Normal human adults have beliefs and intentions upon which they act. But other creatures have functionally similar states. Think of a cat stalking a mouse. It is quite natural to describe the cat as believing that the mouse is around the corner, having the goal of catching it, and racing around the corner in order to pounce on it. Unlike cats and other animals, though, we are not merely passive recipients of our beliefs and intentions; we are capable of reflectively governing them. We do not have this kind of governance over all of our mental states. Although we can reflect on the way things look to us (e.g. the stick in water looks bent), reflection alone is powerless to change our perceptions. They are what they are, whether we reflect on them or not. But normally when we reflect on our attitudes, we do not merely come to know what we in fact believe or intend; we determine what we shall believe or intend. This influence that reflection has over our attitudes is unmediated—we can normally determine what we believe or intend by reflection alone.

However, only certain kinds of reflection have an unmediated influence over our attitudes. For example, when we reflectively conclude that it would be desirable to believe what we are already inclined to believe, our reflection alone will not issue in our believing it. But when we reflectively conclude that our evidence confirms what we are already inclined to believe, our reflection alone normally will issue in our believing it. Even if you are inclined to believe that you will receive a promotion, for example, and even if you reflectively conclude that you will be happier if you believe that you will receive a promotion, you will not form that belief on that basis. But if you reflectively conclude that your evidence confirms that you will receive a promotion, then you normally will form that belief on that basis. There is of course a parallel here with intention. When we step back and try to
determine what to intend, we find that only reflection on the desirability of an act can move us to intend it. As the toxin case makes vivid, reflection on the desirability of intending an act, as opposed to the desirability of the act itself, will not issue in an intention to perform the act.  

Although these are fairly obvious facts about reflection, I think the correct explanation of them reveals something interesting about how we conceive of our attitudes. It may be objected right away that these are contingent facts, so explaining them cannot yield any insight about how we conceive of our attitudes—no conceptual truths. After all, it seems that we can imagine these facts not obtaining. All we have to do is imagine that reflection on our attitudes is unconstrained, that we have whatever attitudes we judge desirable to have, and have them just because we so judge. But if we can imagine this, then the fact that our unmediated reflective governance over our attitudes is constrained is an empirical fact, not a conceptual truth.  

But here we should ask why it is that what we seem to be able to imagine—that we form an attitude precisely because we judge it would be desirable to have that attitude—does not in fact occur. I think that once we have the correct answer to this question, we will see that we can’t imagine our attitudes being governed by judgments about their desirability. If we think this is something we can imagine, it is because we haven’t fully conceived of the attitudes involved as beliefs and intentions or understood the nature of reflective governance.  

In this paper I will focus on reflection about what to believe, but I think that much of what I say applies, with a bit of tweaking, to reflection about what to intend as well. Why is it that reflection on the evidence that \( p \), but not on the desirability of believing that \( p \), can issue in believing that \( p \)? In the past I have argued that it is because our reflection about what to believe is structured by our acceptance of the following norm:

**The Norm of Truth:** Believing that \( p \) is correct iff it is true that \( p \), and incorrect iff it is untrue that \( p \).  

The norm of truth explains the evidential constraint on reflection about what to believe because that norm is contained in the very *concept* of belief. Anyone who applies this concept to his reflection must also attempt to conform to the *norm of truth*. Evidence that bears on the truth of \( p \), not on the desirability of believing that \( p \), determines whether, by believing that \( p \), one would be complying with this norm.  

Of course, aiming to arrive at a belief doesn’t typically involve explicitly asking oneself what to believe or representing the standard of correctness associated with belief. The norm is not explicitly represented because it is not a premise in reasoning. Its role is different. The aim of arriving at a correct belief provides the motive for such reasoning; the norm thus determines what can count as a premise in such reasoning.
If this is the right explanation of why we can’t decide to have a belief just because it would be desirable to have it, then our inability to believe at will is not merely a contingent fact about us. Anyone who seeks to reflectively determine what to believe must apply the norm of truth, and the application of this norm forecloses believing at will. If we think we can imagine a creature that does believe at will, that is because we are failing to fully acknowledge that the activity of forming a belief is constrained by the truth norm.

The focus of this paper is the objection that this explanation portrays reasoning as overly self-conscious. The objection can be spelled out as follows. When our reasoning concludes in a belief, we are not, as I claim, typically reasoning about the belief, we are reasoning about its content. We do not start our doxastic reasoning with the question whether to believe that p, but directly with the question whether p. And if we do not ask ourselves what to believe when our reasoning issues in belief, then our reasoning need not be governed by our application of the norm contained in the concept of belief. The evidential constraint on reasoning thus cannot be derived, as I claim, from the allegedly normative character of the concept of belief.

In short, my response is that not only can the evidential constraint be derived from the agent’s reflective awareness that his activity, belief-formation, is subject to the truth norm, it must be so derived—no other derivation is possible. To the extent the objection’s description of doxastic reasoning masks this fact, it is misleading.

The paper proceeds as follows. First, I attempt to show that the most plausible non-reflective explanation of the evidential constraint on reasoning fails. Next I argue that the features of reasoning, such as rule-following, that appear to rule out a reflective account of reasoning do not in fact do so. I then clear up a recent misunderstanding about the role of the truth norm in reasoning. Finally, I conclude by describing how our understanding of reflective reasoning may be central to all of our ascriptions of reasoning, even ascriptions of “non-reflective” reasoning.

Teleology

If reasoning is a non-reflective phenomenon, then, it might be thought, we should look to non-reflective mental states to explain its character. This is just what the teleological account of belief attempts to do. It claims that the source of the evidential constraint on reasoning does not lie in our thoughts about belief, but in the nature of belief itself. It is because belief aims at the truth that doxastic reasoning is subject to the evidential constraint.

But what does it mean to say that belief has an aim? Our primary understanding of aims and aim-guided activity is our understanding of intentional
action. We act in order to achieve our ends, i.e. we go to the store in order to get milk. In explaining actions we ascribe aims to whole individuals, not to mental states. Having the aim of getting milk may be a matter of being in a certain mental state such as the intention to get milk, but it is not the intention that has this aim, it is the agent who has the aim in virtue of having this intention. So might the idea that belief has an aim be shorthand for claims about an aim that an agent has in virtue of having a certain intention with respect to his beliefs? In other words, might the idea be that believing that \( p \) involves an intention to believe that \( p \) if and only if \( p \), and that it is in virtue of my having this intention that I count as aiming at truth when I believe that \( p \)?

As it stands, this thought cannot explain how I arrive at a belief by reflection. According to the view under consideration, I only have the intention that guides my reflection if I already have a belief about \( p \). At best, the invocation of this intention can explain reflection I engage in once I already believe that \( p \). But how can it explain the reflection by which I come to believe that \( p \)? It can do so if we follow the advice of some functionalists in the philosophy of mind and individuate beliefs by their causal antecedents: if an attitude counts as a belief that \( p \) only if it was brought about by an intention to have it if and only if \( p \), then the only kind of reasoning that can bring about this attitude is reasoning that is guided by this intention.

The problem with this explanation is that it over-generalizes. If an intention to believe the truth always plays a stand-by role in belief-formation, but it is only active in instances in which we have doubts about the reliability of a source. Questions about the reliability of a source do not prompt me to hesitate about whether to fantasize that I will be promoted, but they do prompt me to hesitate about whether to believe that I will be promoted. The intention to believe the truth, it might be claimed, explains this difference.

But what are we to say about cases in which non-evidential considerations cause us to believe things? Wishful thinking is an unfortunate fact of our mental lives. In such cases it looks as if the intention to believe the truth, if it were operative, should have stopped us from believing what we wanted to believe. But if the intention to believe the truth doesn’t exert an influence in these cases, in what sense is it constitutive of belief as such?
Might the teleologist say that in these cases the normal influence of the intention to believe the truth was obstructed, and thus that it is constitutive of belief that it is *normally* under the control of this intention? This would allow for *abnormal* cases of belief-formation that are not evidentially constrained, such as beliefs induced by wishful thinking. So far, so good. But what the teleologist cannot explain is why there are no abnormal cases of one particular kind of belief-forming process, reasoning. If the intention constitutive of belief can be overridden or prevented from having its normal influence, there ought to be clear instances of doxastic reasoning in which this intention fails to play its normal role. These would be instances of evidentially unconstrained doxastic reasoning. But there are no such instances. Although of course we sometimes arrive at a belief by reasoning badly, when we do so our reasoning constitutes bad evidential reasoning. What we do not, and cannot, do is arrive at a belief by reasoning on the basis of non-evidential reasons. We cannot reason our way to the belief that we will be promoted merely from the judgment that it would make us happy to believe this. Our wishes can affect what we believe, but they cannot be the reasons upon which we base our beliefs. Yet, for all the teleologist has said, this ought to be possible. Contrary to what his view implies, the evidential constraint applies to all instances of doxastic reasoning.9

Here is the basic problem. The teleologist must supply some intelligible connection between the aim constitutive of belief and the mechanisms that causally regulate belief. She must commit herself to something along the following lines: Beliefs have the aim of accurately representing the way things are, and it is because they have this aim that they are causally regulated by evidentially sensitive mechanisms. Such a view, however it is cashed out, is committed to the claim that whenever we have a belief, we aim at truth, even when the belief is not consciously held or even consciously accessible.10 Whatever explanatory work is done by this aim thus is independent of whether it is consciously accessible. But it is this feature of the view that prevents it from explaining the unique way in which conscious reasoning is constrained. As I said before, it is a striking fact that, of the processes that causally regulate our attitudes, conscious reasoning is the only one that is constrained in the ways that I described. I may be led to unconsciously believe that $p$ because I find it desirable to have this belief, but wishful thinking exerts its powerful effects only when it is masked. When I consciously deliberate about whether to believe that $p$, I cannot take the desirability of believing that $p$ to be my reason to believe it. What requires explanation, then, is why, in contrast to other attitude-forming processes, conscious reasoning is constrained. According to the teleological view, I have an aim whenever I have a belief. Given that all beliefs have aims, these aims should have the same effect on all belief-forming processes. It thus is difficult to see how these aims can explain the unique constraint on one particular type of belief-forming process.11,12
The Authority of the Truth Norm

In fact, though, we needn’t look at non-reflective beliefs to see why the teleological view won’t work. With respect to any aim-constituted activity, there are two kinds of deliberative questions one can ask. One can deliberate about how to achieve the aim, or one can deliberate about whether to continue to have the aim. In the midst of playing chess, I can deliberate about what move to make next. This decision will be guided by the aim of checkmating my opponent, an aim that is constitutive of playing chess.\textsuperscript{13} But I can also deliberate about whether to abandon the game and do something else instead. The decision to abandon the game is a decision in part to abandon chess’s constitutive aim.

If the teleological view of belief were correct, there thus should be two kinds of deliberation possible with respect to belief. We should be able to deliberate about how to achieve a belief’s aim, but we should also be able to deliberate about whether to abandon its aim. We should be able to retain or abandon a belief by focusing on the reasons determined by the aim involved in the belief, but we also should be able to retain or abandon it just by focusing on the reasons for having or not having the aim itself. For example, there should be cases in which we stop believing that $p$, not because we find ourselves with new evidence against the truth of $p$, but simply because we decide to give up the aim of truth with respect to our acceptance of $p$. Yet this is something we cannot do. I currently believe that I’m going to die on the basis of the evidence that I, like everyone else, am mortal. But believing that I am going to die makes me unhappy. Nonetheless, this fact cannot move me to abandon my belief. I cannot decide to stop believing that I am going to die just because I determine that it is desirable to stop aiming to represent the truth with respect to my mortality.

If we could abandon a belief because we judged that it was undesirable to continue to hold the belief, this would mean that we could abandon the belief at will. But we can neither form nor abandon a belief at will. Although the teleological view correctly implies that we cannot do the former, it incorrectly implies that we can do the latter.

The fact that we can neither form nor abandon a belief at will shows that the evidential constraint applies not only to cases in which we form beliefs, it applies as well to cases in which we abandon them. There are only two ways of abandoning a belief that $p$: believing that-$p$, or suspending belief. Evidence constrains reasoning about whether to suspend belief just as much as it does whether to believe that-$p$.

This feature of doxastic reasoning reflects an important feature of the truth norm. Unlike the norms governing a game, whose authority depends on the desirability of playing the game, the authority of the truth norm does not depend on the desirability of having a belief. This may be obscured by the fact that reasoning preceding a belief sometimes occurs in two distinct stages.
There are occasions when we decide to deliberate about whether to believe some proposition because we realize that the benefits involved in making up our mind now about the issue outweigh the burdens. We determine that we have the time, information, and peace of mind to form a belief about the matter. It is only after we have made a determination that forming a belief is worth doing that we go ahead and engage in deliberation governed by the truth norm. Because reasoning sometimes involves these two stages, it might look as though the authority of the truth norm depends on the prior judgment that it would be desirable to form a belief.

But this would be a mistake. As the previous example illustrates, even if we come to realize that the costs of holding onto a belief are much higher than we had realized, we will not be able to decide whether to *retain* this belief except by considering the evidence. If we conclude that the evidence confirms the proposition we believe, then we will retain it; if we conclude otherwise, then we will not. This, I take it, is just what the deliberative perspective is like. But it *wouldn’t* be like this if the authority of the truth norm were contingent on the value of having the belief. If its authority depended on the desirability of having beliefs, then we would be able to ignore it whenever we judged that a belief it governed was too costly. But we can’t do this.

In this way, the authority of the truth norm is like the authority of moral norms. Having made a promise to pick up a friend from the airport, you find yourself obliged to do so, even though you’d like to stay home and watch the game. That it is inconvenient for you to be required to leave now for the airport may lessen the motivational force of the obligation, but it will not diminish its reason-giving force. Thinking that it would be inconvenient to pick up your friend cannot lead you to conclude that you are no longer obliged to do so.

We really can’t help having beliefs, and once we have them it is incoherent to reject the norm of truth that governs them. This norm thus has a kind of authority that teleological accounts of belief cannot explain, and that meta-normative theorists ought to find as puzzling as they find the norms of morality.

**Is Reflection Irrelevant to Reasoning?**

Recall that the objection to the normative account’s depiction of reasoning is that it is overly self-conscious. Reasoning is a non-reflective process; it does not involve the higher-order attitudes to which the normative account appeals. It might appear that this objection is supported by the fact that when one reasons, one’s focus is on the content of one’s belief—the proposition believed, not the fact that one believes it. However, the normative account doesn’t deny that our reasoning is focused on the contents of our beliefs. If reasoning about whether to believe that *p* involves applying the norm of
truth, one must focus exclusively on the status of \( p \) in order to comply with
the norm. One applies the norm just in virtue of applying the concept of be-
lief; this accounts for the seamless transition between thinking about whether
to believe that \( p \) to thinking about whether \( p \).

However, there is a better way of developing the objection.\(^{14} \) I claim
that the evidential constraint on reasoning is brought about by the subject’s
application of the norm of truth, and the subject’s application of this norm
is entailed by the fact that his deliberation is aimed at determining what to
believe. But even reasoning that does not involve the subject’s conception of
what he is doing—first-order reasoning—is evidentially constrained. Some-
thing else thus must explain its evidential character. Furthermore, whatever
explains its evidential character is also present in self-conscious reasoning.
Thus, it might be thought, once we have explained the evidential character
of first-order reasoning, there will be no need to add anything to explain the
evidential character of self-conscious reasoning.

It does seem to be true that one kind of truth-preserving reasoning—
deductive reasoning according to the rules of logic—cannot be explained
by invoking the subject’s conception of what he is doing. Here, very briefly,
is why. All deductive reasoning involves rule-following. For example, when
we reason from the beliefs that \( p \) and if \( p \) then \( q \) to the belief that \( q \), we
follow the rule of modus ponens. Although it is notoriously difficult to give
necessary and sufficient conditions for a movement of thought to count
as an instance of rule-following, a familiar lesson from the failure of such
attempts is that rule-following cannot be explained by our attitudes about
what we are doing. As Paul Boghossian has pointed out, attitudes, including
attitudes about the rules one is following, only bear on one’s reasoning via
the application of a rule.\(^{15} \) If we were to try to explain a particular instance
of rule-following by invoking an attitude about the rule we were following,
we would need to presuppose a second instance of rule-following in order to
explain how this attitude brought about the first instance of rule-following.
Rule-following thus appears to be metaphysically prior to reflection. So if
the rule-following character of deductive reasoning is what accounts for the
evidential constraint on doxastic reasoning, then any account, including the
normativist one, that attempts to explain the evidential constraint in terms
of the subject’s conception of what he is doing, is doomed.

Whether the rule-following character of deductive reasoning is the source
of the evidential constraint on such reasoning depends on what the correct
explanation is of why we follow the deductive rules that we do. Why do we
follow the rules of logic as opposed to the rules of counter-logic? If the answer
is that the rules of logic are truth-preserving but the rules of counter-logic
are not, then it looks as if the evidential constraint on deductive reasoning
cannot be explained by the fact that we follow the rules of logic, because it
is needed to explain why we follow these rules.
However, a different explanation of why we follow the rules of logic is possible: following the rules of logic is constitutive of concept-possession. According to this hypothesis, we would not count as thinkers (i.e., possessors of concepts) unless we followed the rules of logic. There is no alternative to reasoning logically; reasoning counter-logically is unintelligible. If this is so, then we do not reason deductively in order to form true beliefs, or for any other reason.

If conformity to the rules of logic is a necessary condition for the possession of concepts, we need not see the evidential constraint on deductive reasoning as the product of our reflective application of a norm to our reasoning. Instead, we can view the evidential constraint as the upshot of the preconditions for reasoning. In order to reason we must possess concepts, and in order to possess concepts we must follow the rules of logic. And it is because we follow the rules of logic that our reasoning is evidentially constrained.

This kind of argument, if it works, may provide a non-reflective explanation of the evidential constraint on deductive reasoning, but can it explain the evidential constraint on non-deductive reasoning? I don’t think so. I doubt that we can explain why we follow inductive rules by appealing to the unintelligibility of doing otherwise. It may be plausible that in order to possess concepts, we must reason logically as opposed to counter-logically, but it is not plausible that in order to have concepts, even empirical concepts, we must also reason inductively, as opposed to counter-inductively. Although counter-induction may be irrational, it seems perfectly coherent.

Following the rules of induction, unlike the rules of logic, does not appear to be a necessary condition for thinking. Yet non-deductive reasoning is as evidentially constrained as deductive reasoning. We can imagine following different non-deductive rules than the ones we actually follow—we can imagine reasoning counter-inductively—but we cannot imagine following evidentially unconstrained rules. When we inductively arrive at the belief that the sun will rise tomorrow on the basis of our evidence that it has always done so in the past, we are reasoning under the evidential constraint: if we do not think that these facts are evidence that the sun will rise tomorrow, we will not arrive at that belief. Why is this so? Why can’t our reasoning be evidentially unconstrained in such cases? Why can’t we decide to believe that the sun will rise tomorrow on the ground that it makes us happy to believe it? It cannot be because following the rule of induction that we actually follow is constitutive of having the concepts involved in beliefs about the sun, because this is not so. Reasoning differently about whether the sun will rise tomorrow on the basis of our evidence that it has always done so in the past appears intelligible, even if irrational.

The general requirements on concept-possession have been exhausted; they cannot help us to explain why all doxastic reasoning is evidentially constrained. The natural place to look for an explanation of this constraint
is in the nature of belief. If a teleological account of the nature of belief could explain the evidential constraint, then there would be no need to advert to the subject’s conception of what he is doing. But, as I have argued earlier in the paper, a teleological account of belief cannot explain this constraint.

The Role of the Truth Norm

The normative account has no difficulty explaining the evidential constraint on doxastic reasoning and the authority of the truth norm. Engaging in doxastic reasoning involves attempting to conform to the standard of correctness for belief. The only way to conform to this standard is to conform to the evidential constraint. The authority of the truth norm is not dependent on whether one has the aim associated with belief, so one cannot escape its grip by abandoning this aim. One has the aim because one accepts the norm, not the other way round. And one accepts the norm because one recognizes that one has, or is attempting to have, the attitude to which it applies.\(^\text{17}\)

Recently, several philosophers have objected that the mere fact that belief involves a constitutive norm won’t explain the evidentially constrained character of doxastic reasoning.\(^\text{18}\) Unless we also have an overriding motive to comply with this norm, nothing follows about how we will reason, and the mere fact that this norm is constitutive of belief does not entail such a motive. After all, the mere fact that there is a constitutive norm of promising that one should keep one’s promises doesn’t entail that someone who makes a promise can’t decide to break it, or that he can’t decide to make a promise he intends to break. Unfortunately, people do such things all too often. So the mere fact that the norm of truth is constitutive of belief does not entail that we will attempt to comply with it.

Furthermore, there is a good reason to think that normative judgments cannot entail overriding motives. It is a central fact about norms that they can be violated. As exemplified by the case of promise breaking, we don’t always do what we ought to do, or even what we judge that we ought to do. So even if we did accept a norm for our beliefs, this would not guarantee our compliance with it. And if our acceptance of the norm did guarantee compliance, then it wouldn’t really be a norm, since then we couldn’t violate it.

I think this objection is based on two mistakes, one about what we are attempting to do when we engage in deliberation, and another about the constitutive norm for promising. When we deliberate about whether to believe that \(p\), we are attempting to form a belief on the basis of reasons. Thus the deliberative question whether to believe that \(p\) must be a normative question. The hypothesis that the norm of truth is constitutive of belief is meant to explain why this question can only be settled by determining whether there is evidence that \(p\). In applying the concept of belief, we recognize that believing that \(p\) is correct iff \(p\). If this is so, we cannot see ourselves as deliberating
whether to believe that $p$ unless we seek to determine whether $p$. If we seek the answer to any other question, we can no longer see ourselves as arriving at a belief on the basis of reasons. We can seek to determine whether it would be desirable to believe that $p$, but if we do, we can no longer see ourselves as deliberating about whether to believe that $p$. The problem is not that we will be deliberating badly about whether to believe that $p$, it is that if we focus on whether it would be desirable to believe that $p$, we will no longer be able to see ourselves as even attempting to determine what we have reason to believe. This doesn’t mean that we will actually believe every proposition for which we have sufficient evidence is true. But it does mean that we cannot believe a proposition on the basis of deliberation—on the basis of our attempt to believe what we have most reason to believe—unless we judge that we have sufficient evidence that it is true.

But then why are we able to make a promise without regard for the constitutive norm for promising? Why are we able to decide to make promises that we have no intention of keeping? We would not be able to do so if the constitutive norm of promising was that one has most reason to keep one’s promises. But this is not promising’s constitutive norm. Promising is a social exchange between a promisor and promisee. The primary norm governing this social practice is that the promisor has an obligation to the promisee to fulfill his promise. Of course for there to be a point to this practice, for promises to be of any value to those to whom they are given, it must be the case that people by and large see themselves as having reasons to fulfill their promissory obligations. But it does not follow that for every instance of promising, a promisor has such a reason. Although I must see myself as obliged to pay back my loans, there may be times when I don’t treat this obligation as a reason to do so.

Here is another way of making the point. In the first instance, promising is a social practice. Individual instances of promising only count as such against the background of a social practice in which most people take themselves to have decisive reason to keep most of their promises. The constitutive norm of promising thus operates at the level of the social practice. A practice would not count as a practice of promising unless most people treated their promises normatively in this way. But believing is an attitude, not a practice. Whether an instance of belief counts as such does not depend on whether the believer is part of any particular social practice. If there is a constitutive norm for belief, it governs each instance of belief, not the social practice of believing.

**Conclusion**

Even though we haven’t found a non-reflective account of reasoning that can explain its evidential character, isn’t it just implausible that the
The explanation of the evidential character of reasoning involves an agent’s conception of what he is doing. No more so, I’m tempted to say, than that the explanation of action involves an agent’s conception of what he is doing. Although we may be willing to call some bodily movements actions even though the agent of those bodily movements doesn’t understand what he is doing, nonetheless, we understand those movements as actions only by reference to the paradigm cases of action, intentional action. Similarly, I am tempted to argue, we may be willing to call some instances of thinking reasoning even though they don’t involve the agent’s conception of what he is doing, but we understand those cases as instances of reasoning only by reference to paradigm cases of reasoning, reflective reasoning. We call some but not all non-reflective processes “reasoning” because they resemble reflective reasoning in being rule-guided and conforming to the evidential constraint. Although it may be true that, due to their rule-following character, some aspects of reasoning must be non-reflective, other aspects of reasoning are fundamentally reflective. Non-reflective reasoning is not prior in the order of understanding to reflective reasoning; rather, what we are willing to call “non-reflective reasoning” depends on its resemblance to reflective reasoning.

Notes

1. For discussion of material in this paper, thanks to Conor McHugh, David Owens, Casey Perin, Matthew Silverstein, Sharon Street, and David Velleman, as well as audiences at the University of Reading, the Aims and Norms Workshop at the University of Southampton, and the Normativity and Reasoning Workshop at NYU Abu Dhabi.

2. Of course your reflection alone might non-rationally cause your believing it, as in cases of wishful thinking or self-deception. But in these cases your reflection alone does not issue in your believing it—as I understand the relation of issuing in—because you do not reflectively endorse it on the basis of your reasoning.

3. Similarly, being offered money to desire a plate of mud makes it desirable to desire a plate of mud, but because it does not make the plate of mud desirable, reflecting on this fact alone will not issue in the desirable desire.


5. In Shah (2008) I argue that the best explanation for the desirability constraint on practical reflection is similar to the best explanation for the evidential constraint on doxastic reflection. Just as reflection aimed at determining what to believe is structured by the norm of truth, reflection aimed at determining what to intend is structured by acceptance of the following norm:

   \textbf{The Norm of Desirability:} Intending to $\varphi$ is correct only if $\varphi$-ing is desirable, and incorrect if $\varphi$-ing is undesirable.

This norm explains our reflection about what to intend because this norm is contained in the very concept of intention. Anyone who possesses this concept, and applies it to his reflection, must also apply the norm of desirability. That
$\varphi$-ing is desirable is by itself pertinent to determining whether, by intending to $\varphi$, one would be complying with this norm; that it would be desirable to intend to $\varphi$, however, is not.

6. Compare Stroud (2011, 91): “There is no suggestion that such assessments or evaluations that are essential to belief and action take the form of an agent's explicitly formulating some evaluative thoughts “before the mind.” The presence of the attitudes is shown by what the person says and does, and in what it takes to make sense of someone as believing such-and-such or doing something or other intentionally.”

7. See Zalabardo (2010, 8) for a nice formulation of this objection. I also have heard this objection numerous times in conversation.

8. Notice that this explanation of doxastic reasoning seems to require that such reasoning involves thinking about one's attitudes. If the intention that guides such reasoning is an intention to believe that $p$ if and only if $p$, then the intention is about one's belief. But it was on the basis of denying that reasoning is about one's attitudes that the teleologist rejected the normative account. I leave it to those teleologists who think that reasoning is first-order to explain how their view avoids making reasoning objectionably second-order.

9. See Owens (2002, 296–300) for another argument that ascribing an aim of truth cannot explain the evidentially constrained nature of doxastic reasoning.

10. As I said in footnote 6, if the aim is an intention to have a belief or intention that meets certain conditions, then even the teleologist will have to admit that reasoning involves second-order thoughts. In order to explain the purportedly first-order character of reasoning, the teleologist will have to describe the aim in such a way that it does not involve a second-order attitude of the agent. One way to do this is to claim that attitudinal aims are constituted by the mechanisms that causally regulate the attitudes. This would make the teleological view a version of functionalism in the philosophy of mind—the view that attitudes are individuated in part by their characteristic inputs. The teleologist might also add that these attitudes were evolutionarily selected precisely because they were regulated by these causal mechanisms. This would allow them to apply the notion of proper function and corresponding notion of malfunction to these attitudes. However, the notion of proper function cannot explain the constraints on doxastic and practical reasoning. In general, an object or process’s proper function does not constrain how people use the object or engage in the process. Even if truth is the proper function of belief and initiating desirable action is the proper function of action, why can’t we ignore these facts in deliberating about what to believe or intend in the same way we can ignore the proper function of a kidney in determining how to use it in a medical experiment?

11. I lay out this objection in more detail with respect to the aim of belief in Shah (2003, 460–5).

12. Pamela Hieronymi’s view (2006, 2009), according to which to believe that $p$ is to affirmatively answer the question whether $p$, suffers from this problem as well: her view does not have the resources to explain why we can be caused to believe that $p$—i.e. answer affirmatively the question whether $p$—by non-evidential considerations even though we can’t take non-evidential considerations as our reasons for believing that $p$—affirmatively answering the question whether $p$. In fact according to our ordinary notion of answering a question, we can answer
a question on the basis of non-evidential reasons—e.g. a friend asks, “Do you think I’m fat?” But I do not see how there is anything in Hieronymi’s technical sense of the notion of answering a question that can explain these features of belief either.

13. I am excluding the possibility that you are “playing to lose.” A teleological account of chess would need to be modified to take account of this possibility.

14. The thoughts sketched in the rest of this section require a fuller treatment than I can provide here. I intend to develop them in future work.

15. See Boghossian (2008).

16. Nor can the general requirements on concept-possession help us to explain the desirability constraint on practical reasoning, since it is even more implausible that the very possibility of thinking requires obeying this constraint.

17. The reason that the norm of belief doesn’t constrain non-deliberative belief-forming processes is that these processes don’t operate through an agent’s conceptualization of them: the effectiveness of non-deliberative processes in producing beliefs doesn’t depend on their being conceived by the agent as exercises of her capacity to determine what to believe.

18. The objection I am about to present is similar to objections that have been made recently by Steglich-Petersen (2006, 506–507), Gluer and Wikforss (2009, 45–52), McHugh (forthcoming) and Sharadin (ms).

19. This is not to question whether reflection can play a role in reasoning. Reflection needn’t be epiphenomenal with respect to all aspects of reasoning in order to be irrelevant to the specifically evidential character of reasoning. In order for reflection to be effective it must obey the constraint; that much is shown by the failure of reflections on the desirability of a belief to enter into our reasoning. But reflection that does obey the evidential constraint needn’t also be responsible for that constraint.

20. In discussion, some people have said to me that they don’t think that non-reflective reasoning obeys the evidential constraint. If this were so, then I don’t think there would even be a temptation to think that the evidential character of reasoning can be explained with purely non-reflective elements.

21. For an excellent articulation of this explanatory strategy with respect to the category of action, see Ford (2011).

Bibliography


McHugh, Conor. (Forthcoming). ‘Normativism and Doxastic Deliberation,’ Analytic Philosophy.


Sharadin, Nate. (ms). ‘Transparency in Doxastic Deliberation.’