

## Knowing Our Desires

Despite the fact that the term “introspection” is used to name the distinctive first-personal method via which we know our own mental states, it is now widely held that we do not look *inwards* to find out what we believe. Instead, it is thought that we look *outwards* towards the world – we attend to the objects and properties and states of affairs that our beliefs are about. If I want to know, for example, whether I believe that there is coffee in the cupboard, I think about the cupboard and its contents, not about my own mind. The rough idea here is that we *see through* our mental states to the world; our investigation of our mental states is *transparent*

This kind of outward looking account is now generally known as a *transparent account*. Unlike how we know about others' mental states, a transparent method will not involve attending to your behavior. And unlike internal scanning mechanism accounts of self-knowledge, a transparent method will not involve attending to anything mental. These two parts of transparency are made clear in a statement of transparency from Richard Moran's *Authority and Estrangement*:

A statement of one's belief about X is said to obey the Transparency Condition when the statement is made by considerations of the facts about X itself, and not by either an "inward glance" or by observation of one's own behavior (Moran 2001: 101).

### **Transparent Accounts of Desire Introspection**

Although Evans and Moran here state the Transparency Condition in a way that explicitly restricts it to belief, this condition has been thought to extend to other mental states as well:

If asked whether I am happy or wishing that p, whether I prefer x to y, whether I am angry at or afraid of z, and so on, my attention would be directed at p, x and y, z, etc. (Bar-On 2004: 106).

If I am asked (by myself or others) whether I want p to be the case, my attention will be directed at p being the case. To address the question, I do not try to, so to speak, scan my own mind in search of a state that I can identify as the relevant desire. Rather, I concern myself with the outside world by focusing on the intentional object of the desire (Fernandez 2007: 524).

...[in introspecting my preferences] often my eyes are still "directed outward – upon the world." I can investigate my preferences by attending to the *beer* and the

*wine...* (Byrne 2005: 100)

There's still more to say here – what exactly do we do once we have fixed our attention on the beer and wine? Well, one thing we sometimes seem to do when we think about what we want is to look at the pros and cons of an option – things which make it a good thing to do, or make it valuable, or desirable (or undesirable). And this is in fact what has recently been suggested as a transpa

you might deny that such desires exist, as Byrne does, and point out that the sense of “valuing” here need not entail thinking the thing in question to be very important; we can find having a cup of tea, or singing in the shower, valuable in the sense that there is

– the drink *appears* to me as desirable or valuable – I just have other reasons to think that this appearance is not to be trusted – ignoring the reasons to the contrary, it might *seem* like a good idea, but I have defeating evidence – thus I judge that it’s not the thing to do. I think that an account based on *appearances* of value – rather than judgments – is a more promising account of desire introspection.

The idea behind this account is, roughly, that when looking out to the world (in thought or in cases of actually looking) *things appear to us to be desirable*. In the same way that the chocolate cake looks brown, it also looks *tasty* (thus desirable in the tastiness sense). Yet this account is *transparent* because our attention is on the object of desire, and not on ourselves.

The problem with taking inference from judgments of desirability or value as the method via which we know our desires stems from the fact that there are several *different* ways in which we can end up with a belief that something is desirable – and it is generally the way in which we come to this belief, and how that belief is sustained, that affects the reliability of the inference from value to desire. One way in which we come to believe something is desirable or valuable is via testimony. You might tell me that the sushi at this restaurant is good; my dentist might tell me that flossing is desirable. When I come to believe that things are valuable on the basis of testimony, sometimes I also then come to desire them. But not always. This is what I think happens in some cases of desire-free valuing – I come to believe that something is valuable but this fails to generate in me the appropriate desire.

On this view, there are two kinds of desire-free valuing: firstly, there are cases where you have come to believe that something is desirable through outside sources, although

you do not see the thing *as* desirable – such as when my dentist tells me it is desirable to floss. Secondly, there are cases where you *once* saw the thing as desirable and have retained a memory of that (and judge that the value of the thing has not changed), although the thing now *no longer* appears to you to be desirable. For example, I might believe that continuing my once-favourite hobby is valuable, although for the moment it seems to have lost its lustre.

Valuing-free desiring, on the other hand, can be accounted for as cases where we see something as desirable yet do not judge it to be so, perhaps because we have competing evidence that such things are *not* desirable. Another glass of wine might *look* attractive to

something, yet have no motivation to get it – and *vice versa*. Appearances of desirability or value seem to fit better with judgments of what one *likes* than what one feels motivation to get. If motivation and liking are realized by different physical systems, as is suggested by Berridge (2004) and Robinson and Berridge (1998), then it is possible to not feel motivated to do the very thing that appears to you to be valuable. This is a threat to the projective account if, as seems likely, people with decoupled liking and motivational systems would *know* (first-personally) that they lack motivation for things they like.

Even if the properties projected were something like *attraction* – something like a magnetic *pull* – to have access to whether you feel pulled towards you would need to direct your attention inwards. To know what one is *motivated* to get or do one has to look inwards, to an inner push or pull.

This account also faces a challenge from experiments on early childhood theory of mind. If we inferred what we want from appearances of desirability, such that those appearances seemed to be desire-independent objective properties of the desire's object, we would expect children to often make mistakes and attribute their own desires to others. Yet at roughly the same age that they start talking about desire, they recognize that different people want different things.

This is in contrast to their talk of belief, where they often mistakenly attribute their own beliefs to others – young children have been found to not attribute beliefs to others that they know to be false, even if it seems that they have very good reason to think that the other person would have a false belief. Contrasting their own beliefs with those of others typically does 701lse bliktd 50 0 0 Tm /F3..hct





- Berridge, Kent. 2004. "Pleasure, Unfelt Affect and Irrational Desire". In *Feelings and Emotions: The Amsterdam Symposium*, A. Manstead, N. Frijda and A. Fischer (ed.), 243–62. Cambridge: Cambridge University Press.
- Byrne, Alex. 2005. "Introspection". *Philosophical Topics* 33: 79-104.
- Evans, Gareth. 1982. *Varieties of Reference*. Oxford University Press.
- Fernandez, Jordi. 2003. "Privileged Access Naturalized". *The Philosophical Quarterly* 53: 352-72.
- Fernandez, Jordi. 2007. "Desire and Self-Knowledge". *The Australasian Journal of Philosophy* 85: 517-36.
- Gertler, Brie. 2007. "Do We Determine what We Believe by Looking Outward?"  
Available at SSRN: <http://ssrn.com/abstract=996267>
- Moran, Richard. 2001. *Authority and Estrangement*. Princeton University Press.