, though. Take this: I believe my boss is an incompetent man, but I refrain from using this in my day-to-day reasoning, as I fear the consequences this may have. I *do* believe it, but I am wise enough to voluntarily limit its influence. The idea here is that I am disposed to reason with that belief, but I decide not to do so. A similar phenomenon takes place when beliefs get compartmentalized because of some failure of our cognitive systems. Here is an example discussed above and offered by Lewis (1982). A resident of Princeton has the following three inconsistent beliefs: she believes that Nassau St. runs North-South; he believes that the railroad tracks run East-West; and finally she believes that Nassau Street and the railroad tracks run parallel to one another. What's going on here is, according to Lewis, that the agent can harbor contradictory beliefs: one is that the street and the tracks run North-South, and the other is that the street and the tracks run East-West. How is that possible? The idea is that the agent fails to bring those beliefs to bear, she simply does not see that they are in contradiction because she does not activate them at the same time, in the same contexts. Her beliefs are, that is, compartmentalized (Stalnaker, 1984; Egan, 2008). As a result they both get to be activated and used in reasoning, albeit in different situations. It is important to notice that compartmentalization is *not* a feature of belief itself (as it is for attitudes such as assumptions and hypothesis), but the result of an abnormal situation that limits the range of influence of a particular belief. Absent or lifted these external influences, belief gets back to its job of being a default premise.¹⁵³

¹⁵³ I argued at length for this view of belief in the second chapter of the dissertation. There the reader will find a much more nuanced analysis of the role of belief in reasoning.

Given the role belief has in reasoning, it is to be expected that belief has a part to play in reaching reasoning's constitutive goal. I argued that reasoning aims at arriving at a conclusion supported by normative reasons, so it seems natural to see belief as providing those reasons. Since belief is a default disposition to use a proposition as premise, then it means that it is a default disposition to use a proposition *p* as a *reason* in reasoning.¹⁵⁴ To illustrate, since I believed that Trump would be an incompetent President, I was disposed to use this a reason not to vote for him in the 2016 Presidential Elections. And now that he has been elected President of the United States, my belief about his abilities is used as a reason for me to believe that the economy is probably not going to substantially improve over the next years. And so on. So now let's see what that means for the Explanandum.

BELIEVING THROUGH REASONING

At last, here is my explanation for the Explanandum. Above I sketched an account of reasoning as a thought process guided by an aim. This is a constitutive aim, in the sense that it is necessarily present in all instances of reasoning. But it is also a fundamental or basic aim, in the sense that it should be able to explain other key goals we may have when reasoning. Thus, it should not come as a surprise if the aim of reasoning is able to account for the specific, constitutive goal we have when reasoning toward beliefs (vs. the goal we have when we reason toward different attitudes).

¹⁵⁴ A similar view has been expressed by Scanlon who maintains that believing that *p* “involves recognizing [*p*] as having the status of something that is to be relied on in further theoretical reasoning by providing reasons for accepting what it entails, and to be relied on as a premise in practical reasoning” (Scanlon, 2007: 104).

Now, the aim of reasoning is in general to have one's conclusion sufficiently supported by (factive) normative reasons. So a good premise is a premise that is (at least) *true* because it is a premise that does what it is supposed to do: it supports the conclusion of reasoning. Imagine then a situation in which I want to form a belief that I can use in future reasoning. Clearly, I am going to strive to form a good premise. Since the goal of reasoning is reaching a conclusion supported by reasons, I will not treat a proposition as a reason or adopt it as premise if I doubt its truth. That would go against the aim of reasoning.

At the same time, though, when we strive to form a good premise, a premise that can be a reason, we can't literally strive for truth, as that would require that we aimed at being *certain* of each proposition we treat as a reason. But we can't aim at being certain that our premises are true, otherwise we would have basically no premises, given that probability 1 is really hard to obtain. In fact, when we treat a proposition as a reason we are most of the time not *certain* that it is a reason and thus true. We use it *as if* we were certain of it¹⁵⁵. We said belief is a *default* disposition to use a proposition as a reason in reasoning. It is what we use as a reason in normal cases. That means that, since I want to use the believed proposition p as default premise, I want that p is supported enough to give me some reliability when treating it as true in reasoning. So when forming a belief that we want to use as default in further reasoning, we need to strive to form a belief that is – by our own lights – supported by sufficient reasons for its truth. To be used by default, a premise must be a proposition that appears/ is not doubted to be supported by sufficient *epistemic* reasons, reasons that by and large can guarantee its truth.

¹⁵⁵ As pointed out by Ross and Schroeder, this seems to be a heuristic method of reasoning that requires a degree of approximation. I also argued for this claim in the first chapter.

How much is sufficient evidence for a proposition to be treated as a default premise? A proposition can be treated as a default premise in reasoning only if an agent can feel epistemically comfortable relying on in further reasoning in normal contexts, i.e. context in which the stakes are not too high. A plausible view is that sufficient evidence for an agent's belief that *p* is evidence that is strong enough to rule out that the belief that *p* is easily false given the evidence she has, but not so strong that it requires that she is actually certain of it. And, plausibly, 'easily' here is the degree of support in the proposition *p* that rules out we have knowledge that *p* is true.¹⁵⁶ So let me make the reasonable stipulation that I have sufficient evidence in a believed proposition for it to be used as a default premise in reasoning if and only if I would also count as knowing it, if that proposition were true. And additional reason for this assumption is that there is now considerable support for the claim that one appropriately treats a proposition as premise if and only if one knows it (Hawthorne and Stanley 2008). Following the host of intuitions on which this view is based, it seems that people deem that the necessary support to appropriately use a proposition as a premise in reasoning is whatever support that would ensure knowledge. At least, this seems to be true for the normal cases, i.e. situations in which the stakes are not too high. The result of this is the plausible claim that, when forming a belief that we want to use by default in further reasoning, we need to strive to form a belief that is supported enough to count as knowledge, if true.

When forming a belief that we want to use in further reasoning, we need to strive to form a belief that is supported enough to count as knowledge if true. That's my claim so far. But why should we necessarily care about using a belief in future reasoning?

¹⁵⁶ This view is supported, among others, by Dretske (1971), Pritchard (2005), Sosa (1999), and Williamson (2000).

Why can't I form a belief through reasoning with disregard for its truth? I mean, it seems I should be able to do that, at least while promising to myself that I will limit its influence.

I believe the answer is in the phenomenon of reasoning itself. When I reason toward an attitude I must aim at forming an attitude I *subscribe* to. When one engages in reasoning one is trying to arrive at a conclusion one endorses, a reasonable conclusion by one's light. That was one of the elements of reasoning I mentioned above.

Naturally, a belief I endorse is also a default premise I endorse, i.e. a default premise I see as legitimate. How could I endorse a belief, but then not rely on it? Presumably, if I see belief B as an expression of my rational agency, it seems that I can't, at the same time, refuse to use belief B as a guide in my deliberation.

As a result, one cannot both engage in reasoning *and* try to form a belief that one believes one would not be able to use as a default premise. Aiming at forming a belief with disregard for its sufficient epistemic support means trying to form a belief that we endorse, but that we know we would not be able to use as a legitimate default premise. But this is contradictory. So we can't have that aim. The aim of having an epistemically supported belief is a constitutive aim of reasoning, and it is an aim that trumps any other aim we have.

And this explains the weak Explanandum: when reasoning toward a belief we aim at forming a belief at least sufficiently supported by epistemic reasons. As a result, we can be motivated to form a belief that p by evidential considerations pertaining to the truth of p (unless we suspect that these considerations are not sufficient reasons that p is true). In addition, non-epistemic considerations cannot motivate us to form a belief, if we

suspect that we don't have enough evidence that *p* is true. This is why non-epistemic considerations have limited direct impact on the formation of beliefs; in contrast, epistemic considerations are enough to motivate us to directly form a belief, unless deemed to be insufficient to guarantee that that belief can be used as a default premise in reasoning.¹⁵⁷

In conclusion, since we can't but want to have true premises and thus true beliefs when reasoning *in general*, when trying to form a default premise, i.e. a belief, we can't but try to form a belief for which we can't be easily mistaken. Aiming at having *epistemically* sufficiently supported belief is the necessary aim we have when reasoning towards beliefs. The source of that aim is reasoning's fundamental goal of rationality, the role belief plays in it and the fact that, when we try to form a belief in reasoning, we can't but aim at forming a belief we *can* use as a default premise in future reasoning.

OBJECTIONS

OBJECTION The analysis above aims to explain why we can't form a belief disregarding its epistemic support. However, such a proposal does not seem applicable to suspension of judgment. More specifically, it seems that, in most cases at least, epistemic considerations dominate deliberation that ends with suspending judgment as well. To illustrate: in most cases, even if we believe that it would be for us beneficial to not have a certain belief (e.g. I have high probabilities to lose the next match, I have an incurable illness), it seems that it is almost impossible to come to suspend judgment if we

¹⁵⁷ The fact that I necessarily do care about having beliefs that are sufficiently justified does not close, for me, the question of whether I *ought* to care. So nothing I am saying here bears on the normative question about how we should form beliefs.

also have strong evidence that that belief is true. It is unclear how the view proposed above can explain that, since suspension of judgment does not have a role in reasoning.

To address this objection first let me point out that the aim of reasoning applies even when, in reasoning, we end up suspending an attitude or withholding it. It is true that talking about sufficient reasons may sound problematic when referred to suspension of judgment and withholding of attitudes. The idea, though, is that when reasoning we want our conclusion to *correctly respond* to the reasons we have. If intended broadly enough, that means we can respond to the lack of sufficient reasons for an attitude by dropping it. So, for instance, if I lack sufficient reasons to hold a belief that p, while lacking sufficient reasons to disbelief that p, I respond correctly to my (lacking of) sufficient reason only if I suspend judgment on p. Similarly, if I drop the belief that p even though I have enough evidence to support it, in most cases it means that I miss the goal of reasoning by issuing an unjustified or irrational conclusion.

In addition, the reason why we can't just decide to suspend judgment whenever we want is also related to belief's role in reasoning. If one has sufficient epistemic reasons to believe that p, suspending judgment may compromise future reasoning. Attitudes such as belief are systematically linked to a host of other attitudes. Holding or withholding any such attitude is bound to have effects on the attitudes we have and use in our reasoning. So deciding to withhold a belief that is likely to be true will have an effect on the other attitudes and beliefs we have or might have. The risk of compromising future reasoning is thus very high, and this explains why – even in accordance with Weak Explanandum – the direct influence of pragmatic considerations over suspension of judgment is also severely limited.

OBJECTION: *The view proposed here is that in reasoning we are motivated by the goal of forming a belief only if epistemically supported. This idea seems to fall short of the intuitive claim that we want true beliefs (assuming that sufficiently supported beliefs may be false). If we discover that we reasoned toward a false belief, we usually deem this to be a failure even if the belief was itself sufficiently supported.*

It really does seem that we regard false but sufficiently justified beliefs as a failure or a missed target. And my view does explain why false beliefs are problematic, and why reasoning toward a false belief is a failure. Beliefs are our default dispositions to use propositions as our premises. This means that, since the goal of reasoning is to have conclusions supported by true premises, a false belief fails in what it is supposed to do qua premise. And reasoning toward a belief that turns out to be false is a failure insofar as the goal of belief-formation through reasoning is to form premises for future reasoning. If we use a false belief as premise in reasoning, there is a failing in our reasoning.¹⁵⁸ The practice of reasoning itself – as I depicted – entails that reasoning toward a false belief is a misstep. Hence, what distinguishes my position from the widely-popular, alethic-teleological view is that on my view the value of truth is not intrinsic to the attitude of belief but emerges out of the practice of reasoning.

But why not saying that truth is the goal we have in reasoning towards belief? That has struck many as a plausible position. I find this proposal problematic and generally less convincing than the view I am putting forward. On one version of such a

¹⁵⁸ In a recent paper, Huw Price (2003) tackles a somewhat similar issue and offers an analogous conclusion. He calls truth a “convenient friction” and argues that we take ourselves to be bound by a norm of truth not because there is necessarily such a norm, but because this is a necessary condition for human discourse. In conversational practice “speakers take themselves and their fellows to be governed by a norm stronger than that of justification.” (2003: 168)

proposal, the goal reasoners have is to form a belief *only* if that belief is true. Reasoners could satisfy the goal, though, of only truth by simply avoiding any belief. That is, suspending judgment on any subject matter would just do for them. But of course this is not what reasoners do, and the goal of reasoning is hardly met by simply avoiding forming beliefs all together. Relatedly, saying that we aim only at truth in theoretical reasoning is either too strong or uninformative. If we are *really* after only truth, excluding any other goal, why do we accept to believe anything short of what has subjective probability 1? That would be in fact the best way to achieve the goal to avoid false beliefs. But if certainty were the goal here, we would form basically no belief via reasoning, given that probability 1 is really hard to get. So saying that believers are motivated by the goal of having only true beliefs is ascribing to believers an almost impossible aim: the aim of certainty. Alternatively, what we mean is that we want the truth but only to some degree. But how much do we want that? That becomes the real question.

A final, more promising, move is to say that, for some proposition p they are considering in reasoning, reasoners aim at believing p if and only if p is true. That would deliver the Explanandum, while avoiding the pitfalls of the first formulation of the truth-aim. Indeed we now seem able to explain why we want beliefs that are epistemically justified but not necessarily certain: we want to avoid false beliefs, but also make sure that, if the proposition we are considering is likely to be true, we do end up believing it.

Still, I worry that this formulation of the aim of truth is too strong. I worry about cases in which, unbeknownst to us, we are precluded from reaching the truth. In reasoning we can only operate with the reasons available to us. In some instances, our

prospective may be limited and the reasons we have access to, even when sufficient, may constitute only a small subset of the reasons we could potentially get access to were we in a better epistemic position. In those cases, truth may be outside our reach as there may be pieces of evidence we will never be able to obtain. And it seems intuitive to say that we can't aim at having a goal that is unobtainable. Based on the truth aim, in those cases we would not be able to form a belief through reasoning – which seems quite unintuitive. As a result, the aim of truth seems too strong to capture that kind of reasoning.¹⁵⁹

In addition, there is not clear explanation for why in reasoning we would aim at truth and only at truth. This formulation of the aim of truth completely rules out cases in which, based on pragmatic considerations, we may decide to avoid forming a belief for which we have enough epistemic support. As I mentioned above, given the aim of reasoning, we know why the direct influence of pragmatic considerations over suspension of judgment is also severely limited. But from what I said, we have no reasons to believe that pragmatic considerations have no influence over suspension of judgment. So the view that we aim at truth and only at truth cannot be vindicated within the theoretical frame I argued for in this chapter. This is why I reject it.¹⁶⁰

¹⁵⁹ There are also cases in which truth is not our goal, and overtly so. In particular, scientific reasoning hardly aims at truth. It is widely held, even by scientific realists, that even our best scientific theories are likely false. Although over time our inquiry may be approximating truth, as they are working through the construction of a scientific theory, the goal of truth is likely not what scientists are aiming for. Empirical adequacy may just do it for them. However, there are good reasons to suspect that the attitude scientists have and form through reasoning may not often be belief, but a form of acceptance. Similarly, van Fraassen (1980) has proposed a distinction between acceptance of an empirical theory and believing it to be true, arguing that evidence of a theory's empirical adequacy justifies one only accepting rather than believing.

¹⁶⁰ Whiting (2014) also uses the aim of (pragmatic) reasoning and the role belief has in reasoning to make sense of the 'aim of belief' which, however, he sees as the thesis that subjects are committed to evidentialism (whereas I make no such claim here). His argument is that we aim to act only on the basis of practical reasons, and thus when we aim to form an action through reasoning we must aim to form it as based on reasons. Belief plays the role in premise in reasoning and so belief must be true. This proposal, however, leaves some issues unresolved. First, Whiting says that the aim of belief is to believe that p only

OBJECTION *There is now considerable support for the claim that one appropriately treats a proposition as premise only if one knows it (Hawthorne and Stanley 2008). As a result, one would expect that, when forming a belief that p through reasoning one aims at putting oneself in a position to correctly treat p as a reason. That also means that when forming the belief that p through reasoning, since we are committed to using it as a premise for future reasoning, we must necessarily aim at actually knowing that p. That is, when it comes to formation of belief, reasoning aims at forming a belief only if one knows its content.*

My proposal is compatible with the intuition that something goes wrong if I treat a proposition as a reason in reasoning but fail to have knowledge of its truth. On my view, this is a combination of two things. First, to correctly treat p as a reason we need enough guarantee that p is true. That means that I have sufficient evidence in a believed proposition to use it as a reason only in case that, if that proposition is true, I would also count as knowing it. Second, it follows from my account of reasoning that a piece of reasoning is good when we get to a conclusion supported by reasons. Thus, if in reasoning we treat p as a reason, our reasoning fails if p turns out to be false. So if I have enough evidence that p is true, but p is still false, what is wrong is that my use of p as a

if p is a reason (and thus true). But this is compatible with the fact that (1) they may be completely motivated by pragmatic considerations for suspension of judgment, and (2) they may be motivated by pragmatic considerations along with epistemic considerations in forming a belief. However, both (1) and (2) contradict what Whiting is seeking to explain, i.e. that agents are committed to evidence being the only reason for believing. Second, his account of reasoning is that in reasoning one aims to be *guided* by the facts and, more specifically, by reasons. In addition, the aim of belief is to provide premises for reasoning. As a result, he claims, the aim of belief is to be a reason. But there are certainly premises in reasoning that do not provide guidance: reduction ad absurdum is an example of reasoning where the premises do not provide guidance. Thus, claim that in reasoning we aim at being guided by facts strikes me as false. My claim is, in contrast, that in reasoning we aim at conclusions supported by facts. Finally, as I explained above, in forming a belief we can't simply aim only at truth.

reason failed to deliver a conclusion supported by actual reasons. Putting these two things together we see some support for the idea that knowledge is required for using a premise in reasoning.

The view I favor is that in reasoning we aim at forming a belief that we can use as premise in future reasoning. Whereas, of course, we prefer to have premises we can *appropriately* use in future reasoning, arguing for the aim of knowledge incurs in some of the problems I raised above for the aim of truth. More specifically, the claim that we aim only at forming beliefs that would amount to knowledge raises the worry that forming no beliefs whatsoever may be enough to fulfill that aim. And adding that we aim at forming a belief *if and only if* that belief would constitute knowledge raises concerns about cases in which we are precluded from reaching the truth and so we can't aim at knowledge. For those reasons, I believe it is better to say that our aim in reasoning is to form a conclusion with sufficient support. If that conclusion is a belief, then we aim to form it only if it is sufficiently epistemically supported to guarantee that we can use it as premise in further reasoning.

CONCLUSION

More often than not, having beliefs that are true is definitively a good thing. Nobody will deny that. However, this is a far cry from saying that belief, by its very nature, aims at truth. In fact, in this dissertation, I raised worries for the almost universally accepted view that there is a constitutive, normative and/or teleological, relation between belief and truth, and that epistemic normativity follows from that. I did argue that belief has a necessary aim, but I am not sure it counts as ‘constitutive’ of belief. I did say that truth is important for belief, but I also added that truth is not what believers necessarily aim at when deliberating about what to believe. I did not exclude that norms may govern beliefs, and necessarily so. But I also offered reasons to be skeptical of the idea that these are part of the ‘essence’ of belief which, I claim, is really about the role belief has in reasoning. I tackled all this while trying to explain why we have limited power over our beliefs so that we can’t believe what we want. First, I defended the claim that believing means being disposed to use a proposition as a default premise in reasoning. Second, I focused on reasoning and defined reasoning as the inferential process guided by the constitutive aim of coming to a conclusion supported by sufficient normative reasons. Because it plays the role of a default premise in reasoning, belief is also generally expected to be a normative reason. Arguably, normative reasons are either true or at least epistemically supported propositions/attitudes. Hence, as a result of the constitutive aim we have when reasoning in general and the unique role belief has in reasoning, when deliberating about what to believe we are forced to (try to) form epistemically supported i.e. beliefs that we are willing to use in further reasoning. And this is why, through reasoning, we can’t come to believe what we want, but only what we take to have sufficient evidence for.

BIBLIOGRAPHY

- Adler, J. (2006). *Belief's Own Ethics*. Cambridge, MA: MIT Press.
- Alston, W. (1988). The deontological conception of epistemic justification. *Philosophical Perspectives* 2: 257-299.
- Alvarez, M. (2010). *Kinds of Reasons: An Essay on the Philosophy of Action*, Oxford: Oxford University Press.
- Ancombe, E. (1957) *Intention*, Oxford: Basil Blackwell; 2nd edition, 1963.
- Archer, S. (2015). Defending Exclusivity. *Philosophy and Phenomenological Research*, 92.
- Austin, J.L. (1975). *How to Do Things With Words*, Cambridge, Harvard University Press.
- Bach, K. & Harnish (1979). *Linguistic Communication and Speech Acts*. Cambridge, Mass., MIT Press.
- Bennett, J. (1990). Why Is Belief Involuntary? *Analysis* 50: 93- 101.
- Birch, S. A. J., Vauthier, S. A., Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, 107: 1018-1034.
- Boghossian, P. (2003). Blind Reasoning. *Aristotelian Society Supplementary Volumes* 177: 225–48.
- (2008). Epistemic Rules. *Journal of Philosophy* 105: 472–500.
- (2008). *Content and Justification*. Oxford: Oxford University Press.
- (2012). Blind rule-following. In Crispin Wright & Annalisa Coliva (eds.), *Mind, Meaning, and Knowledge: Themes From the Philosophy of Crispin Wright*. Oxford

University Press

---- (2015). Experience, Phenomenal Character and Epistemic Justification. *Philosophical Issues*, 25:243-251.

---- (2014). What is inference?. *Philosophical Studies*, 169: 1–18.

---- (2016). Rationality, reasoning and rules: reflections on Broome’s rationality through reasoning. *Philosophical Studies*

BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Boyle, M. (2009). Active belief. *Canadian Journal of Philosophy*, 41:

---- (2011). ‘Making Up Your Mind’ And The Activity Of Reason. *Philosophers’ Imprint* 11: 1-24.

Borgoni, C. (2015). Dissonance and Moorean Propositions, *Dialectica*, 69: 107-127.

Bortolotti, L. (2009). *Delusions and Other Irrational Beliefs*, Oxford: Oxford University Press.

Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, 138: 389–414.

Brandom, R. (1994). *Making it Explicit*. Cambridge: Harvard University Press.

Bratman, M. E. (1992). Practical reasoning and acceptance in a context. *Mind*, 101:1–15.

Broome, J. 2001). Are Intentions Reasons? And How Should We Cope with Incommensurable Values? In *Practical Rationality and Preference. Essays for David Gauthier*, edited by C. W. Morris and Arthur Ripstein, 98–120. Cambridge: Cambridge University Press.

- (2013). *Rationality through reasoning*. Oxford: Wiley Blackwell.
- (2014). Comments on Boghossian. *Philosophical Studies* 169: 19-25.
- Burge, T. (2003). Perceptual Entitlement. *Philosophy and Phenomenological Research* 67: 503–48.
- (2010). *Origins of Objectivity*. Oxford: Oxford University Press.
- Bykvist, K. and Hattiangadi, A. (2007). Does thought imply ought? *Analysis* 67: 277-285.
- (2013). Belief, truth, and blindspots. In *The Aim of Belief*, ed. T. Chan, 100-122. Oxford: Oxford University Press.
- Christensen, D. (2004). *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford: Oxford University Press.
- Clifford, W. K. (1999), The Ethics of Belief, in T. J. Madigan (ed.), *Ethics of Belief and Other Essays* Amherst, New York: Prometheus Book.
- Cohen, L. J. (1981). Can Human Irrationality Be Experimentally Demonstrated. *Behavioral and Brain Sciences*, 4:317-70.
- (1992). *An Essay on Belief and Acceptance*. New York: Clarendon Press.
- Currie, G. (2000). Imagination, Delusion and Hallucinations. In *Mind and Language*, Coltheart & Davies (eds.), 168-183, Blackwell.
- (2002). Imagination as Motivation. *Proceedings of the Aristotelian Society*, 102:201-16.
- Currie, G & Ravenscroft, I. (2002). *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford University Press.
- Dancy, J. (2004) From Thought to Action. In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Vol. 9. Oxford: Oxford University Press.

- Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy*, 60.
- (1973). Radical Interpretation. *Dialectica*, 27: 313–328
- (1980). *Essays on Actions and Events*. Oxford: Oxford University Press.
- (2004). *Problems of Rationality*, Oxford: Oxford University Press
- deFinetti, B. (1937). La Prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*
- de Sousa, R. B. (1971). How to give a piece of your mind: or, the logic of belief and assent. *Review of Metaphysics*, 25: 52–79.
- Dennett, D. (1978). *Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press
- Dretske, F. 2000. *Perception, Knowledge and Belief*. Cambridge: Cambridge University Press.
- (2001). Norms, History, and the Mental. In D. Walsh (ed.), *Evolution, Naturalism and Mind*. Cambridge University Press
- Doebel, S. & Koenig, M. (2013). Children's Use of Moral Behavior in Selective Trust: Discrimination versus Learning. *Developmental Psychology*, 49: 462-469.
- Douven, I. (2002). Testing Inference To The Best Explanation. *Synthese*, 130: 355–377
- Dub, R. (2017). Delusions, Acceptances, and Cognitive Feelings. *Philosophy and Phenomenological Research*, 94: 27–60
- Dummett, M. (1959). Truth. *Proceedings of the Aristotelian Society*, 59: 141- 162; Reprinted in his *Truth and Other Enigmas* (London: Duckworth 1978), 1- 24.
- Earman, J. & Roberts J. (1999). Ceteris paribus, There is No Problem of Provisos. *Synthese*, 118:439--478.

- Egan, A. (2008). Seeing and Believing: Perception, Belief Formation, and the Divided Mind. *Philosophical Studies* 140: 47-63.
- Egan, A. & Elga, A. (2005). I Can't Believe I'm Stupid. *Philosophical Perspectives*, 19: 77- 94.
- Engel, P. (2000) "Is Truth a norm?" in *Interpreting Davidson*, P. Pagin, G. Segal & P. Kotatko, eds. Stanford, CSLI.
- (2004) Truth and the aim of belief. In *Laws and Models in Science*, ed. D. Gillies, 77–97. London: King's College.
- (2013). In Defense of Normativism about the Aim of Belief. In T. Chan (ed.), *The Aim of Belief*, Oxford University Press
- Enoch, D. (2006). Agency, schmagency. *Philosophical Review* 115: 169-198.
- Evans, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.
- Evans, I. (2013). The Problem of the Basing Relation. *Synthese* 190, 2943-2957.
- Fantl, J. & M. McGrath (2010). *Knowledge in an Uncertain World*. New York: Oxford University Press
- Field, H. (2000). Indeterminacy, degree of belief, and excluded middle. *Noûs*, 34:1–30
- Fisher, A. (1989). Suppositions in Argumentation. *Argumentation*, 3: 401–413.
- Fodor, J. (1985). Fodor's guide to mental representation. *Mind*, 94, 76–100.
- (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT/Bradford.
- (1990). *A Theory of Content and Other Essays*. Cambridge: MIT/Bradford Press.
- Foley, R. (1992). The Epistemology of Belief and the Epistemology of Degrees of Belief. *American Philosophical Quarterly*, 29:111–124. □

- (1993). *Working Without a Net*. Oxford: Oxford University Press.
- (2009). Beliefs, Degrees of Belief, and the Lockean Thesis. In *Degrees of Belief*. Huber & Schmidt-Petri (eds.), Springer. pp. 37-47
- Frankish, K. (2007). Partial belief and flat-out belief. In F. Huber, & C. Schmidt-Petri (Eds.), *Degrees of belief: An anthology*. Oxford: Oxford University Press.
- (2004). *Mind and Supermind*. Cambridge: Cambridge University Press
- Frankfurt, H., (1992). The Faintest Passion. *Proceedings and Addresses of the American Philosophical Association*, 66, 3. 5-16.
- Frege, G. (1979). "Logic." In *Posthumous Writings*, edited by Gottfried Gabriel, Hans Hermes, and Peter Long, 1–9. Oxford: Blackwell.
- Friedman, J. (2013). Suspended judgment. *Philosophical studies*, 162: 165-181.
- Gendler, T. S. (2000). The Puzzle of Imaginative Resistance. *The Journal of Philosophy*, XCVII:55-81.
- (2003). On the relation between pretense and belief. In D. M. Lopes & M. Kieran (Eds.), *Imagination, philosophy and the arts* (pp. 125–141). London: Routledge.
- (2006). Imaginative Contagion. *Metaphilosophy*, 37: 1-21.
- (2007). 2007. Self-Deception as Pretense. *Philosophical Perspectives*, 21: 231-258.
- (2008). Alief and Belief. *Journal of Philosophy*, 105: 634-663.
- Gibbard, A. (1990/92). *Wise Choices, Apt Feelings: A Theory of Normative Judgment* Cambridge. Massachusetts: Harvard University Press, and Oxford: Oxford University Press.
- (2004). *Thinking How to Live*. Harvard University Press.
- (2005). Truth and Correct Belief. *Philosophical Issues*, 15: 338–50.

- Gibbons, J. (2009). Reason in action. In Lucy O'Brien and Matthew Soteriou (Eds.), *Mental Actions*, pp. 72–94. Oxford: Oxford University Press.
- (2013). *The Norm of Belief*, Oxford University Press.
- Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, 98: 254-267.
- (2000). *Adaptive thinking: Rationality in the real world*. New York: Oxford University Press.
- (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking Press.
- (2008). *Rationality for mortals: How people cope with uncertainty*. New York: Oxford University Press.
- (2015). *Simply rational: Decision making in the real world*. New York: Oxford University Press.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. New York: Oxford University Press.
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristics: Why biased minds make better inferences. *Topics in Cognitive Science*, 1:107–143.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103: 650-669.
- Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.
- Gigerenzer, G., Todd, P. M., & the ABC Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Glüer, K. and Wikforss, Å. (2009). Against content normativity. *Mind* 118: 31-70.

- (2013) Against belief normativity. In *The Aim of Belief*, ed. T. Chan, 80-99. Oxford: Oxford University Press.
- (2013a) Aiming at truth: on the role of belief. *Teorema* 32: 137- 162.
- Godfrey-Smith, P. (2004). "On Folk Psychology and Mental Representation," in H. Clapin, P. Staines, and P. Slezak (eds.) *Representation in Mind: New Approaches to Mental Representation*. Amsterdam: Elsevier Publishers, pp. 147-162.
- Goldman, A. (1979). What is justified belief? In G. S. Pappas (Ed.), *Justification and knowledge*. pp. 1-23. Dordrecht: D. Reidel. Reprinted in Sosa & Kim (Eds.)
- (2006). *Simulating Minds: The Philosophy, Psychology, and Neuro- science of Mindreading*. Oxford University Press, Oxford.
- Graham, P. (2011). Perceptual entitlement and basic beliefs. *Philosophical Studies* 153: 467- 475.
- (2012). Epistemic entitlement. *Noûs* 46: 449-482.
- Green, M. S. (2000). The Status of Supposition. *Noûs*, 34: 376–399.
- Griffiths et al. (2012). How the Bayesians Got Their Beliefs (and What Those Beliefs Actually Are): Comment on Bowers and Davis (2012). *Psychological Bulletin*, 138: 389–414.
- Haidt J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108: 814-834
- Hahn, U. & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review* 114:704–32.
- Harman, G. (1986). *Change in View: Principles of Reasoning*. Cambridge, Mass.: MIT Press.

- Harris, P. L. & Corriveau, K. H. (2011). Young children's selective trust in informants. *Philosophical Transactions of the Royal Society B*, 366, 1179-1187.
- Hattiangadi, A. (2006). Is meaning normative? *Mind and Language* 21: 220-40
- Hawthorne, J. & L. Bovens (1999). The Preface, the Lottery, and the Logic of Belief. *Mind*, 108: 241–264.
- Hawthorne, J. & Stanley, J. (2008). Knowledge and Action. *Journal of Philosophy* 105: 571–90.
- Hieronymi, P. (2005). The Wrong Kind of Reason. *Journal of Philosophy*, 102:437 - 45
- (2009). Two kinds of agency. In L. O'Brien and M. Soteriou (Eds.), *Mental Action*, pp. 138–162. Oxford: Oxford University Press.
- Hlobil, U. (2014). Against Boghossian, Wright and Broome on inference. *Philosophical Studies* 167: 419–429.
- Horwich, P. (1998). *Meaning*. Oxford: Oxford University Press.
- Huddleston, A. (2013). Naughty Beliefs, *Philosophical Studies* 160: 209-222
- Hunter, D. (2011). Alienated Belief. *Dialectica* 65: 221-240.
- Hursthouse, R. (1999). *On Virtue Ethics*. Oxford University Press
- Jackson, F. (1991). Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics* 101: 462–82.
- James, W. (1896). The Will to Believe. *The New World*
- Jeffrey, R.C. (1970). "Dracula Meets Wolfman: Acceptance vs. Partial Belief", in: M. Swain, ed., *Induction, Acceptance, and Rational Belief*, Dordrecht: Reidel, pp. 157–85
- Jones, M., & Love, B.C. (2011). Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition.

Behavioral and Brain Sciences, 34:169-188.

---- (2011). Pinning down the theoretical commitments of Bayesian models of cognition (response to commentaries). *Behavioral and Brain Sciences*, 34: 215-231.

Johnson-Laird, P. (2008). *How We Reason*. Oxford: Oxford University Press.

Ichikawa, J.J. & Jarvis, B. (2012). Rational Imagination and Modal Knowledge. *Noûs*, 46: 127–158.

---- (2013). *The Rules of Thought*. OUP Oxford.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Macmillan.

Kahnemann, D & Tversky, A. (Eds.). (1982). *Judgments under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.

Kaplan, M. (1981a). A Bayesian theory of rational acceptance. *Journal of Philosophy*, 78: 305–30.

---- (1981b). Rational acceptance. *Philosophical Studies*, 40: 129–45.

---- (1995). Believing the improbable. *Philosophical Studies*, 77: 117–45.

---- (1996). *Decision Theory as Philosophy*. Cambridge: Cambridge University Press.

Kripke, S. 1980. *Naming and Necessity*. Oxford: Blackwell.

Kelly, T. (2002). The rationality of belief and other propositional attitudes, *Philosophical Studies* 110:163-96.

---- (2003). Epistemic rationality as instrumental rationality: A critique, *Philosophy and Phenomenological Research*, 66: 612–640.

Kiesewetter, B. (ms). *The normativity of rationality*, Phd Dissertation

Kind, A. (2016). *The Routledge Handbook of Philosophy of Imagination*. Routledge

Koenig, M., Clément, F., & Harris, P.L. (2004). Trust in testimony: children's use of true

- and false statements. *Psychological Science*, 15: 694–698.
- Kolodny, N. (2005). Why Be Rational? *Mind* 114: 509–563.
- (2008). The Myth of Practical Consistency. *European Journal of Philosophy*, 16: 366–402.
- Korsgaard, C. (1997). The Normativity of Instrumental Reason. In *Ethics and Practical Reason*, edited by Garrett Cullity and Berys Gaut, 215–254. Oxford: Clarendon Press.
- (2008). *Self-Constitution. Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
- Kyburg, H. E. (1961). *Probability and the Logic of Rational Belief*, Middletown, CT: Wesleyan University Press
- Lance, M. N. (1995). Subjective probability and acceptance. *Philosophical Studies*, 77: 147-179.
- Langland-Hassan, P. (2012). Pretense, imagination, and belief: the Single Attitude theory. *Philosophical Studies* 159:155-179
- Lee, M. D. & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, 11:343–52.
- Leslie, A. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50, 211–238.
- Lewis, D. (1982). Logic for Equivocators. *Nous* 16, 431-441. Reprinted in D. Lewis, *Papers in Philosophical Logic*, Oxford, Oxford University Press, pp. 97-110.
- Littlejohn, C. (2012). *Justification and the Truth-Connection*. Cambridge: Cambridge

University Press.

Locke, D. (2015). Practical Certainty. *Philosophy and Phenomenological Research*, 90:72-95

Lord, E. (2010). Having Reasons and the Factoring Account. *Philosophical Studies* 149 (3): 283–296.

Lynch, M. (2009). The values of truth and the truth of values. In *Epistemic Value*, eds. A.Haddock, A. Millar and D. Pritchard, 225-242. Oxford: Oxford University Press.

Mackie, J. 1977. *Ethics: Inventing Right and Wrong*. Oxford: Oxford University Press.

Makinson, D. C., (1965) *Paradox of the Preface*, *Analysis* 25 205-207.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.

Mascaro O. and Sperber D. (2009) The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, 112, 367–380.

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

McHugh, C. (2012). The Truth Norm of Belief. *Pacific Philosophical Quarterly*, 93:8-30.

---- (2012). Belief and Aims. *Philosophical Studies* 160:425-439.

---- (2013). Normativism and Doxastic Deliberation, *Analytic Philosophy*, 54: 447-465.

---- (2014). Fitting Belief. *Proceedings of the Aristotelian Society* 114:167-187.

---- (2015) The Illusion of Exclusivity. *European Journal of Philosophy* 23:1117-1136.

McHugh C. & Way J. (2015). Broome on Reasoning. *Teorema* 34: 131–141.

---- (2016). What Is Good Reasoning? *Philosophy and Phenomenological Research* 94

---- (2016). What Is Reasoning?, *Mind*

---- (2016c). Against the Taking Condition. *Philosophical Issues* 26:314-331.

- Mercier, H., & Sperber, D. (2011). Why Do Humans Reason? Arguments for an Argumentative Theory. *Behavioral and Brain Sciences* 34: 57-74.
- Millar, A. (1994). *Understanding People*. Oxford: Oxford University Press.
- Millikan, R. (1984). *Language, Thought and Other Biological Categories*. MIT Press
- Moran, R. (2001). *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton University Press
- Nayding, I. (2011). Conceptual evidentialism, *Pacific Philosophical Quarterly*, 92: 39-65.
- Neta, R. (2013). What is an inference? *Philosophical Issues*, 23: 388-407.
- Neuman, Y. (2003) Go ahead, prove that God does not exist! On high school students' ability to deal with fallacious arguments. *Learning and Instruction*. 13:367 – 80.
- Neuman, Y., Weinstock, M. P. & Glasner, A. (2006) The effect of contextual factors on the judgment of informal reasoning fallacies. *Quarterly Journal of Experimental Psychology*, Section A: Human Experimental Psychology 59:411 – 25.
- Newell, B. R. & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing “one-reason” decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29:53–65.
- Nichols, S. & Stich, S. (2000): A cognitive theory of pretense. *Cognition*, 74:115–147.
- (2003). *Mindreading*. Oxford: Oxford University Press.
- Nichols, S., Stich, S., Leslie, A. and Klein, D. (1996). Varieties of off-line simulation. In *Theories of Theories of Mind*. P. Carruthers and P. Smith (eds). Cambridge: Cambridge University Press.
- Nickerson, R.S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises.

Review of General Psychology. 2: 175–220

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84: 231-259.

Nolfi, K. (2013). Why is Epistemic Evaluation Prescriptive? *Inquiry* 57:97-121.

---- (2016). How to be a normativist about the nature of belief. *Pacific Philosophical Quarterly*.

Noordhof, P. (2001). Believe What You Want. *Proceedings of the Aristotelian Society* 101 (3):247-265.

Noordhof, P. and Sullivan-Bissett, E. (2013). A defence of Owens' exclusivity objection to beliefs having aims. *Philosophical Studies* 163: 453-457

Norton, J. (2010). There Are No Universal Rules for Induction. *Philosophy of Science*, 77: 765-777

Nozick, R. (1993). *The Nature of Rationality*. Princeton, NJ: Princeton University Press.

O'Brien, L. (2005). Imagination and the Motivational View of Belief. *Analysis*, 65: 55–62.

Owens, D. (2003). Does belief have an aim? *Philosophical Studies* 115: 283-305

Papineau, D. (2013). There are no norms of belief. In *The Aim of Belief*, ed. T. Chan, 64-79. Oxford: Oxford University Press.

Parfit, D. (2001). Rationality and Reasons. In *Exploring Practical Philosophy: From Action to Values*, edited by Egonsson, Josefsson, Petersson, and Rønnow-Rasmussen, Aldershot: Ashgate.

---- (2011). *On What Matters*. Oxford: Oxford University Press.

Pennington, N. & Hastie, R. (1993) Reasoning in explanation-based decision- making.

Cognition 49:123–63.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge University Press.

Pettit, P. (1993). *The common mind: An essay on psychology, society and politics*. New York: Oxford University Press.

---- (2016) Broome on reasoning and rule-following, *Philosophical Studies*

Pollock, J. (1987), Defeasible Reasoning, *Cognitive Science* 11: 481-518.

Price, H. (2003). Truth as convenient friction. In de Caro & Macarthur (eds.), *Journal of Philosophy*. Columbia University Press. pp. 167--190

Pritchard, D. (2005). *Epistemic Luck*. Oxford: Oxford University Press.

Railton, P. (1994). Truth, Reason, and the Regulation of Belief. *Philosophical Issues*. 5: 71–93.

Rakoczy, H., & Tomasello, M. (2009). Done wrong or said wrong? Young children understand the normative directions of fit of different speech acts. *Cognition*, 13: 205-212

Ramsey, F. P. (1926). Truth and probability. In his *The Foundations of Mathematics*, Routledge and Kegan Paul (1931)

Rawls, J. (1995). Two Concepts of Rules, *Phil. Rev.*

Raz, J. (2005). The Myth of Instrumental Rationality. *Journal of Ethics and Social Philosophy* 1:1–28

---- (2009). Reasons: Practical and Adaptive. Reprinted in: *From Normativity to Responsibility*, 36–58. Oxford: Oxford University Press (2011).

Rey, G. (1988) Toward a computational account of akrasia and self-deception. In McLaughlin, B. P. and Rorty, A. O., editors, *Perspectives on Self-Deception*. University

of California Press.

---- (2007). Resisting normativism in psychology. In *Blackwell Debates in Philosophy of Mind*, eds. J. Cohen and B. McLaughlin, 69-84. Oxford: Blackwell

Rosen, G. (2001). Brandom on modality, normativity and intentionality. *Philosophy and Phenomenological Research* 63: 611-623.

Ross, J & M. Schroeder. (2014). Belief, credence, and pragmatic encroachment. *Philosophy and Phenomenological Research*, 88: 259-288.

Samuels, R., S. Stich and M. Bishop (2002). Ending the rationality wars: How to make disputes about human rationality disappear. In R. Elio (ed.) *Common Sense, Reasoning and Rationality*, Vancouver Studies in Cognitive Science, New York: Oxford University Press.

Sargent, M. (2009). Answering the Bayesian Challenge. *Erkenntnis*, 70: 237-252.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

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

---- (2007). Structural Irrationality. In *Common Minds. Themes from the Philosophy of Philip Pettit*, edited by Geoffrey Brennan, Robert Goodin, Frank Jackson, and Michael Smith, 84–103. Oxford: Oxford University Press.

Schroeder, M. (2008). Having Reasons. *Philosophical Studies* 139: 57–71.

---- (2009). Means-End Coherence, Stringency, and Subjective Reasons. *Philosophical Studies* 143: 223–248.

Schwitzgebel, E. (2010). Acting Contrary to our Professed Beliefs, or, The Gulf Between Occurrent Judgment and Dispositional Belief, *Pacific Philosophical Quarterly*, 91: 531-

553.

Searle, J. (1969) *Speech acts*. Cambridge University Press.

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

Sellars, W. 1963. *Science, Perception, and Reality*. London: Routledge.

----- (1968) *Science and Metaphysics: Variations on Kantian Themes*, Routledge & Kegan Paul Ltd; London, and The Humanities Press; New York

Setiya, K. (2013). Epistemic agency: some doubts. *Philosophical Issues* 23: 179–198.

Shah, N. (2001). Thinking Through Belief (Phd Dissertation)

---- (2003). How Truth Governs Belief , *Philosophical Review*, 112: 447-82.

---- (2006). A New Argument for Evidentialism, *The Philosophical Quarterly*, 56: 481–498

---- (2008). How Action Governs Intention, *Philosophers' Imprint*, 8:1–19

Shah, N. and Velleman, J. D. (2005). Doxastic Deliberation , *Philosophical Review*, 114: 497-534.

Sharadin, N. (2016). Nothing but the Evidential Considerations?. *Australasian Journal of Philosophy*, 94: 343-361

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63:129–138.

Sinhababu. N. (2013). Distinguishing belief and imagination. *Pacific Philosophical Quarterly*, 94:152–165

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Smithies, D. (2012). The normative role of knowledge. *Noûs* 46: 265-288.

- Sosa, E. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume One*. Oxford: Oxford University Press.
- (2009). *Reflective Knowledge: Apt Belief and Reflective Knowledge, Volume Two*. Oxford: Oxford University Press.
- (2011). *Knowing Full Well*. Princeton, NJ: Princeton University Press.
- Sperber, D. (1997) Intuitive and reflective beliefs. *Mind and Language* 12:67– 83. [
- (2000) *Metarepresentations: A multidisciplinary perspective*. Oxford University Press.
- Steglich-Petersen, A. (2006), The Aim of Belief: no Norm Needed, *Philosophical Quarterly*, 56: 500-516.
- (2008), Does Doxastic Transparency Support Evidentialism?, *Dialectica*, 62: 541–547
- (2008). Against essential normativity of the mental. *Philosophical Studies* 140: 263-283.
- (2013) The no guidance argument. *Theoria* 79: 279-283.
- (2009). Weighing the Aim of Belief, *Philosophical Studies* 145: 395–405.
- Stein, E. (1996). *Without Good Reason*. Oxford: Clarendon Press.
- Stich, S. (1983). 1983. *From Folk Psychology to Cognitive Science*. Cambridge, Mass.: MIT Press.
- (1990). *The Fragmentation of Reason*. Cambridge, MA: MIT Press.
- (1991) Do true believers exist? *Proceedings of the Aristotelian Society*, Suppl. 65: 229–44.
- Toribio, J. (2011). What we do when we judge. *Dialectica* 65: 345-367.

- (2013). Is there an ‘ought’ in belief? *Teorema* 33: 75-90.
- Tucker, C. (2012). Movin’ on up: Higher-Level Requirements and Inferential Justification. *Philosophical Studies* 157: 323–40.
- Turri, J. (2011). Believing For a Reason. *Erkenntnis*. 74: 383-397
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211: 453–458.
- Valaris, M. (2014). Reasoning and regress. *Mind*, 123:101–127.
- (2016) What Reasoning Might Be. *Synthese*.
- Van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- Van Leuween, N. (2009). The motivational role of belief. *Philosophical Papers*, 38, 219–246
- (2014). Religious Credence is not Factual Belief. *Cognition*, 133, 698-715.
- Velleman, D. (2000). *The possibility of practical reason*. Oxford: Oxford University Press.
- Wason, P. C., and Evans, J. S. B. T. 1975. Dual processes in reasoning? *Cognition*, 3: 141–54.
- Way, J. (2009). Two Accounts of the Normativity of Rationality. *Journal of Ethics and Social Philosophy* (Discussion Note): 1–8.
- Wedgwood, R. (2002). The Aim of Belief. *Philosophical Perspectives*, 16, 268-297
- (2006). The Normative Force of Reasoning. *Noûs*, 40: 660–686.
- (2007). *The Nature of Normativity*. Oxford: Oxford University Press.
- (2009). The normativity of the intentional. In *The Oxford Handbook of the Philosophy of Mind*, eds. A. Beckermann and B. McLaughlin, 421-437. Oxford: Oxford

University Press.

---- (2013). The right thing to believe. In *The Aim of Belief*, ed. T. Chan, 123-139.

Oxford: Oxford University Press.

Whiting, D. 2010. Should I believe the truth? *Dialectica* 64: 213-224.

---- (2012) Does belief aim (only) at truth? *Pacific Philosophical Quarterly* 93: 279-300.

---- (2014). Reasons for Belief, Reasons for Action, the Aim of Belief, and the Aim of Action, in *Epistemic Norms: New Essays on Action, Belief, and Assertion*, ed Clayton Littlejohn and John Turri

Williams B. (1973). Deciding to Believe, in *Problems of the Self*. Cambridge: Cambridge University Press

---- (1979). Internal and External Reasons. Reprinted in: *Moral Luck*, 101–113.

Cambridge: Cambridge University Press (1981).

Williams J. N. (2004). Moore's Paradoxes, Evans's Principle, and Self Knowledge. *Analysis* 64: 348–353.

Williamson, T. (1996). Knowing and Asserting. *Philosophical Review* 105: 489-523.

---- (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Wittgenstein, L. (1953), *Philosophical Investigations* (Oxford: Blackwell).

Worsnip, A.L. (ms) *Rationality's Demands on Belief*. PhD Dissertation.

Wright, C. (1992). *Truth and Objectivity*. Harvard University Press.

---- (2004). Warrant for nothing (and foundations for free)? *Aristotelian Society Supplementary Volume*. 78: 167–212

---- (2014). Comment on Paul Boghossian, "What Is Inference". *Philosophical Studies*,

169: 27-37.

Yablo, S. (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, 53: 1–42.

---- (2002a). Go Figure: A Path Through Fictionalism. *Midwest Studies in Philosophy*, 25: 72–102.

---- (2002b). Abstract Objects: A Case Study. *Noûs*, 36 (Supplementary Volume 1): 220–240.

---- (2005). The Myth of the Seven. In *Fictionalism in Metaphysics*, M. Kalderon (ed.), New York: Oxford University Press, pp. 88–115.

Zalabardo, J. (2010). Why believe the truth? Shah and Velleman on the aim of belief, *Philosophical Explorations*, 13: 1-21

Zangwill, N. (1998): Direction of fit and normative functionalism. *Philosophical Studies* 91: 173-203.

---- (2005). The normativity of the mental. *Philosophical Explorations* 8: 1-19.

---- (2009). Normativity and the metaphysics of mind. *Australasian Journal of Philosophy* 88: 1-19.

CURRICULUM VITA

Marianna Bergamaschi Ganapini will be an Assistant Professor in the Philosophy Department at Union College starting Fall 2017. In 2004 she received her B.A. from the University of Florence (Italy). She holds also a doctoral degree from the University of Florence, and she was a visiting student at Georgetown University in 2007-2008. Her main areas of research are philosophy of mind and epistemology. She is one of the recipients of Dean's Teaching Fellowship for Spring 2015. In 2014 she was awarded with a Sachs Fellowship, which is a dissertation-year fellowships awarded to students who are making substantial progress toward completing their dissertations. Over the last few years she has given presentations at several national and international conferences and workshops. Before finishing her PhD., she wrote two articles, one on philosophy of language and the other on metaethics, which are published in *Philosophia* and in the *International Journal of Philosophical Studies*.