Is Introspection Inferential?

MURAT AYDEDE

University of Florida
Department of Philosophy
330 Griffin-Floyd Hall
P.O. Box 118545
Gainesville, FL 32611-8545

maydede@phil.ufl.edu

ABSTRACT: Coming to know about our own experiences and their qualitative
color is a matter of introspection. An increasingly influential view identifies such
introspection with a species of “displaced perception” (DP), as Fred Dretske, one of its
most articulate proponents, has called it. This view makes our introspective knowledge
essentially inferential. I present the view, and argue that there is no sound inference to
the target introspective judgments if introspection is a matter of inference in the way
proposed by the DP model. I end the paper by reflecting on the form of the
introspective judgments and their relation to an inferential view.

Suppose there is a red ball against a uniformly gray background moving toward my left. I am
seeing the moving red ball. I am having a visual experience that carries the information (among
other things) that [the ball] is red.¹ Now supposing that I have the concepts RED and SEEING,
and all my other cognitive (including introspective) mechanisms are intact and working
normally, the job is to say exactly how I do come to know that I am seeing [the ball] as red.
How do I come to know, as I shall sometimes put it, that I am seeing red?

¹ ‘[…]’ is to indicate that the linguistic position is to be read transparently. I will
sometimes omit the brackets in what follows. Also, I will not use ‘seeing’ in the success sense,
but rather in the sense of ‘having visual experience’. Words in capitals will refer to concepts
understood as mental representations somehow realized in the brain. I will call the first-order
de re beliefs about the observable properties of objects directly obtained from perceptual
experiences “perceptual beliefs”. I will also assume, as most defenders of the view under
criticism do, that colors are objective qualities of objects.
1 The Displaced Perception Model of Introspection

There is an account answering this question that has been gaining some popularity in recent years. The most explicit formulation of this account is due to Dretske (1995, 1999).\(^2\) Dretske sometimes calls the account the “displaced perception model” (DPM) of introspection. The DPM treats introspection of experiences as a species of conceptually articulated inferential knowledge acquired on the basis of ordinary perception.

Dretske’s examples of displaced perception are suggestive in this regard. I look at the pointer on my scale when I am on it. I see that the pointer points to ‘170’, and I come to know something about myself, say, that I weigh 170 lbs. I look at the fuel gauge on the dashboard of my car, I see the pointer pointing to ‘E’, and I come to know that the gas tank is empty. In each case, what I come to know is not what I directly perceive. I infer my belief about myself or about the fuel tank from a perceptual belief that I acquire by perceiving a different object (i.e., the pointer) via an appropriate connecting belief. In the former case, this is the conditional belief that

\[ \text{If the scale’s pointer points to numeral } x \text{ when I am on it, then I weigh } x \text{ lbs.} \]

In the latter example, the connecting belief is that

\[ \text{If the pointer of the fuel gauge points to ‘E’, then the gas tank is empty.} \]

The fundamental idea of DPM is that my introspective knowledge that I am seeing red is likewise an inference from a perceptual belief that I acquire as a result of my seeing the ball. I will argue in what follows that this view must be incorrect because there is no sound inference yielding the target introspective knowledge.

---

\(^2\) William Seager (forthcoming) and Tye (1995, 2000, and his piece in this volume) explicitly endorse versions of this account. Shoemaker’s view of introspection (1996) can also be seen as a variation of this line. I will restrict the account of introspection, as does Dretske himself, to introspection of experiences. For an attempt to extend the account to cover propositional attitudes, see Seager (forthcoming).

\(^3\) It is clear that such conditionals ought to be hedged by ceteris paribus clauses. I will omit this for convenience in what follows.
DPM is in essential harmony with representationalism about experiential qualia. Representationalism about qualia is the externalist view that qualia are not intrinsic features of experiences that we come to be acquainted with, and thus come to know about, by engaging in introspection: rather, qualia are relational in that they are essentially the representational content of experiences. Experiences represent the world (perhaps non-conceptually) in a certain way. Thus, what we come to know when we come to know about the qualitative content of experiences is how they represent the world. So when I come to know about my visual experience of the red ball moving to my left, what I come to know is how it represents the ball: it represents it as red and moving to my left against a uniformly gray background.

Because DPM treats the introspective knowledge of experiences as beliefs inferentially obtained, it is a version of the Higher Order Thought (HOT) theory of introspection, in that introspection is a matter of having conceptually articulated thoughts (beliefs) about one’s experiences. But since to have such introspective thoughts is (partially) to have thoughts about the representational content of experiences, DPM requires that the introspector have the relevant concepts articulating the content of the experience. On this view, if the introspector can’t have conceptually articulated perceptual thoughts (beliefs) about her environment, she can’t introspect the experiences that give rise to such perceptual thoughts. In other words, she can introspect the qualitative content of her experiences only if, or only to the extent to which, she has the perceptual concepts that apply to the objects that are represented in her experiences. Thus, according to DPM, to say that perceptual concepts are required for the introspection of

---

4 Because its presentation would unnecessarily complicate things for the purposes of this paper, I will omit narrow representationalism here.

5 HOT is supposed to contrast with the Higher-Order Perception (HOP) view of introspection. Not all HOT theories are inferential. See Rosenthal (1997) for an example: he defends a non-inferential HOT theory of introspection (in addition to defending a HOT theory of state consciousness). Tye (2000, this volume), after explicitly endorsing the DPM of introspection, flatly denies that his version is inferential. But as far as I can see, he has only stipulation to back up this rejection: he simply states that introspection is a reliable process that takes the representational content of experiences and delivers introspective knowledge that an experience of the relevant kind with that content occurs. This is unsatisfactory if we are trying to understand, as we do, how introspection works. I will continue to hold that Tye is an inferential theorist insofar as he wants to stick to DPM.
perceptual experiences is not to say that we can’t experience/perceive without those concepts. Without perceptual concepts we are blind only to our experiences, not to the world.\textsuperscript{6}

Notice that DPM also requires that for introspection one have concepts about the perceptual modes through which one gathers the information about one’s environment. In addition to having concepts that apply to the objects represented in one’s experiences, one also needs to have the concepts of experience, sensation, perception, or more specifically, of seeing, hearing, etc. This requires that one have intentional concepts like the concept of representation or “aboutness”. This is important to keep in mind for what follows.

2 Connecting Beliefs

According to DPM, introspection of experiences works as follows. We first perceive our environment, and from the information gathered this way we somehow infer that this information is being obtained by perceiving/experiencing (seeing, hearing, etc.) the environment. DPM is therefore committed to there being a premise in this inference believed to connect this information, $P$, about the environment to the introspective belief that one is having an experience with the content $P$. Let’s see what this connecting belief might be in the case of my coming to know that I am seeing that [the ball] is red.

**STEP 1:** I form the first-order visual/perceptual belief that

(B) [the ball] is red.

\textsuperscript{6} There is a certain sense in which we are also blind to our environment without such concepts if our intentional behavior is determined only by conceptually articulated intentional states (like beliefs and desires), or more fundamentally, if the only way the information in experiences can be used by the experiencer is by way of converting this information into conceptually articulated states. This is why, in my view, Dretske’s non-cognitivism about perceptual experience does not cut much ice vis-à-vis Dennett’s cognitivism (or, vis-à-vis
This is a perceptual de re belief. According to most versions of epistemological externalism, and according to Dretske in particular, this belief state is a knowledge state if its tokening is an appropriate causal consequence of my visual state that carries the information (in analog form) that [the ball] is red. The qualification “appropriate” expresses that the belief state must be caused by the physical feature of the relevant brain state (realizing the visual state) in virtue of whose tokening the visual state carries the relevant information.\textsuperscript{7} Call the tokening of this physical feature, $Pv$, i.e., $v$’s being $P$.\textsuperscript{8} The belief state too is supposed to be a physical structure realized in the brain.

\textbf{STEP 2:} I infer (come to believe) that I am seeing [the ball] as red. Or, better, I infer the \textit{introspective belief} that

\begin{enumerate}
\item[(IB)] I am seeing that [the ball] is red.\textsuperscript{9}
\end{enumerate}

But how is this inference supposed to go after I come to see (believe) \textit{that} [the ball] is red? What is the connecting belief? As far as I can see, there are two candidates:

The \textit{connecting belief} that

\begin{enumerate}
\item[(CB1)] if [the ball] is red, I am now seeing it as red,
\end{enumerate}

\textsuperscript{7} As emphasized by Dretske in many places, but see especially Chap. 5 of his (1981).
\textsuperscript{8} I characterize the DPM view as a naturalist/physicalist theory only because all its advocates do so; the view itself does not require physicalism.
\textsuperscript{9} Or perhaps more accurately: that I am currently having a visual experience [purporting to indicate] that something is red. The reason for the transparent position here is that I may not have the specific concepts to represent to myself that the visual experience is \textit{as of} a red ball. All that is needed is that I have some concept of a visual experience qua representor or qua something purporting to indicate the instantiation of the properties of being red, moving, ball, etc.
or, that

(CB2) if I have now come to believe that [the ball] is red, I am now seeing (visually experiencing, representing, etc.) it as red.

Let’s first take up (CB1). In the absence of other crucial factors, to hold (CB1) (and the like) is epistemically irresponsible because it is often false, and known to be so, unless, of course, we somehow try to build into the de re nature of the perceptual belief, whose content is expressed here by the antecedent, that it is obtained visually/perceptually. But doing this would be circular: we are trying to understand how we pass from a belief about the ball’s redness to a belief about its experience, or rather, how we infer having a visual experience from having a perceptual belief (namely, B) the direct epistemic source of which is that experience. So I take it that as a matter of fact most people (perhaps all), being more or less epistemically responsible, just lack a connecting belief of the (CB1) sort. I take this to be an empirical but obvious truth. If so, to the extent that people do have introspective knowledge of their experiences, introspection cannot, de facto, be inferential in the way proposed by DPM if we take (CB1) as the relevant model for connecting beliefs.

The point, however, is not simply the threat of epistemic irresponsibility (and, to that extent, irrationality) by believing something that is obviously false. After all, a DPM theorist can still insist that we must have connecting beliefs like (CB1) — if not consciously or explicitly, then unconsciously or implicitly — if we have introspective knowledge and introspection is inferential. Rather, it is that we could not possibly have introspective knowledge if one of the crucial premises of the inference we essentially draw on to arrive at this knowledge is routinely false. I assume — as do the advocates of DPM — that we do have extensive introspective knowledge about our experiences and what they are like. But if introspective beliefs are essentially inferentially formed and the inference is routinely unsound, then those beliefs aren’t
knowledge. It is just epistemic luck on our part if they turn out to be true. I don’t think anyone would like to hold such a view.

Let me take up (CB2). First a clarification: as stated, it is not clear how (CB2) allows (IB) to be inferred from (B) since its antecedent won’t match the content of the first-order belief (B), and hence will not allow the required *modus ponens* inference. The essential idea behind suggesting (CB2) as the relevant connecting belief is the intuition that the introspector should not only have a belief about the redness of the ball, but he must also be aware that he has such a belief: i.e., he must also believe that he believes that [the ball] is red. So (CB2) assumes this extra step, and in doing so, it already makes use of our capacity to introspect our first-order perceptual beliefs.

This is problematic depending on what sort of account one is assuming about the introspection of propositional attitudes. If this account is inferential, then it seems that it too requires a connecting belief, and it is any one’s guess what that connecting belief might be. Such a connecting belief, presumably, would be analogous to (CB1): just replace ‘seeing it’ there with ‘believing it to be’. But again the resulting connecting belief would be false. It is obvious that there are many times the *de re* proposition that [the ball] is red is true, but I lack a belief to that effect. To the extent to which this is so, however, holding (CB2) is, again, epistemically irresponsible, and all the points I raised above against (CB1) apply to this case *mutatis mutandis* — including the point about the possibility of introspective knowledge.

On the other hand, if introspecting first-order perceptual beliefs is not inferential, then it is not clear what the wisdom might be in insisting that introspecting experiences is inferential, especially when, from the common sense point of view, an inferential view of introspection is already very implausible to begin with. Nevertheless, it is still consistent to hold a view according to which the introspection of beliefs is not inferential but introspecting experiences is.\(^{10}\)

\(^{10}\) See Seager (forthcoming) for the naturalness of extending the inferential view to propositional attitudes. In fact, it is so natural that any doubt as to whether introspection of
But (CB2) won’t stand anyway. I have so many different ways to come to believe that [the ball] is red — and like most (or perhaps all) people, I know that. If (CB2), as it stands, is meant to be a kind of generalization sufficient (when combined with a belief in its antecedent) to yield the desired introspective knowledge, it is, again, false. The connecting belief in question must allow me to infer how (through what kind of experience or sensory modality) I have come to form the belief in its antecedent, which is precisely the introspective knowledge we are after. But I don’t see how anything like (CB2) can do that without being circular.

However, one way to defend (CB2) is this.\footnote{This is suggested to me by some remarks made by Fred Dretske in personal communication. I don’t think Dretske presently endorses such a view.} It may be that we’ve come to associate certain sensory modalities or kinds of perceptual experiences with certain properties we detect whose concepts we deploy in our first-order perceptual judgments. For instance, we may have learned to associate color properties with visual experiences, or with the visual sensory modality; sound properties with hearing; etc. There are a variety of ways this could occur independently of introspection. It may be that concepts of sensory experiences are formed or acquired just as other theoretical concepts are, partly on the basis of inference to the best explanation and partly by relying on third-person observation or experimentation methods: for instance, we may come to form the concept of an experience and of its kinds by observing others and how they react to their environments. Then, assuming we have the introspective capacity to come to believe the antecedent of (CB2) as discussed above (which already requires a capacity to introspect first-order perceptual beliefs), we can infer the consequent, which is presumably the introspective knowledge the DPM theorist is after. More accurately, the idea is that when I believe that I have come to non-inferentially hold a belief of the form ‘[… is \(F\)’, where \(F\) stands for a sensible property of objects detected in perception and known (on independent, third-person grounds as stated) to be associated with a particular sensory...
modality by a connecting belief of the (CB2) variety, I can infer that I am having a sensory experience (in the associated modality) representing something as $F$.

In effect, this amounts to a proposal to modify (CB2) in a way that would give us the connecting belief that

(CB2*) if I have now non-inferentially come to believe that [the ball] is red, I am now seeing (visually experiencing, representing, etc.) it as red.

The claim is that such connecting beliefs may be true and justifiably held if the introspector has a theory about what properties go with what kinds of sensory experiences. If I have the capacity to know that I have just formed a belief about the instantiation of such a property and know that I have formed this belief non-inferentially, then on the basis of this theory I can infer to what sensory experience I must have had (or am still having if I keep holding the perceptual belief). And all this can happen very fast and reliably.

On this view, introspection is like the descriptive or theoretical knowledge we gain about our bodies. Indeed, the situation here is somewhat analogous to acquiring the concept of a kidney from third person channels by description or by learning the relevant biological theory, and then self-applying it on the basis of an inference that might go like this: “Every member of the human species has kidneys; I am human, hence I have kidneys”.

This attempt to justify (CB2) is implausible. For one thing, it already assumes a great deal of introspective knowledge not only about first-order perceptual beliefs but also about the nature of their sources, specifically about whether they are formed inferentially or non-inferentially. At this point, one might wonder: if we can introspectively track whether we acquire our beliefs non-inferentially, what is the point of insisting that we cannot track what their particular perceptual/experiential sources are (so we must infer them)? If we have assumed this much introspective capacity without relying on DPM, why not go all the way and
say that we can come to know about the experiential sources of our perceptual beliefs without relying on DPM?

Second, there is the problem of common sensibles. If there are sensible qualities that can be detected by more than one modality, DPM seems to be in trouble. Suppose I have come to (non-inferentially) believe that this is spherical (moving, rough, etc.). How can I tell whether I have seen it or felt it (or touched it or even heard it)?

Third, there is the problem of so-called intransitive bodily sensations and their introspection. On the version of DPM under discussion, we come to know whether we are in pain by way of a self-attribution of an experience that is known to be associated with a quality of an object or condition standardly detected by having that kind of experience. So, supposing that pain experiences detect bodily injury or damage, I come to know I am in pain by first detecting the bodily injury, hence coming to hold a belief that, say, [my liver] is injured (in such and such a way), and then by forming the introspective belief that I have come to believe that [my liver] is injured, and finally, on the basis of my theory detailing the relevant associations, I infer that I am in pain. This is almost a *reductio* of the view of introspection under discussion. For note that our concept of pain is a concept that applies, *in the first instance*, to token experiences, and not to injured bodily parts that they might represent. Any case in which we know that we are in pain without having — and knowing whether we have — a belief that a certain part of our body (including internal organs) is injured or damaged is a refutation of this view.

A similar point applies to other bodily sensations, like itches or tickles. Like the concept of pain, the concepts such as ITCH, TICKLE, and so forth, apply to particular kinds of experience rather than of what they might be representing. So, suppose a spot on my back

---

12 Or, perhaps there are no common sensibles, as indeed Dretske has said he holds (personal communication; cf. his (2000)). But this seems to be a hefty price to pay — especially within an information-theoretic psychosemantics.

13 It has come as a surprise to me to learn that many (especially those with representationalist leanings) take this claim to be quite tendentious — i.e., the claim that our ordinary/dominant concepts of intransitive bodily sensations like PAIN, ITCH, and TICKLE
suddenly starts to itch intensely for no apparent reason. I come to know that I’m itching by first coming to know that I have a belief that [that spot on my back] is undergoing some physical disturbance, and infer from this, on the basis of my knowledge of the relevant associations (how do we get these?), that I am itching (or, [that part] is itching). It is clear that if there is any inference involved, it is the other way round: first I come to know that I have an experience of a certain sort (itch), then I infer (if I do) that I have a body part that is undergoing some sort of a physical disturbance, and only then do I come to know (if ever) that I have a belief that I have a body part that is undergoing some sort of a physical disturbance.

Note also that in the case of bodily sensations, DPM’s representationalism becomes extremely strained. For in order to know I have an itch (an experience), I must have the concepts that apply to what my experience represents about those body parts that itch. However, it is clear that I have no idea what it is that I am representing about the spot on my back when it starts to itch. At best, all I can surmise is that some kind of disturbance is going on. I can equally say the same thing about a tickle, or even about a mild pricking pain there. This shows that knowing the phenomenal content of these experiences is not a matter of knowing their representational content through the concepts applicable to the properties that the represented object/condition possesses.14

---

apply, in the first instance, to experiences, rather than to their object. To clarify the claim, suppose that John’s current excruciating experience (call it, E) represents a physical condition in his leg (call it PC) so that our ordinary concept of pain applies to PC in his leg in the first place, then saying the following would be correct: (a) he would not have any pain if he had E but no PC in his leg (as in the case of some phantom limb pains and chronic pain cases); and conversely, (b) he would have pain if he had PC but no E (as would be the case if he had taken pain killers — sic!). On reflection, it is self-evident to me that saying these things would be incorrect given our ordinary/dominant concept of pain. Similarly with other intransitive bodily sensations. Note that this is a factual claim, not a terminological legislation on my part. Furthermore, this fact, in and of itself, does not argue against representationalism about such sensations — although, as I argued in my (2001), it can effectively be used in an argument against a purely representationalist position about pain.

14 This is, of course, not to say that phenomenal content is not representational. In my (2001), I have elaborated extensively on this line against a thoroughgoing representationalism about bodily sensations, and about pain in particular.
Perhaps the major difficulty with this way of defending (CB2) is that it removes any sort of epistemic authority that the introspector has about his own mental states. Whether or not introspective access is infallible, it is certainly true that normally I am the most authoritative epistemic source about my own experiences. The privileged access we normally have to our own conscious experiences is not preserved under this proposal, because anyone who knows that I have just non-inferentially come to believe that this ball is red, has exactly the same authority as I do about the nature of the experience I have just had. Actually having the experience in and of itself bestows no special epistemic authority under the present proposal: it is all a matter of knowing which associations exist and what perceptual beliefs one has acquired.

There are various other, equally serious, problems with this attempt to justify (CB2). But I will spare the reader going over them, when the initial plausibility of this attempt, as I’ve said, is already extremely low.

3 Conclusion

I have done my best to come up with other sorts of connecting beliefs of the required sort, but I am unable to think of any other possibility that is even remotely plausible. The difficulties discussed so far point to the fundamental problem of DPM: its inferential character. There just don’t seem to be any connecting beliefs of the required sort.

I have said that viewing the introspection of experiences as essentially involving an inference is already very implausible. One source for this initial implausibility might be related to the form introspective judgments seem to assume. (IB) essentially amounts to a self-attribution of a visual experience representing the world in a certain way. However, consider the introspective belief of the following form:
that [this] is how the redness of [the ball] is visually experienced (visually registered, represented, etc.\textsuperscript{15} — where ‘[this]’ rigidly picks out the feature, say, $P_v$, of the visual experience that represents the redness of the ball).

An introspective belief of this form seems to capture our intuitions better than (IB) above. Indeed, we seem to be capable of making such introspective judgments. On the face of it, a query about knowing what it’s like to see red is best answered by such a judgment.\textsuperscript{16}

(IBa) is supposed to express a \textit{de re} belief expressing a singular proposition \textit{about} $P_v$ (=about a feature of the experience), whereas (IB) is a \textit{de re} belief expressing a singular proposition \textit{about} me: it says of M.A. that he is having a visual experience representing red. Normally, having the visual experience of red seems to suffice to form a belief like (IBa) without any intervening beliefs — provided that one has the appropriate concepts, RED and VISUAL EXPERIENCE and the like.

Note, however, that the force of (IBa) can be obtained by a self-attribution: “I am now having a visual state representing red by being in \textit{this} sort of state”, where the demonstrative rigidly picks out $P_v$ or the relevant feature of the experience. What matters, therefore, doesn’t seem to be the form \textit{per se}, but rather the \textit{special way} we can directly refer to or pick out the relevant feature of the experience, and assuming a naturalist framework, the relevant brain state, $P_v$. This way seems to be essentially perspectival as well as quasi-indexical.\textsuperscript{17}

Indeed, I don’t think that the issue at hand ultimately comes down to the proper form of introspective judgments. For it is possible to make the singular proposition expressed by (IBa) true by a correct use of it in highly contrived circumstances which don’t involve introspection.

\textsuperscript{15} See note 9 above.

\textsuperscript{16} Or, by remembering making such a judgment — if there were no times we were capable of making such judgments, it is not clear whether we could claim to know what it is like to see red at all.

\textsuperscript{17} ‘Quasi-indexical’ because it also seems to be predicative and not purely demonstrative. See Aydede & Güzeldere (ms.) for a fuller elaboration of what kind of perspectival quasi-
For instance, if I am a scientist who knows what $Pv$ is (say, identical to experiencing red) and how to detect it, then I can monitor my brain states via fancy instruments while having experiences of the relevant sort. Pointing to a screen monitoring the tokening of $Pv$ (or, pointing to a mirror reflecting the image of it), I can then utter a version of (IBa): “this is how redness is experientially represented (by me)” (perhaps seeing the red blood in or around $Pv$ if the monitoring devices require open brain operation). This would be expressing a singular proposition appropriately expressed by a *de re* belief — if $Pv$ is a natural kind. But the belief would hardly be introspective.

However, issues about form are not irrelevant to assessing DPM. For, barring the kind of highly contrived situations I have just mentioned, and assuming that we can and do correctly make such introspective judgements, it is not clear how (IBa) can be (or, need be) inferred from anything, let alone from a conditional connecting belief of the relevant sort with a belief in its antecedent. (And the same goes for its self-attributive variant). So, to the extent that there is introspective knowledge assuming the kind of form (IBa) illustrates, DPM is inadequate.\textsuperscript{18}

indexicality is involved in introspecting our experiences of secondary qualities, and for a general information-theoretic account of introspection of phenomenology.

\textsuperscript{18} I would like to thank Brie Gertler, Eric Margolis, Philip Robbins, and Güven Güzeldere for their helpful comments. Also many thanks to Fred Dretske for sharing his time so generously.
REFERENCES


AYDEDE, M. and GÜVEN GÜZELDERE (ms.). Information, Introspection, and Phenomenal Consciousness. Available at

http://web.clas.ufl.edu/users/maydede/introspection.pdf


http://www.scar.utoronto.ca/~seager/intro.pdf


