

# Knowing What One Wants

Krista Lawlor

Stanford University

Pre-print; please cite *Philosophy and Phenomenological Research* 79 (1): 47-75 (2009)

## 1. Introduction

Katherine, a young woman, stands by her son's crib, watching him sleep. "Have another"—she hears the words in her head. She is startled. Was she thinking about that? She supposes she was, but she wonders: was that a directive from some part of her? Or was it a question? Or was it something else?

When Katherine hears these words, *have another*, she realizes that she has been thinking on and off about the possibility. So, on hearing the words, she thinks perhaps she has already been deliberating, and here is the answer: *have another*. But then, if this is an answer, it does not feel decisive. And so she thinks perhaps she has *not* been deliberating, but merely entertaining a possibility without thinking squarely and focusedly about it—here then is not an answer, but rather "the question is being called." A decision of some importance must be taken, and it is time to think squarely about it.

What kind of question is being called? The question Katherine faces is not just the question of what her ends are to be, or of what she should all-things-considered pursue. She also faces the question, a distinct question, of what it is that she wants.

That there are two distinct questions here may be difficult to discern. The question of what one wants often blurs into the question of what *to* want. The reasons for this are complex: often, the question of what one wants does not even arise outside the sphere of practical

reflection. It is not until one asks oneself, "What end might I pursue here?" that the question of what one wants even takes shape.

Nonetheless, these are distinct problems for us as human agents: knowing what to want, and knowing what one wants, plain and simple. The question Katherine faces is not only about what her ends are to be, but also about whether an impulse counts as a desire, as something that might even compete as a candidate end.

So how does one know what one wants? It seems clear that one does not detect one's desires, through overt mental acts of scanning one's consciousness, say. Neither does one need to wait and see what one does, in order to have evidence of one's desire. In light of such facts, many philosophers have been led to argue that self-knowledge of desire owes to the presence of some kind of *constitutive* relation between the attitude known—the desire—and the reflective attitude involved in knowing about it.<sup>1</sup> As a consequence, these philosophers are committed to the idea that knowing what one wants is not a cognitive accomplishment: that is, knowing one's desire is not a matter of successfully finding out about desires one has, through cognitive effort; likewise, failing to know one's desire is not a matter of unsuccessfully trying to find out about desires one has, but is instead a matter of failing to make up one's mind, or failing to have the desire in the first place. So knowledge of one's desire, for these philosophers, is not a cognitive accomplishment.

I think that in many cases, knowing what one wants *is* a cognitive accomplishment. In what follows, I will argue that this makes better sense of our actual paths to self-knowledge.

How does one know what one wants, when there is no behavioral evidence one goes by, no inner desire-detector? I believe that we haven't paid enough attention, in trying to answer this question, to

how we actually seek and gain knowledge of what we want. We have not, for instance, let ourselves see the central role of imagination, in such characteristic acts as noted above ("calling the question"), in knowing what one wants.

My aim in this paper is to examine one route to self-knowledge of desire. In brief, I think we'll see that *inference from internal promptings* is a routine means by which we know what we want.<sup>ii</sup> In some cases, the internal promptings are simple sensations; in other cases, where desires are more complex, one's internal promptings may include imaged natural language sentences and visual images (which in turn may figure in specific kinds of imaginative rehearsal). One makes sense of such promptings by making an inference about their likely causes. Because the inference concerns causes, I will call this path to self-knowledge, "causal self-interpretation." In cases when one employs causal self-interpretation, knowing what one wants is a cognitive accomplishment.

The plan for the paper is as follows. In the next section, I briefly canvass three theories of knowing what one wants. A common theme in these theories is that self-knowledge of desire, like knowledge of other attitudes such as belief, is based on constitutive relations, and so is not in routine cases a cognitive accomplishment. I believe we have good reason to think constitution relations cannot tell the whole story of our self-knowledge, and that this is revealed when we have a closer look at what is involved in knowing what one wants. In section 3, I pay close attention to the experience of getting (and trying to get) self-knowledge of one's desires. This examination suggests that causal inference is a routine means by which we know what we want, and so self-knowledge is, in routine cases, a cognitive accomplishment. In the remainder of the paper, I'll consider why

philosophers have been reluctant to grant that inference is a route to self-knowledge.

## 2. Knowledge of desire: existing accounts

The kind of self-knowledge we are considering is knowledge of one's *conscious* desires. We might distinguish between desires that are conscious in the sense that they are objects of awareness, or *on one's mind*, and those that are conscious in the sense that there are no systematic barriers to their becoming objects of awareness. Desires that are conscious in either of these senses contrast with desires that are unconscious—viz. desires of which one is unaware, and for which there are systematic barriers to one's becoming aware. My claims will not be about unconscious desires. It is probably uncontroversial that we might only know about unconscious desires through some sort of inference. I wish to defend the idea that we often arrive at self-knowledge of conscious desires through inference. (Hereafter, in speaking of desires, I'll mean *conscious desires*.)

So, how do we know our conscious desires? There are many competing theories of our capacity to know the contents and states of our minds. Some of these theories focus on knowledge of sensations, experiences and other occurrent phenomenal mental events.<sup>iii</sup> Such theories have no direct application to the question of how one knows one's attitudes, like belief and desire. Alternatively, we have theories that focus on one's basic self-knowledge of the content of one's thoughts, or the nature of one's actions.<sup>iv</sup> Proponents of such views are careful to limit the scope of their accounts of self-knowledge. The accounts aren't meant to explain how one knows one's attitudes, such as beliefs and desires.<sup>v</sup>

There are three main accounts of self-knowledge that do focus on the attitudes. I'll call them the *self-constitution* account, the *rational supervenience* account, and the *rational-self-constitution* account.

(i) the self-constitution account

On the self-constitution account, self-interpretation or the act of attempting to know one's own mind, puts in place the very facts known. So, one's interpreting oneself as having a desire puts the desire in place. One might arrive at such a view by adopting an interpretivist view about the attitudes. The interpretivist holds that mental facts supervene on mental ascriptions.<sup>vi</sup> Interpretivism about the mental applies reflexively, to oneself, someone might hold.

Charles Taylor has such a view about the emotions, and, because of the connection he traces between emotions and desires, his is a view on which desires are constituted by self-interpretation.<sup>vii</sup>

In brief, Taylor argues that humans are essentially self-interpreting creatures. Our emotions in particular are shaped by the act of articulating them in language. Consequently, Taylor claims, self-interpretation constitutes the facts about one's emotions:<sup>viii</sup>

Our understanding of them or the interpretations we accept are constitutive of the emotion. The understanding helps shape the emotion. And that is why the latter cannot be considered a fully independent object... ("The concept of a person", p.101)

Now, on Taylor's view, one's desires are inextricably bound up with related emotional responses:

One way we can characterize our feelings is in terms of the things we want to do or have or experience, the

consummations desired: I want to eat, to sleep, to be at home again. These provide alternative ways of saying that I am hungry, sleepy, homesick...

But we can sometimes go deeper into our feelings, make more articulate what is involved in our desires, if we can express the [emotional] imports which underlie them and give them their point. When I am ashamed, I want to hide... I could try to say what I feel purely in terms of a consummation desired, by saying something like, 'I want my pride back.' But these terms would not be understood, either, without a grasp of the [emotional] import which is articulated in the whole vocabulary of shame, the shameful, dignity, pride, respect. ("Self-interpreting Animals" p. 56.)

Knowledge of one's desires and knowledge of one's emotions are mutually re-enforcing, on Taylor's view, and both emotions and desires are *constituted* (in part) by one's self-interpretive efforts. The basic point, for our purposes, is just this: on this account, knowing what one wants owes to the fact that self-interpretation—the act of attempting to know one's own mind—puts in place the very facts known.

(ii) The rational supervenience account

On the rational-supervenience account, by contrast, interpreting oneself as having a desire does not put the desire in place, but one enjoys a kind of guarantee of self-knowledge all the same. On the rational-supervenience theory, simply being rational (and having relevant conceptual resources) suffices to give one self-knowledge. This is because the mental state involved in knowing one's desire

*supervenies* on the desire known, or as we might say, *desire is self-intimating*, at least in rational creatures.

Sydney Shoemaker espouses the rational-supervenience theory.<sup>ix</sup> Shoemaker holds that for knowing one's desires, as well as one's beliefs, being rational is enough.<sup>x</sup> How could being rational suffice to give one self-knowledge? Consider the case of belief: Lacking an inner sense organ built to detect one's beliefs, one may instead simply ask oneself what is the case:

...one thing that will be true of a rational agent, whether her intentions are honest or dishonest, is that she will answer affirmatively to the question 'Do you believe that P?' if and only if she will answer affirmatively to the question 'Is it true that P?'<sup>xi</sup>

Shoemaker's idea is that, if one is rational, and has the relevant concepts, questions about what one believes are questions about the state of things believed.

Shoemaker goes on to claim that knowledge of one's desires is much like knowledge of one's beliefs. Shoemaker gestures at the sort of question a rational person might put to herself, in order to know what she wants:

Similar remarks apply to desire. The rational agent who wants X and has a normal mastery of language will, *ceteris paribus*, respond affirmatively to the question 'Shall I give you X?' and given her mastery of the concept of desire, she will respond affirmatively to the question 'Do you want X?' if she will respond affirmatively to the question 'Shall I give you X?'<sup>xii</sup>

Being rational and having the relevant concepts of desire, Shoemaker argues, suffices to answer. Consequently, knowledge about one's desires supervenes, in the rational agent, on the desire itself.<sup>xiii</sup>

(iii) the rational self-constitution account

Finally, on the rational-self-constitution account, one comes to have or change one's desires in virtue of rationally making up one's mind about what to desire. As with the pure self-constitution theory, self-interpretation (conceived now as a making up of one's mind) puts in place the very facts known, and so one also enjoys a kind of guarantee of self-knowledge. Failure to know one's desire is possible, but only in the sense of failing to make up one's mind about what one desires, not in the sense of having desires one does not know about.

Richard Moran has recently put forward a compelling defense of a variant of this view.<sup>xiv</sup> Let us begin with this core claim of Moran's: ordinarily, if one is rational, a *descriptive* question about what one's attitude is, in fact presents a *deliberative* question about what one's attitude should be, in light of one's reasons.<sup>xv</sup> Do I have a fear of being unemployed? In answering I weigh the reasons for and against the appropriateness of fear of this outcome, asking instead, "Is unemployment to be feared?" I make up my mind about this question, and the result is my *having* the attitude in question. I do fear it. If one is rational, self-ascribing an attitude involves answering a deliberative question about what one's attitude is to be, making up one's mind and thereby knowing the attitude.

The resulting account of self-knowledge of desire, setting aside many niceties, is this: If one is rational, faced with the question of *whether one desires that p*, one takes a "deliberative stance", just as one does towards one's beliefs, fears and regrets, answering instead



the question of *whether p is to be desired*. Ascribing a desire that p to oneself, then, in the special way one does when one knows it, is the result of seeing reasons for one's desire, and making up one's mind, "I want this", in light of those reasons.<sup>xvi</sup>

Of course, one does not always form attitudes after consciously entertaining one's reasons. Indeed, many times, one barely notices one's reasons for the attitudes one has.<sup>xvii</sup> So, in explicating Moran's view we must note this important feature: it isn't the case that in order to have a desire, one must actually make up one's mind about it, for every desire that one has; rather, the view is that, if one is to *know about a desire*, and self-ascribe it in an authoritative way, then one must make up one's mind about it; one must answer the *deliberative question* of whether p is to be desired. In the moment of knowing one's desire, one must see reasons for one's desire, and as it were, form it afresh in light of those reasons.

It is important to note that on Moran's view, if one cannot in this way avow one's desire that p, it is still possible both to have a desire that p, and know that one does; but one's knowledge in that case will not be "ordinary" self-knowledge.<sup>xviii</sup> It will be "third-personal" or "theoretical"—not the kind of knowledge we seek to understand in understanding self-knowledge. For example, taking the word of a trusted source (one's therapist say) one might come to self-ascribe a desire for revenge. But in such a case, not seeing reasons for the desire, one will be estranged from the desire, and one's knowledge of it won't be "ordinary" self-knowledge. According to Moran, it is only upon taking the deliberative stance, and making up one's mind, "I want this", that one enters a distinctively first-personal relation to one's desire. (A relation that, if not epistemic, is in any case the relation we care about in caring about self-knowledge.<sup>xix</sup>) Thus with Moran's rational-

self-constitution view we get both an account of self-knowledge, and a re-working of the idea of self-knowledge. Knowing what you want is understood not as an epistemic feat, but as a practical accomplishment.<sup>xx</sup>

Although these accounts differ substantially on important matters, the following is characteristic of each of them: On each of these accounts, one might not know every desire one has. And one can experience certain kinds of effort, in making the self-interpretations that yield self-knowledge. Nonetheless, in normal cases knowing one's desire is not a matter of successfully finding out about or discovering desires that one has, through cognitive effort. Rather, one either has the knowledge effortlessly (as on Shoemaker's view), in virtue of having the desire, or one's effort is of a practical kind (as on Moran's view). Likewise, failing to know one's desire is not a matter of unsuccessfully trying to find out about desires one has, and thereby suffering an epistemic failure; rather, failing to know one's desire is instead a matter of failing to make up one's mind, or failing to have the desire in the first place. In these ways, knowledge of one's desire is not a cognitive accomplishment, on any of these views. Rather, each of these accounts makes self-knowledge the product of constitutive relations among mental states.

We should be careful to note that each of these accounts will allow that one could come to know about a desire that one has through some alternative means—one might take one's behavior as evidence, or take the word of one's therapist, say, in self-ascribing a desire. But these must be deviant cases, not cases of "ordinary" self-knowledge.<sup>xxi</sup> Each of the foregoing views, then, concurs on this central claim: In normal cases, if one knows one's desire, that is the result of a

*constitutive, not a cognitive, relation* between the attitude known—the desire—and the reflective attitude involved in knowing about it.

I want to question this central claim of the foregoing views—I think that sometimes, in quite ordinary, psychologically normal cases, knowing what one wants rests on a *cognitive, not a constitutive, relation*. Knowing what one wants can be a cognitive accomplishment, in the sense that one finds out about an independently constituted object of knowledge (one's desire), through means that are routinely epistemic (namely, through inference). We see this, I think, when we focus on our experience of coming to know what we want.

### 3. Searching for self-knowledge

Too little attention has been paid to the experience of getting (and trying to get) self-knowledge, especially of one's desires. So I will spend some time with the phenomenon.

Perhaps our lack of attention owes to the fact that many times, knowing what one wants is easy. I know I want to be over this flu, I want to feel better. I know I want something cool to drink. How do I know these things? It seems so easy, there's not more to say than I just do know. Sometimes one's desires are for things so simple (that my head stop aching, that my thirst be relieved), the idea that the desire is self-intimating is very plausible.

As noted above, Shoemaker holds that it is in the nature of all desires, like beliefs, that they be self-intimating—to have them is to know about them. I'll be saying more about his account below, in section 5. First however, I want to note that some desires are not so easy to know about. And not just one's irrational or unconscious desires.<sup>xxii</sup> A person might wonder: Do I want to have my mother-in-law

live with me? Do I want to stay in this line of work, or pursue this other? Do I want to move to the city? One faces questions about conscious (or at least not deeply unconscious) desires routinely, and oftentimes answers about these desires are not immediately forthcoming. Second, I want to note that the line between such desires and simpler desires is of course not sharp—there is a wide spectrum here. And desires that are less about major life decisions and more about mundane matters can also fail to be self-intimating: Do I want to go for a run now? Do I want to buy these shoes? Is that lecture one I really want to hear?

Of course, it is open to the defender of self-intimation to insist that one does not have the relevant desire until one has the relevant self-knowledge. But this seems an inaccurate description of our experience of searching for self-knowledge. And, for the moment, it is our experience that I am interested in. Often one feels that one does in fact want or not want some particular thing, but cannot say immediately, all the while feeling that if one could only discover what it is that one wants, one would be better off. Katherine, for instance, may feel that there is a fact of the matter about her desire for another child—she really does or doesn't want one, and she feels that it is of great importance for her to find out which. Although this is a difficult case, on our spectrum of cases, it is the one I want to start with.

So how will Katherine find out what she wants? Now that the question has been called, Katherine starts noticing her experiences and thoughts. She catches herself imagining, remembering, and feeling a range of things. Putting away her son's now-too-small clothes, she finds herself lingering over the memory of how a newborn feels in one's arms. She notes an emotion that could be envy when an acquaintance

reveals her pregnancy. Such experiences may be enough to prompt Katherine to make a self-attribution that sticks. Saying "I want another child", she may feel a sense of ease or settledness.

More likely, however, her experiences do not prove decisive. Since she has for months felt wistfulness and wonder at her son's growing up, the aforementioned emotions and rememberings are routine, and haven't settled the question for her. More likely, then, if she makes a self-attribution at this point at all, it will have a provisional quality. She'll say to herself, "I think I want another," and see if it sticks—does she resist the self-attribution or not? And she might resist. Instead of a sense of ease, Katherine might find herself saying, "No, that's not it" or, "...still not sure," and she may find herself simply revisiting the same thoughts and imagings.

Alternatively, if the ascription doesn't stick, Katherine may cease her self-scrutiny, saying she can't really know her mind here: "I just don't know, it's so hard." Interestingly, agnosticism is not a stable position over time. Over time, if she never arrives at a self-ascription one way or another, Katherine will feel pressure to self-ascribe some desire or other, now under the demand that it be consistent with her inaction, for instance, "I wanted it, but not enough", or some such. Another alternative: Katherine may begin to wonder whether she really has a single desire one way or the other, or whether she isn't ambivalent, both wanting and not wanting. Ambivalence, like agnosticism, is not usually a stable self-ascription: "I'm just ambivalent, I want it and I don't", while compelling in the moment, also quickly leads to the demand, from self and others, "Well, which do you want *more*?"<sup>xxiii</sup>

If Katherine's provisional self-ascription doesn't settle the question, but she still feels there is an answer to the question that

has been called, she will continue to notice her fleeting imaginings, and reactions to them. Perhaps now with greater urgency, she'll sometimes shift from the passive experience of fleeting imaginings to a more active prompting of her imagination. Sotto voce, or in her mind's ear, she will prompt herself ("I keep thinking about this—why?"). She'll concentrate on her imaginings, trying to replay them, or to fill them out.

Some of Katherine's active self-questioning and prompting permits her to discover further imaginings. Some prompting results in Katherine finding herself re-playing or trying to replay specific compelling imaginings. This allows her to notice further features of the imaginings she has already had. For instance, she may notice that her imaginings are all about having a child just like the one she's already got. On noticing this, she might find herself with the fantasy of freezing him in time, and hanging on to his childhood. These images will suggest a different desire—it's not really a second child she wants, but this one all over again. So she'll say, "what I really want is to have *him* all over again."

Active prompting, as well as passive experience of imaginings can be repeated. That is, Katherine's imaginings may be *rehearsed*, both actively and passively: she may passively experience repetitions of imagined scenes, but she may also actively rehearse them, noting details and seeking to direct their content. If Katherine's mental life includes bouts of imaginings that are repeated or rehearsed, she may, even when not directly experiencing them, recall their flavor and content. For instance, she imagines a newborn, imagines its cries, and waking in the night for feedings. These imaginings she finds pleasing to linger over. Later in the day, she recalls the experience of imagining these things, and notes how she felt in rehearsing them.

It is worth pausing to note this fact: Given that desires can be complex, tracing their presence in one's thoughts, imaginings and fantasies can take a lot of reflection, and data to reflect upon. People tend to appreciate this fact about self-knowledge to greater and lesser degrees. Not everyone is as comfortable as Katherine is, in our example, letting her capacities for imaginative rehearsal have free rein. For some, not knowing what one wants is anxious-making, so the first self-ascription they make is the one they cling to, and try to live in accord with. For others, the opposite problem of self-knowledge arises—they never feel certain, in light of possible alternative self-ascriptions, in calling an impulse a desire of theirs. It could, after all, be another, more complex desire instead. (The dangers of reflecting too much or too little are the stuff of commonsense moral instruction, and recent social psychology.<sup>xxiv</sup>)

It would all be easier, we might find ourselves saying, if only one could know *that* one knows. Or, if only there were simple markers that attached to one's mental states, "desire", "mine." It doesn't work that way unfortunately. So what tells Katherine she is right, when she settles on one self-ascription or another? In the best case, after making a self-ascription, Katherine will experience a sense of ease, and she finds that her unbidden imaginings change: the inner voice that first called the question isn't heard from again; the pangs of envy or longing cease, and are replaced with new imaginings.

Now this might tempt us to say, knowing just is (or is the functional equivalent of) having made a self-attribution one can live with. However, being able to live with a self-attribution is no certain mark of having made the right self-attribution, or of knowing one's mind. The mark is confounded by the fact mentioned above—some people will find any way they can to live with an attribution, simply to have

the question settled. Another confounding fact is that living with a provisional self-ascription for a time is another means we have of finding out what we want. After one makes a provisional self-ascription, one may try the attitude on, both in imagination and in action. For instance, Katherine might say to herself, "I think I want another child", and even if that doesn't seem quite right, she may act in ways that fit the attitude (not donating old baby clothes to charity *just yet*). This is partly an effort to try on the desire. So, finding that one can, for a time anyway, live with a self-attribution is not a definitive mark of having identified what one wants.

While there may be no definitive mark of having identified what one wants, there are characteristic marks. If Katherine's mental life includes bouts of imaginings, passively repeated and actively rehearsed, Katherine may find the self-ascription, "I want another child" unavoidable, especially if the self-ascribed desire makes sense of why her mental life is so taken up with these particular imaginings. Second, this self-ascription may bring a new quality to her imaginings: now she feels anticipation, or hope, or exhilaration where before she felt no such future-directed emotions. She finds herself putting different questions to herself: not "do I want this?" but "how might we?" (We might even imagine that now, if she hears "have another", it sounds like an assertion in her mind's ear.) When this happens, Katherine's self-ascription of desire sticks. It finds no further challenges in her imaginative life, and begins to condition her further experiences, intentions and plans. There are then at least two characteristic marks of having identified a desire: (a) the self-ascribed desire is an explanation of the current shape of one's mental imagery and (b) the self-ascription effects characteristic changes in the further course of one's imaginative rehearsals.



There are no doubt many routes to self-knowledge of desire. Once we stop to think about it, no single means seems to be our exclusive means. One might for instance hear some unexpected bit of news, and be keenly disappointed, or overjoyed. Only then may one realize that one wanted the thing in question.<sup>xxv</sup> One might actively propose a course of action—heading to the beach, say—and test one’s reaction: a cool reaction suggests lack of desire.<sup>xxvi</sup> There are many ways to discover one’s wants. What I have tried to do here is not to provide a comprehensive account of all the ways one might come by knowledge of one’s desires, but to examine more closely the ins and outs of a particular, fairly common, avenue. It is the avenue we take when matters are not so simple, when our desires do not intimate themselves directly, or when we don’t stumble across them in the course of practical reasoning. In such a case, the avenue examined above is, I think, a routine way for the search for self-knowledge to play out.<sup>xxvii</sup>

In sum, it seems that causal self-interpretation, specifically, *inference from internal promptings*, is a routine means by which we know what we want. In some cases, the internal promptings are simple sensations; in other cases, where desires are more complex, one’s internal promptings may include imaged natural language sentences and visual images (which in turn may figure in specific kinds of imaginative rehearsal). And if inference is a means by which we know what we want, then we must disagree with the “no-accomplishment” views surveyed above, in their central claim. Sometimes, in quite ordinary, psychologically normal cases, knowing what one wants rests on a cognitive, not a constitutive, relation.

Of course, we’d like to go beyond how it seems to us that we know what we want, and have a fuller understanding of this route to self-

knowledge. In the next section, I'll begin to articulate some of the theoretical commitments and explanatory burdens we have, in taking seriously our experience about this particular route to self-knowledge of desire.

#### 4. Causal self-interpretation

My task in this section is to briefly sketch some of the theoretical commitments and explanatory burdens that we take on, in acknowledging that sometimes we know what we want through inference from internal promptings. I won't in the end be able to settle all the theoretical questions that arise, but I'll make a start and in some cases point to issues that will require further work.

Briefly, let's notice some features of our experience of trying to discover what we want: First, inner *imagings*, as we might call them, both of *natural language sentences* (i.e. inner speech), and of *images*, play a central role. (When I speak of imagings from here on, I have in mind both natural language sentences and images.)<sup>xxviii</sup> We enjoy both the passive experience of and the active prompting of such imagings. Further, although fleeting imagings may be important for self-knowledge, imagings that figure in *imaginative rehearsals* are also of importance in our search for self-knowledge. Such rehearsals involve occurrences of sequences of imagings, and both the passive experience and active prompting of these. Interestingly, we may recall or "replay" particular imagings that figure in such rehearsals so as to discern further features of them.<sup>xxix</sup>

Second, one's search for self-knowledge ends, perhaps only provisionally, with a *self-ascription*, which is the product of an inference about the source of these imagings. A self-ascription of a

desire is a judgment that one desires that *p*, which may or may not be publicly asserted; it has several features: (i) more than one self-ascription is made as part of a whole story about a *range* of one's imagings and emotional reactions to them, where (ii) the whole story is about the *best explanation* of one's imagings and reactions; (iii) nothing tells one for *certain* that one is correct in one's inference; however, provisional self-ascriptions may be tried on, imaginatively, and in action; and (iv) certain *characteristic changes* in one's imagings count as one's having 'settled the question.'

In brief, then, self-knowledge of desire is in routine cases a matter of self-interpretation of one's imagings, where that self-interpretation is a causal inference to the best explanation of one's inner life. (Again, not all knowledge of desire comes by the same route; but in central cases, this sort of inference is our route.)

Given this brief theoretical description, there are many things we need to say to fill out the story about how inferences about one's internal promptings can be a route to self-knowledge. In the remainder of this section, I consider just a few issues, focusing attention on the nature of the relevant internal promptings, especially the notions of imagination and imagings, and on the nature of the relevant inference.

First, there is a great deal that we will have to understand about the *imagination*—the faculty that allows us to entertain the relevant imaged sentences and images, sometimes even brief scenes, complete with characters and events. For instance, it seems that we enjoy both the passive experience of entertaining imagings that are unbidden, as well as the active direction of some of our imagings.<sup>xxx</sup> So we will have to understand better how people can take active control over their imaginative life, as they seek self-knowledge. I suspect

that we may be able to get a grip on the nature of the active direction of imaginative rehearsals as a species of *pretense*<sup>xxxii</sup>: One actively participates, prompting a flow of structured imaginative rehearsals, where the structure of one's imagings is settled by (tacit) principles of generation and systematic possibilities for discovery. This is clearly an area for further research.<sup>xxxiii</sup>

Another large issue: we will have to understand how desires are linked to relevant imagings (and other internal promptings). This too is a very large issue, and one I cannot hope to address here; I'll only lay out some of the territory.<sup>xxxiii</sup> First, the question of what links desire and imaging (and other internal promptings) is partly a metaphysical matter, concerning the very nature of desire. (Some of the questions that arise here: Is the relation causal, with the desire causing the relevant sensations and mental imagery? And if that is so, how does the desire do its work—must related beliefs also come into play? Or are the relevant imagings just part of what it is to have the desire?) The question of what links desire and imaging is also partly an epistemological matter. With an image, we are tempted to say, one undertakes to inform oneself of one's desires. Here, for instance, is Colin McGinn, speaking of the "daytime images that 'pop into the head'":

...the unconscious has thrown a pebble into the pool of consciousness, no doubt with a motive in mind...I might find that a recurrent image of someone is actually prompted by my *wanting* to think about that person—though this was not at first evident to me. 2004, p. 180, fn 18.

One issue that arises here is that saying as much threatens to generate a paradox: how one can inform oneself of something? Note that a similar paradox confronts us in making sense of self-deception: It seems that

self-deception is a matter of deceiving oneself, and of course deceit requires intention. But then, how can I deceive myself intentionally, since that in turn requires already knowing what I am supposedly trying to prevent myself from knowing? The paradox of *self-informing* runs in a parallel fashion: It seems that with an image one intentionally informs oneself of some desire (or other mental state). But, how can I inform myself intentionally, since that requires already knowing what I am supposedly trying to induce myself to know?

I want to briefly suggest a possible answer to this epistemological question, concerning the link between desire and imaging. What I say will be speculative, but I hope suggestive for further research into the relation between desire and imagination.

The question is, How might we understand the apparently self-informing quality of imagings? Here is one possible answer: Images make sense, and are engrossing, in ways that invite us to experience them as intended, as missives, we might say. But this is an illusion. Images are not missives sent from oneself to oneself, or from some agent within oneself to oneself. Why do they make sense, then, in ways that invite us to experience them this way? Here is a possible answer. Suppose that one's mind includes many systems, attentional, computational (comprised of many task specific computational modules), memorial, affective, visual, auditory, speech-producing and consuming, and so on. One also has self-monitoring systems, built to register bodily states and needs (thirst, hunger, temperature and so on). The information stored and processed continually by all these systems is quite enormous; but attention is limited, as are resources for forming intentions about what to do and what to think. In this competition for agential, person-level, resources, it gives an attitude, a need, or an emotion a decided edge if it can cause representations that have a

powerful pull on the attention. The attention is commanded by representations with auditory and visual aspects, and held by information that tells a coherent story. So if a desire can cause an image or images that catch and hold the attention, relevant systems of intention-formation may more readily engage in ways favorable from the perspective of the desire. Typically causing the right kind of images gets the desire a better chance of being fulfilled. The idea is that images are products of the self's desire (and other attitudes), designed as it were, by the desire itself or the systems responsible for the desire. A desire, (and related systems) has part of its functional role to grab the attention and command other of person-level resources (not just intention-forming resources, but computational resources, and so on). No agency needs to design the image so that it speaks just so about the desire. It is enough that having such effects in creature like us is a way for desires to get their way with us.

The foregoing is of course entirely speculative. I offer it here as one possible way of understanding the role of imagings in our mental lives, specifically as a way of answering a puzzle about what accounts for the self-informing quality of imagings. For our present purposes, we will leave the matter for further study.

Much more might be said about the imagination and its role in self-knowledge. But now I want to turn briefly to some issues about the sort of *inference* from imagings and other internal promptings that is involved in knowing what one wants.

First, and most important to note, the nature of the inference is not rationalizing, but causal. The images one reflects upon routinely have some narrative coherence, and speak of one's desires in an intelligible way. But when one reflects on one's imaginative rehearsals, one's questions are not of the form, What would make this

imagining the rational one to have, given my current situation? The questions one asks oneself take the form, What makes me keep thinking of this, Why do I keep coming back to that? When one answers this sort of question, by inferring one's desire as cause, one is not rationalizing one's behavior. Of course to make the relevant inference one must understand how imagings speak of one's attitudes. And this, like any other understanding, involves the grasp of some general principles (broadly speaking, principles of folk psychology). But inferring that a given desire is the cause of one's imaginative rehearsals does not require one to apply an intentional or rationalizing theory of the mental.

Note that to say as much is not to settle the metaphysical questions I sketched above. Whether or not one's desire *actually causes* one's imagings and other internal promptings is a separate question. The point I note here is that our inference is structured in such a way as to *suppose* that desires cause imagings and other internal promptings.

It is also worth stressing that one's inference is holistic, in the sense that one makes a target self-ascription in light of many assumptions about a range of one's imagings and reactions to them. Katherine self-ascribes the desire for another child, along with feelings of *envying* others their children, *fearing* that it might not work out, and so on. If envy and fear are not themselves compelling self-ascriptions, the ascription of desire may not stick. Also, forestalling a worry, let me add that the cognitive processes of inference can be very swift, barely rising to consciousness. I have described in elaborate detail various imagings and inferences, at the risk of making the work of inferring one's attitude seem ungainly and extremely laborious—nothing *ordinary* people could or would engage in.

But my description should not in this way distract us. Compare: a cognitive psychologist's description of the processing involved in negotiating a room full of furniture, were it complete, would look incredibly complex and difficult to perform, seemingly nothing *ordinary* people could or would engage in.

We have now taste of the work that lies ahead, and to give such a taste was my main aim in this section. Let's recap where we have come to. In section 2, we saw three accounts of self-knowledge of desire, on which routine or ordinary self-knowledge is never a cognitive accomplishment; in section 3 we looked hard at the experience of coming to know what one wants, which suggested that making an inference from internal promptings (including imagings) is in fact an ordinary route to self-knowledge. In section 4, we explored briefly some of our theoretical commitments and explanatory burdens, if we take seriously the idea that causal self-interpretation is a route to self-knowledge.

In the final section of the paper, I want to consider the question, Why have philosophers been reluctant to grant that inference is a route to self-knowledge? This is a pressing question for us, having come this far. If reflection on our experience tells us that inference (from internal promptings of one kind or another) is a routine means of knowing what we want, why do philosophers routinely deny that inference from evidence has anything to do with self-knowledge?

##### 5. What's so bad about inference

It is commonly held that while inference might play a very important role in knowing about other people's attitudes, it doesn't play much of



a role at all in knowing about one's own. Here is one typical expression of the idea:

In the case of others, I have no choice but to *infer* what they think from observations about what they do or say. *In my own case, by contrast, inference is neither required, nor relevant. Normally, I know what I think—what I believe, desire, hope or expect—without appeal to supplementary evidence.* Even where such evidence is available, I do not consult it. I know what I think directly.<sup>xxxiv</sup>

Most acknowledge that the claim here needs some restriction: inference from the evidence (of one's actions, or internal promptings) might play a limited role in self-knowledge, when the attitude one knows about is unconscious, say, or an attitude one held in the past, or is in some other way an attitude one holds at some remove from one's current consciousness. The commonly held claim is that, *barring such cases—where one's relation to one's attitude is arguably more like that of a third-person—if one is rational, one knows one's attitudes without inference from evidence.*

I think that this claim is mistaken, even in its restricted form. Our examination of our actual routes to self-knowledge of desire tells directly against this idea. But this raises a question: why do philosophers routinely deny that inference from evidence has anything to do with self-knowledge? If our experience tells us that it is routine to infer our desires, why is the very idea so roundly dismissed?<sup>xxxv</sup> In this final section, I'll consider some possible sources of resistance.

First, perhaps philosophers have simply overlooked that the relevant "supplementary evidence" from which one makes one's inference can take the form of *internal* goings-on (imagings, sensations and so

on), and have fixed instead on behavioral evidence, of the kind we must use in attributing desires to others. Inference from evidence, when evidence is restricted to behavior, does in fact seem irrelevant in normal cases of knowing what one wants (and other mental states and attitudes). One certainly does not have to wait to see one's hand reaching for the glass to know one wants a drink. Perhaps then, there's been a simple failure to see that, although we use *different data*, in third-person and first-person ascription of attitudes, we might use the *same means*, namely inference. The problem with this hypothesis, however, is that we don't find a lot of evidence of philosophers overlooking internal promptings as a possible source of data. Paul Boghossian (the source of the quotation above) for instance, acknowledges internal promptings as potential data for inferences.<sup>xxxvi</sup> So no such simple hypothesis explains the resistance to the idea that inference is a routine route to knowing what one wants.

Here is another, more likely, source of resistance. As our quotation suggests, as some would have it, it just seems that one makes no inferences, but rather knows *immediately*—where that is to say *non-inferentially*—about one's desires. The claim then, is that the first-person phenomenology of seeking self-knowledge turns up no inferences.

In response, I think we can safely agree that this may be how it seems sometimes. Especially in knowing about sensations this is true. One feels a pain, and one just knows one is in pain. One feels prickly heat on one's hand, and one just knows it. Inference here doesn't seem to have any place. And even in the case of knowing some of one's desires, as I noted above, knowing what one wants can seem immediate, where by that we mean, not based on inference from anything more basic. One knows one wants to be over the flu, or that one wants something cool to drink. Sometimes it seems, there's not more to say about *how it*

seems one knows these things other than, one just knows. But other times, there is more to say about how it seems one knows what one wants. In the cases we have considered above, it's pretty clear that it seems to one that one makes sense of the items and events in one's inner life by seeing them as traces of one's desires.<sup>xxxvii</sup>

Beyond the phenomenology, or how it seems to one that one knows what one wants, philosophers express resistance to inference on more reflective grounds. A closer look at some no-accomplishment theorists reveals a couple of distinct arguments against the role of inference in self-knowledge. Before closing I want to look at two such arguments.

First, an argument from Shoemaker. On Shoemaker's rational supervenience account, recall, knowledge of one's desires is much like knowledge of one's beliefs, in that there is a characteristic question a rational person might put to herself, in order to know what she wants:

The rational agent who wants X and has a normal mastery of language will, *ceteris paribus*, respond affirmatively to the question 'Shall I give you X?' and given her mastery of the concept of desire, she will respond affirmatively to the question 'Do you want X?' if she will respond affirmatively to the question 'Shall I give you X?'

Being rational and having the relevant concepts of desire, Shoemaker argues, *suffices* to answer; so inference from internal promptings is just irrelevant to knowing what one wants. In other words, Shoemaker's claim here can be understood to be that since we *can employ exclusively non-inferential means* in knowing what we want, inference is irrelevant to self-knowledge.

In response to this claim, it is not enough to simply appeal to our cases, and to the fact that we sometimes do employ inference in knowing what we want. Shoemaker's claim is consistent with this fact. All he needs, in order to discredit the importance of inference as a routine means of self-knowledge, is to show that it is extraneous, something over and above what actually suffices for the rational person to know her desires. In response, then, I take a different tack: I question whether Shoemaker's own account of self-knowledge is as free of inference as he suggests. Let's consider for a moment how the question "Shall I give you X?" gets answered. (Of course, we might wonder about whether Shoemaker identifies the right form of question to ask oneself, in order to prompt self-knowledge of desire. There are many things one might want, that cannot be things one would have given to one. But leave these points aside.) The issue is, Can this (or any likely candidate) question concerning desires really get answered in a parallel way to the question concerning beliefs? That is to say, can it get answered immediately, or just in virtue of one's being rational? I don't think so.

Rationality and conceptual resources are not enough, I suggest, to provide an answer to the question (whatever the question might be) in the desire case. Suppose Katherine is asked, "Shall I give you another child?" (Supposing for argument's sake that this were a suitable question, asked by a suitable person.) Are we to suppose that Katherine simply knows the answer to this question? That is hardly likely. In fact, the question will be answered, if it is answered at all, as a result of Katherine's attending to and interpreting the flow of internal imagings it provokes, in the way we have considered above.

In sum, if Shoemaker were to follow through on his own suggestion about the case of desire, I believe he would see that more than mere

rationality and conceptual competence are needed for self-knowledge of desires. Some kind of sensitivity to one's imaginings, and inference from them, is also routinely needed.

So far, then, we have seen little reason to discriminate against inference as a means of knowing what one wants. I will consider one last argument against inference, this one from Richard Moran. Moran's position is fairly complex, so I'll lay out his position here with some care.

Moran encourages us to look closer at our intuition that self-knowledge is immediate. He suggests that what we find is not a claim against inference per se, but a claim against any route to self-knowledge that would purport to deliver knowledge but leave one alienated, estranged from, or unable to endorse the attitude known. Self-knowledge has two aspects, epistemic and practical, Moran argues:

The dimension of endorsement is what expresses itself in one aspect of first-person authority, where it concerns the authority of the person to make up his mind, change his mind, endorse some attitude or disavow it. This is a form of authority tied to the presuppositions of rational agency and is different in kind from the more purely epistemic authority that may attach to the special immediacy of the person's access to his mental life. (92)

And while these two aspects of self-knowledge are "not unrelated", no amount of one can produce the other:

...alienation...would remain possible on even the most generous epistemology for self-knowledge, so long as it was construed purely theoretically [i.e. as resting on an empirical means, like inference]. We may allow any manner of inner events of consciousness, any exclusivity and privacy, any degree of

privilege and special reliability, and their combination would not add up to the ordinary capacity for self-knowledge. (93)

Someone might think that Moran's central claim here is irenic, merely expressing a call not to lose track of the non-epistemic, practical aspect of endorsement or commitment that comes along with knowing one's mind.<sup>xxxviii</sup> But Moran's position is not irenic: for on Moran's view, "ordinary self-knowledge" requires more than mere epistemic authority; it requires the dimension of non-epistemic, practical authority. So Moran is committed to the idea that inference alone (or any other empirical means) could in no case be enough to produce ordinary self-knowledge. Inference might produce a bit of knowledge with a high degree of "purely epistemic authority", but inference cannot produce endorsement of or commitment to the attitude one thereby knows about. One might know one's attitude, but be alienated from it, if one knows it by inference, or any other empirical means. And for that reason, inference cannot be one's route to ordinary self-knowledge.

Moran illustrates these claims using the case of *feeling betrayed*, and finding out about one's feeling by *being told about it*, but the moral Moran would draw is supposed to hold for a wide range of attitudes, like desire, and a wide range of "theoretical" or "empirical" means of finding out:

The person might be told of her feelings of betrayal, and she may not doubt this. But without her capacity to endorse or withhold endorsement from that attitude, and without the exercise of that capacity making a difference to what she feels, this information may as well be about some other person, or about the voices in her head. (93)

Being in a position to endorse the attitude one knows about is necessary for having ordinary self-knowledge, according to Moran. Since

one might infer one's desire or feeling of betrayal, and not thereby find oneself in a position to endorse it, the resulting knowledge falls short of ordinary self-knowledge—indeed, it “may as well be about some other person.” So, according to Moran, inference alone (or any other empirical means) cannot produce ordinary self-knowledge of desire (or other attitudes).

In responding to Moran's challenge here, I think we must question the idea that *endorsement* is required if one is to know one's desires.<sup>xxxix</sup> Making endorsement a requirement on self-knowledge of desire has the effect of identifying two kinds of self-knowledge that should be kept apart.

Briefly, recall, that on Moran's view, endorsement comes with seeing reasons for one's attitude, and this is precisely what one does when one takes up a “deliberative stance” toward the question “whether I desire p”, asking instead the question “whether p is to be desired.” Ascribing a desire that p to oneself, in the special way one does when one knows it, Moran claims, is the result of seeing reasons for the desire, and making up one's mind, “I want this”, in light of those reasons. In other words, to know *what one wants*, one deliberates about *what to want*, in light of one's reasons. The question of what one just plain does want, regardless of one's reasons, is replaced, on Moran's view, by the question of what to want. Here I think, we have two kinds of self-knowledge that need to be kept apart. Knowing what one wants—however unreasonable one's wants might be—is an independent and legitimate goal of self-knowledge.<sup>x1</sup>

There are compelling reasons in favor of distinguishing these two kinds of self-knowledge: after all, one can explicitly set aside the question of whether a thing is feasible, or whether it would make other ends impossible, and so on, and intelligibly ask oneself, Do I really

want this? One can ask oneself, Am I moved to think about this because I think it's good for someone I care about, or because of social pressures, and so on, or am I moved to think about this because it's what *I* want? Second, one might decide that, although one does want the thing in question, it's just not the thing to do. Judgments about attainability or suitability of an object may have an effect on what one can desire, to be sure. But there is nonetheless conceptual space to prize apart desire and desirability. Likewise, one might realize that although one is actively pursuing an end, treating it as something to want, it is not what one wants. Finally, knowing what one wants is sometimes a necessary first step in knowing what to want, even if this is not always the order in which we proceed.

These considerations suggest that the two questions—what I want, and what to want—should be distinguished. Knowing what one wants is not simply or only a matter of knowing what to want. For this reason, we should not accept Moran's demand that any route to self-knowledge should produce endorsement of the attitude known, at least in the case of desire.<sup>xli</sup> That demand suggests that knowing what one wants is a matter of knowing what to want, or what might be rationally endorsed as one's desire. But these are two different kinds of self-knowledge.

It seems, then, that we have no reason to discriminate against inference as a route to knowing what one wants.

## 7. Conclusion

I have argued that one's self-knowledge of desire owes to one's ability to sleuth out desires as causes of certain characteristic kinds of mental imagings that fill the stream of conscious life. Sometimes one is passive before one's unbidden imagings. Sometimes one actively



participates, prompting a flow of structured imaginative rehearsals. In interpreting the import of these rehearsals, one does a special sort of interpretive work: one doesn't use rationalizing psychology; one interprets one's imagings for their likely cause, for the trace of desire they carry. One engages in causal self-interpretation.

It's worth noting that causal self-interpretation is also a means of self-knowledge of other mental states than desire. When we attend to ordinary experience, as many have noted, we find a rich field of discrete elements. Many figures, scenes, and imaged phrases pass through one's mind in any given stretch of wakefulness. And often, almost without noticing as much, one interprets these imagings (sometimes incorrectly) as being caused by specific mental states or attitudes: in the midst of one's daily activity, one pictures lemons, and where they're to be found at the store, *because* (one thinks) one intends to buy some. One suddenly imagines the house and fields where one grew up, *because* (one thinks) the summer sun just now reminds one of the summer sun there. One hears the words "Point Reyes" in one's inner ear, *because* (one thinks) one wants to see that part of California again. Assigning a cause is part and parcel of identifying the mental attitude that underlies the imaging. That is to say, assigning a cause is a way of identifying one's imaging as the product of an intending, a remembering, or a desiring, to a specific effect. Assigning a cause is knowing one's mind.<sup>xlii</sup> One's interpretations run the risk of being outright wrong, of course. When they're right, they count as significant cognitive accomplishments.<sup>xliii</sup>

## Bibliography

- Bar-On, Dorit. *Speaking My Mind*, Oxford: Oxford University Press, 2004.
- Bratman, Michael. "Two Problems about Human Agency", *Proceedings of the Aristotelian Society*, vol.101, no.3, April 2001, pp.309-326.
- Burge, Tyler. "Content Preservation", *Philosophical Review*, vol.102, no.4. 457-488.
- Peter Carruthers. *Language, Thought and Consciousness*, Cambridge: Cambridge University Press, 1996.
- Chisholm, Roderick. *The First Person*, 1981.
- Currie, G. and Ravenscroft, I. *Recreative Minds*, Oxford: Oxford University Press, 2002.
- Darwall, Stephen. "Self-Deception, Autonomy and Moral Constitution", in McLaughlin and Rorty (eds.) *Perspectives on Self-Deception*, Berkeley: University of California Press, 1988.
- Davidson, Donald. "Deception and Division" in LePore and McLaughlin, *Actions and Events*, 1985
- Dretske, Fred. "Introspection", *Proceedings of the Aristotelian Society* 94: 263-278
- Gallois, Andre. *World Without the Mind Within: An Essay on First Person Authority*. Cambridge: Cambridge University Press, 1996.
- Gendler, Tamar. "The Puzzle of Imaginative Resistance" *Journal of Philosophy*, 97:55-81, 2000.
- Hampton, Jean. *The Authority of Reason*, Cambridge: Cambridge University Press, 1998.
- Kim, J. Brandt, R. and Morgenbesser, S. "Wants As Explanations of Actions", *Journal of Philosophy*, vol. 60, no. 15.

- Kolnai, Aurel. "Deliberation is of Ends" in *Ethics, Value and Reality, Selected Papers*, Williams, B, Wiggins, D (eds.) Indianapolis, Hackett Publishing, 1978.
- Kosslyn, Stephen. *Image and Brain*, Cambridge: MIT Press, 1996.
- Lawlor, Krista. "Knowing Beliefs, Seeking Causes" *American Imago*, vol 65, no.3, pp. 335-356, 2008.
- McDowell, John. *Mind and World*, Cambridge: Harvard University Press, 1984
- McGeer, Victoria. "Is 'Self-Knowledge' an Empirical Problem? Renegotiating the Space of Philosophical Explanation", *Journal of Philosophy* 93, no. 10, 1996.
- McGinn, Colin. *Mindsight*. Cambridge: Harvard University Press, 2004.
- Moran, Richard. "Interpretation Theory and the First-Person", *The Philosophical Quarterly*, vol.44, no 175, pp.154-173.
- , *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press, 2001
- Sartre, Jean-Paul. *The Psychology of Imagination*. Washington Square Press, 1966.
- Shoemaker, Sydney. *The First Person Perspective and Other Essays*, Cambridge: Cambridge University Press, 1996
- Taylor, Charles. *Human Agency and Language: Philosophical Papers* vol.1, Cambridge: Cambridge University Press, 1985.
- Velleman, David. *Practical Reflection*, Princeton: Princeton University Press, 1985.
- "On the Aim of Belief" in *The Possibility of Practical Reason*, Oxford: Oxford University Press, 2000.

Walton, Kendall. *Mimesis as Make-Believe*, Cambridge: Harvard University Press, 1990.

Williams, Bernard. "Internal and External Reasons" in *Moral Luck*, Cambridge: Cambridge University Press, 1981.

Wollheim, Richard. "On Freudian Unconscious", Presidential Address, *Proceedings and Addresses*, APA November, 2003, v.7, issue 2.

---

<sup>i</sup> See for instance, Shoemaker's "On Knowing One's Own Mind" in *The First Person Perspective and Other Essays*, Cambridge: Cambridge University Press, 1996; Victoria McGeer "Is 'Self-Knowledge' an Empirical Problem? Renegotiating the Space of Philosophical Explanation"." *Journal of Philosophy* 93, no. 10 1996; Andre Gallois, *World Without the Mind Within: An Essay on First Person Authority*. Cambridge: Cambridge University Press, 1996. Richard Moran's view in *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press, 2001 is related, but not a straightforward constitution view. I'll say more about his theory below.

<sup>ii</sup> I don't intend to argue that this is our only route to self-knowledge. In fact I believe we have many epistemically distinct means by which we come to know what we want.

<sup>iii</sup> For instance, Chisholm's direct perception account, on which some psychological properties are "self-presenting" (*The First Person*, 1981). See also McDowell's modified "inner-sense" account of phenomenal states (*Mind and World*, Cambridge: Harvard University Press, 1984).

<sup>iv</sup> For the former, see for instance Tyler Burge "Content Preservation", *Philosophical Review*, vol.102, no.4. 457-488. For the

---

latter, see for instance Velleman's *Practical Reflection*.

<sup>v</sup> But see Dorit Bar-On (*Speaking My Mind*, Oxford: Oxford University Press, 2004) for an attempt to broaden the scope of basic self-knowledge of thought contents, or at least the sort of authority that attends avowals of such self-knowledge, to other non-basic cases. It should be noted that her "neo-expressivist" theory is an account only of the special authority of self-ascriptions, and as such is consistent with many accounts of self-knowledge—even deflationary accounts of self-knowledge, on which no special epistemic accomplishment backs the authority of one's claims. Since it is not a theory of self-knowledge *per se*, I don't discuss neo-expressivism in this paper.

<sup>vi</sup> See Moran's "Interpretation Theory and the First-Person" (*The Philosophical Quarterly*, vol.44, no 175, pp.154-173.) for interesting discussion of the application of interpretivism to one's own mental life.

<sup>vii</sup> Charles Taylor *Human Agency and Language*, Cambridge: Cambridge University Press, 1985.

<sup>viii</sup> The range of "self-shaping recognitions" is broad: "I deeply love her; or, I am jealous; or, I really don't care." ("Self-interpreting Animals" in *Human Agency and Language: Philosophical Papers vol.1*, Cambridge: Cambridge University Press). Note that Taylor registers the following caveat:

...this is not to say that we can change our emotions arbitrarily by applying different names to them. We are not talking about a process which could be arbitrarily undertaken. It is not just applying the name that counts, but coming to 'see-feel' that this is the right

---

description; this is what makes the difference...that language is constitutive of our ...emotions says nothing about the order of causation. p.71

<sup>ix</sup> See especially Shoemaker's "On Knowing One's Own Mind" in *The First Person Perspective and Other Essays*, Cambridge: Cambridge University Press, 1996. (And for talk of "supervenience", see p. 34.)

\* That is, being rational, plus having the relevant conceptual capacities, and linguistic mastery, suffices for self-knowledge.

<sup>xi</sup> (1996 p. 237.) Shoemaker here follows Gareth Evans, Wittgenstein, and others, in noting the tight connection between belief ascription and fact stating. Note that Shoemaker has put forward, in a number of papers over the years, several arguments for the supervenience of self-knowledge on rationality. Some of these arguments are quite complicated, involving the conceptual possibility of "self-blindness", and the role of the attitudes in practical reason. Trying to summarize all his arguments here is too large a project. What I try to do is distill a central theme, using Shoemaker's own admittedly overly simple tests.

<sup>xii</sup> (1996, p.237) It is interesting to note that in an earlier paper, Shoemaker despaired of finding any general purpose test:

...there seems to be no formula for satisfying the request, 'Tell me some of your desires' that is comparable with the one I suggested for satisfying the request 'Tell me some of your beliefs', namely treating this as equivalent to 'Tell me some things that are true.' 1996, p.46

---

He also earlier seemed to despair of providing an argument about the supervenience of self-knowledge on rationality, when it comes to desires, because of the notorious human capacity to see desire and judgments of desirability pull apart (1996 pp. 47-48).

<sup>xiii</sup> (1996, p.34). As Shoemaker himself notes, all such question/answer pairs establish is a connection between giving *expression* to beliefs and desires and the ability to *self-ascribe* them. Nonetheless, the question/answer test just noted suggests the broad outline of the argumentative strategy that Shoemaker pursues in his various arguments. The strategy is to claim that mere attention to abstract facts about the evidential basis for one's attitudes (and/or how one should give evidence to others in light of one's attitudes), of a kind we expect rational agents to enjoy, constrains one's self-ascriptions in such a way that one's self-ascriptions can and must count as knowledgeable.

<sup>xiv</sup> See Moran (2001), and also McGeer (1996), Gallois (1996).

<sup>xv</sup> Moran (2001, p58ff).

<sup>xvi</sup> As Moran says: "In characterizing the two sorts of questions one may direct toward one's state of mind, the term 'deliberative' is best seen at this point in contrast to 'theoretical,' the primary point being to mark the difference between that inquiry which terminates in a true description of my state, and one which terminates in the formation or endorsement of an attitude. And so to speak of the person's role in forming his attitudes is not to invoke a kind of willful or wishful capacity for self-creation. A person adopts this role insofar as he can answer questions of the sort 'What am I to believe here?' and thereby come to believe something, or answer a

---

question of the form 'Is this what I really want?' In terms of considerations of what is worth wanting..." (2001, p63).

<sup>xvii</sup> As Moran notes: "This is not to say that one normally arrives at one's beliefs (let alone one's fears or regrets) through some explicit process of deliberation. Rather, what is essential in all these cases is that there is logical room for such a question, about regret as much as about belief, and that the actual fear or regret one feels is *answerable* to such considerations (2001, 63).

<sup>xviii</sup> See especially p.67ff, and p.91ff.

<sup>xix</sup> For further statements about the value of deliberative self-knowledge, see his papers "The Authority of Self-Consciousness." *Philosophical Topics* 26, no. 1 (1999): 179-200 and "Making up Your Mind: Self-Interpretation and Self-Constitution." *Ratio* 1 (1988): 135-51.

<sup>xx</sup> The epistemic and practical feats are often related, Moran argues. See (2001, p93ff)

<sup>xxi</sup> I should note that Taylor is not so far as I know explicitly committed to this claim, but it is an implicit commitment of his view, and it is an explicit claim in Moran. Likewise Shoemaker suggests that only unconscious (and likely irrational) attitudes will be known through cognitive effort.

<sup>xxii</sup> As noted above I leave unconscious desires aside in this paper, however I believe that the way Freud suggests we come by self-knowledge of unconscious desires (namely, "working through") has an analog in the everyday means of knowing one's conscious desires—namely, inference from internal promptings. Wollheim draws a connection between the working through and inference about the product of one's



---

imagination in his "On Freudian Unconscious", Presidential Address, *Proceedings and Addresses*, APA November, 2003, v.7, issue 2.

xxiii That we resist ascriptions of agnotisicism and ambivalence is an important fact about us, deserving of explanation, but I won't try here.

xxiv See for instance Timothy Wilson's *Strangers to Ourselves*, The Belnap Press of Harvard University Press, 2002, for the possible negative consequences of reflecting too much on one's reasons.

xxv See Kim, Brandt, and Morgenbesser, "Wants As Explanations of Actions", *Journal of Philosophy*, vol. 60, no. 15. p.429

xxvi Though of course it may instead suggest a judgment about feasibility, or about whether going to the beach is something one *should* want, given the other things one can or must do. And sometimes, feasibility considerations simply trump, so one needn't go further with trying to sort out whether it's a matter of *really wanting to go but can't*, or just plain *can't*.

xxvii Moreover, I think that the case of desire is not unique. Something of the same sort goes on in knowing about other of one's attitudes, like fear and resentment, and possibly also belief (see my "Knowing what one believes" manuscript).

xxviii See Peter Carruthers, *Language, Thought and Consciousness*, Cambridge: Cambridge University Press, 1996, for recent work on the nature of imaging and its relation to conscious thought.

xxix Note Bernard Williams claims that elements in one's motivational set can be affected, *added* and *removed*, by imaginative reflection, even if they are unmotivated or intrinsic desires. ("Internal and External Reasons" in *Moral Luck*, Cambridge: Cambridge

---

University Press, 1981, p. 110.) This is a claim deserving of further empirical study.

<sup>xxx</sup> Despite our capacities for active direction of our imaginations, I believe we ultimately have to reject any voluntarism about the imagination that would have us in control of all our imagings. But this is an argument for another day. For defense of voluntarism, see Jean-Paul Sartre, *The Psychology of Imagination*. Washington Square Press, 1966, and Colin McGinn, *Mindsight*, Cambridge: Harvard University Press, 2004.

<sup>xxxi</sup> Unfortunately my remark is at this time largely speculative. About pretense and imaginative experience, see Walton *Mimesis as Make-Believe*, Cambridge: Harvard University Press, 1990. Recent work has been done on pretense and belief, and on the experience of affect within pretenses. See Velleman "On the Aim of Belief" in *The Possibility of Practical Reason*, Oxford: Oxford University Press, 2000; Gendler "The Puzzle of Imaginative Resistance" *Journal of Philosophy*, 97:55-81, 2000 Currie, G. and Ravenscroft, I. *Recreative Minds*, Oxford: Oxford University Press, 2002. I am not aware of work on pretense as a route to self-knowledge, especially of desire, but the topic seems directly related to these current investigations.

<sup>xxxii</sup> And it's an area where empirical research will be relevant. For a fine introduction to work on the imagination see Paul Harris, *The Work of the Imagination*, Blackwell, 2000.

<sup>xxxiii</sup> Some of the relevant territory will require further empirical research as well, but I will mention only philosophical questions here.

<sup>xxxiv</sup> Paul Boghossian (1989, p.7 emphasis added). Here, too, is Richard Moran: "The type of access we ordinarily take ourselves to have

---

here is special in at least two basic ways. First, a person can know of his belief or feeling without observing his behavior, or indeed *without appealing to evidence of any kind at all*. And second, rather than this nonreliance on evidence casting doubt on the reliability of such reports, judgments made in this way seem to enjoy a particular epistemic privilege not accorded corresponding third-person judgments that do base themselves on evidence." (2001, p.10 emphasis added)

xxxv For a notable exception, see Goldman, "The psychology of folk psychology", Behavioral and Brain Sciences (1993) v16, 15-28.

xxxvi As does Moran. Boghossian's own argument against inference as a source of self-knowledge, by the way, is limited in scope, taking aim only at the claim that *all* self-knowledge is inferential; this is obviously not the claim I am here defending.

xxxvii It is important here to remark on some ways the first-person phenomenology of our efforts to know our own minds can also be misleading. I think it is clear that we are first-hand familiar with making causal interpretations, or inferences to the best explanation, about why we are experiencing certain imagings or sensations. (As our cases above illustrate.) But we don't always readily describe our experience of coming to know our minds in these terms. Even when one infers one's desire from various internal promptings, if one only casually attends to one's experience, it may seem to one that one knows one's desire immediately (without inference). How do I know I want to go for a run? If asked, I may say, "I just do!", even if what I have done is move from my experience of various mental promptings (stirring sensations, imaged scenes of a path) by inference to an ascription of the desire. Why then do I report my desire as immediately known? It seems likely that this is because, once the desire is ascribed, these

---

experienced imagings are experienced as so intimately bound up with *having* the desire that they aren't experienced as ever having been independent symptoms, by means of which the self-ascription was made.

xxxviii Clearly, Moran is making no charge against inference or any other "theoretical" or "empirical" means of finding out what one desires, as a legitimate route to securing "epistemic authority" (2001, 92).

xxxix I want to stress that were Moran's claim merely to the effect that self-knowledge has practical implications, and that *to know one's mind is thereby to acquire various practical commitments*, responsibilities and so on, I think we should all find it possible to agree. One might hold that self-knowledge brings commitments and responsibilities in its wake, even if one also holds, as I do, that self-knowledge routinely rests upon inference. Moran's claim is not that one acquires practical commitments to attitudes one knows oneself to have; rather the claim is that knowing one's mind comes with taking up the relevant practical commitments and endorsements.

x<sup>1</sup> Of course, although the question of *what one wants* and the question of *what to want* are distinct, there are many reasons why we might confuse them. We might confuse the two kinds of self-knowledge because one often faces the two questions—what do I want, and what should I want—at the same time. Also, questions about what one wants sometimes need filling out, and often one doesn't so much as know how to fill out the question, without some prompting from a *practical demand*. Also, we often don't face questions about what we want until we wonder about what to want, or what end to aim at. This fact marks what might seem a surprising reversal of the expected order of self-discovery. Often the question of what one's ends are to be is not

---

settled by prior knowledge of what one wants, but by one's opportunities. As Aurel Kolnai notes:

We deliberate whether we should adopt a suggested purpose—e.g. that of visiting a city, having received an invitation—or in the variant case, deliberate about possible ends for a given mean. Of course we do not form purposes out of nothing, without any conditioning elements whatever in our pre-existing emotive structure and horizon of knowledge. However, for all the constants in our mental and affective outlook which make us receptive to some kinds of stimuli and unresponsive [sic] to others, our actual purpose-formation is largely contingent on occasions and suggestive influence which happen to cross our path. (1978, p.51)

<sup>xli</sup> Moran begins his exploration with the case of belief (and relatedly, later in his study, intention). Belief is, I suggest, a very special attitude, and it is easy to extrapolate incorrectly if one uses it as one's initial fixed point. In the case of belief, it is compelling to suppose that there exists a demand for what Moran calls *transparency*, where by that he means seeing reasons for one's belief in self-ascribing it.

<sup>xlii</sup> I have argued elsewhere that such causal self-interpretation is also a central means by which one knows one's beliefs. See Lawlor (2008).

<sup>xliii</sup> Thanks to Laura Schroeter and David Velleman for insightful comments. I am also grateful to the American Council of Learned Societies for a Charles Ryskamp fellowship, which supported my writing this paper.