

## Harman on Mental Paint and the Transparency of Experience

Erhan Demircioglu\*


Received: 23 January 2019 / Accepted: 26 June 2019

*Abstract:* Harman famously argues that a particular class of anti-functional arguments from the intrinsic properties of mental states or events (in particular, visual experiences) can be defused by distinguishing “properties of the object of experience from properties of the experience of an object” and by realizing that the latter are not introspectively accessible (or are transparent). More specifically, Harman argues that we are or can be introspectively aware only of the properties of the object of an experience but not the properties of the experience of an object and hence that the fact that functionalism leaves out the properties of the experience of an object does not show that it leaves out anything mentally relevant. In this paper, I argue that Harman’s attempt to defuse the anti-functional arguments in question is unsuccessful. After making a distinction between *the thesis of experiencing-act transparency* and *the thesis of mental-paint transparency*, (and casting some doubt on the former,) I mainly target the latter and argue that it is false. The thesis of mental-paint transparency is false, I claim, not because mental paint involves some introspectively accessible properties that are different from the properties of the objects of experiences but because what I call *the identity thesis* is true, viz. that mental paint is the same as (an array of) properties of the object of experience. The identification of mental

---

\* Koç University

 Department of Philosophy, Koç University, Rumelifeneri Yolu, 34450 Sarıyer – Istanbul, Turkey

 [erdemircioglu@ku.edu.tr](mailto:erdemircioglu@ku.edu.tr)



paint with properties of the object of experience entails that the anti-functionalism arguments Harman criticizes cannot be rightly accused of committing the fallacy of confusing the two.

*Keywords:* Functionalism; intentionalism; transparency of experience; mental paint; Gilbert Harman.

Harman (1990) famously argues that a particular class of arguments against functionalism about mental states or events (in particular, perceptual experiences) can be defused by distinguishing “properties of the object of experience from properties of the experience of an object” (31).<sup>1</sup> Functionalism is broadly defined as the view that mental states such as a perceptual experience as of seeing something red should be exclusively accounted for in terms of their relations, relations between those mental states and perceptual input, relations between those mental states and other mental states, and relations between those mental states and behavioral output. The arguments belonging to the class Harman addresses against functionalism rely on as their major premise the thesis that introspective awareness (or attention) reveals that perceptual experiences have some intrinsic qualities, qualities they have “apart from their relations to other things” (33), and conclude on the basis of this idea that given its exclusive concern with the relational features of perceptual experiences, functionalism leaves out and cannot account for those intrinsic qualities.<sup>2</sup> Harman’s reply is, in its essentials, to reject the major premise and maintain that “when we clearly distinguish properties of the object of experience from properties of experience, we see that we are not aware of the relevant intrinsic features of the experience” (49) but aware only of “what are experienced as intrinsic features of the intentional object of experience” (39). And, if the major premise in question is false, as Harman claims it is, then “the fact that functionalism

---

<sup>1</sup> All page references that follow are to this seminal work, unless otherwise noted.

<sup>2</sup> In this paper, I will be almost exclusively concerned with visual experiences (rather than other sorts of perceptual experiences or mental states in general); and by “experience” or “perceptual experience,” I will solely mean visual experience including deceptive as well as veridical experience, unless otherwise noted.

abstracts from the intrinsic character of experience does not show it leaves out anything you are aware of" (41).

The object of one's experience is, on Harman's account, the object one's experience represents as being in a certain way.<sup>3</sup> The object of my current visual experience, for instance, is the coffee cup on my desk which my perceptual experience represents as being red, being located "in front of" the board marker, subtending a particular angle "from here," and so on. The properties of the object of my experience are the properties my experience represents it as having (or, equivalently, the properties I experience it as having).<sup>4</sup> Among these properties of the object of my experience are those properties my experience represents its object as having intrinsically (or apart from its relations to other things in my visual field or to me), and its redness stands out at least as a plausible candidate for being an intrinsic property of (the surface of) the coffee cup (and let us assume, for the sake of the argument, that it *is* an intrinsic property of [the surface of] the coffee cup).

Given Harman's conception of experience as a form of representation, the distinction he wishes to draw between the properties of the object of experience and the properties of the experience of an object amounts to being a specific version of the more generic distinction between the properties of a represented object and the properties of a representation of that object, in which case some *other* specific versions of the latter might prove useful in understanding the former. Harman thus writes:

(In a painting of a unicorn) the unicorn is pictured as having four legs and a single horn. The painting of the unicorn does

---

<sup>3</sup> In his paper, Harman assumes, unproblematically given his purposes, that the intentionality of experiences is to be accounted for in representational terms (34). I will adopt the same assumption in this paper, and use such locutions as "the object of experience," "the object represented by the experience," "what the experience represents," and "the object the experience is of" interchangeably.

<sup>4</sup> Harman writes: "When you attend to [...] your experience of the redness of an apple, you are attending to [...] *a quality of the apple*" (41, emphasis mine). My experience of the redness of an apple is an experience that represents the apple as red, and to say that the apple being the object of my experience is red is to say that my experience represents its object, the apple, as red (more on this below).

not have four legs and a single horn. The painting is flat and covered with paint. The unicorn is not pictured as flat or covered with paint. Similarly, an imagined unicorn is imagined as having legs and a horn. The imagining of the unicorn has no legs or horn. The imagining of the unicorn is a mental activity. The unicorn is not imagined as either an activity or anything mental. (35)

Just as we need to distinguish the properties of the object represented, on different occasions, by the painting and the imagining, which is in this case the properties of the unicorn, from the properties of what is doing the representing (i.e., the painting and the imagining), we also need to distinguish the properties of the coffee cup represented by my visual experience from the properties of what is doing the representing (i.e., my experience). Furthermore, just as the unicorn pictured and imagined on these occasions can be plausibly thought of as having four legs intrinsically and neither the picture, we can suppose, nor the imagining is four-legged, my experience of the coffee cup might well not be red despite its object being intrinsically red. So, that my experience is intrinsically red does not follow from the fact that its object is intrinsically red, and a fortiori, it is false that if the object of my experience is intrinsically red, introspective awareness shall reveal that my experience itself is intrinsically red.

Harman's distinction between the properties of the experienced object and the properties of the experience establishes that the properties of the experienced object are not necessarily properties of the experience, but it falls short of establishing that they never are. This is because the more generic distinction between the properties of the represented object and the properties of the representation of that object does not exclude the possibility that there are cases in which those two sets of properties coincide. In fact, there are some clear cases in which the two coincide. So, for instance, the color of the unicorn represented by the painting is (non-accidentally) the same as the color of that part of the painting of the unicorn representing the color of the unicorn: if this unicorn is represented as white in the painting, (the relevant part of) the painting qua the representation is also white (and non-accidentally so). This means that the argument against functionalism from the intrinsic properties of

experiences has not yet been fully defused at this stage of the dialectic, by the distinction between the properties of the experienced object and the properties of the experience. There is still room for the proponent of the argument to argue that some intrinsic properties like redness might well be of both the experienced object and the experience, in which case the argument from the intrinsic properties of experiences may proceed as before.

A proponent of an argument from the intrinsic properties of experiences does not need the thesis that all properties of the experienced object are necessarily the properties of the experience but could work with the weaker thesis that they sometimes are. Given this, the success of the attempt to defuse that argument requires further considerations to be brought in. And, this is where Harman appeals, though he never uses the term in the paper, to the *transparency* of visual experiences.<sup>5</sup> His famous illustration of Eloise turning her attention to her visual experience of the tree goes like this:

When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experience. And this is true of you too. (39)

Harman's distinction between the properties of the experienced object and the properties of the experience discloses a possible ambiguity in the major premise of a particular class of anti-functional arguments, viz. that when we have perceptual experiences, we are introspectively aware of

---

<sup>5</sup> The use of the term 'transparency' in the context of philosophical discussions about the nature of visual experiences dates back to Moore (1903), which says the following regarding "the sensation of blue:" "The term blue is easy enough to distinguish, but the other element which I have called 'consciousness'—that which sensation of blue has in common with sensation of green—is extremely difficult to fix [...] And, in general, that which makes the sensation of blue a mental fact seems to escape us; it seems, if I may use a metaphor, to be transparent, we look through it and see nothing but the blue" (Moore 1903, 446). The term's wide currency in the recent literature owes to Tye (1992).

(the) intrinsic properties (of *something*); and, it makes clear that those arguments can only be successful if the properties we are thus introspectively aware of are not *only* the properties of the experienced object but also the properties of the experience.<sup>6</sup> It is at the next stage, however, Harman's defusing attempt is completed, where it is maintained that our introspective findings tell us that the *only* properties to turn our attention to in such phenomenological inspections are properties of the experienced objects. According to Harman, we have "no access at all" (39) to the intrinsic properties of our experiences.

What exactly is it that we can never be aware of when we try to turn our attention to our experiences? There are two possible candidates here, depending on whether one takes as one's model for the representational aspects of visual experience the imagining of a unicorn or the painting of a unicorn. The imagining of a unicorn, as Harman notes, is "a mental activity" (35) whose intentional object is a unicorn; and, if we understand the transparency of visual experiences along the dimension of such "activities," then it amounts to the thesis that when we try to turn our attention to our visual experiences of objects, we can never be aware of the intrinsic properties of *the experiencing* of those objects. Let me call this thesis the thesis of *experiencing-act transparency*. On the other hand, the painting of a unicorn is not an activity (notwithstanding the ambiguity of the term 'painting'), let alone a mental activity (though it seems to require some), but, well, "a thing" by virtue of the particular pattern of the paint on which it takes a unicorn as its intentional object; and, if we understand the transparency of visual experiences along the dimension of such "things," then it amounts to the thesis that when we try to turn our attention to our visual experiences of objects, we can never be aware of, to use Harman's gripping term (39), *the mental paint*, those intrinsic properties (of what one might call the 'mental canvas') by virtue of which those experiences are experiences of those objects (or represent what they

---

<sup>6</sup> As Harman notes in (1995), introspective awareness in question is "direct" (Harman 1995, 75) or "non-inferential."

represent). Let me call this thesis the thesis of *mental-paint transparency*.<sup>7, 8</sup>

It seems clear that the thesis of experiencing-act transparency is not equivalent to the thesis of mental-paint transparency. Mental paint consists

---

<sup>7</sup> There are two points I want make about the notion of mental paint I adopt in this paper. First, Harman himself defines mental paint in the way I have just defined, i.e. as an array of properties by virtue of which an experience is of the objects that it is of (or represents what it does). Harman writes: “In the case of her visual experience, I want to say that Eloise is not aware of, as it were, the mental paint by virtue of which her experience is an experience of seeing a tree” (39), and in his response to Block (1995), “The issue is whether we can become directly or introspectively (as opposed to inferentially) aware of those aspects of perceptual experience—the mental paint, etc.—that serve to represent what we experience” (Harman 1995, 76).

Second, the notion of mental paint is sometimes broadly understood as accommodating both properties of the experiencing act and properties by virtue of which an experience represents what it does. For instance, Block (2010) writes: “Are phenomenological characters of perception—e.g., what it is like to experience redness or roundness—philosophically reducible to the redness or roundness of the objects one sees or to representation of redness or roundness? If there is no such reduction, then there can be said to be mental paint” (Block 2010, 23–24). My use of the notion of mental paint is narrower than Block’s; and by “mental paint,” I mean only those properties of an experience by virtue of which it represents what it does.

<sup>8</sup> Of course, if experiences are brain states, then there will be more to experiences than their objects and the properties of those objects just as there is more to sentences, e.g. their syntactical and morphological features, than the objects and properties they refer to (Block 1995, 26). The question Harman is dealing with, however, is not simply whether there is more to experiences than the properties of their objects but what “psychologically relevant” features of an experience are there that we can be introspectively aware of. Harman writes: “According to functionalism, the *psychologically relevant* properties of an internal process are all functional properties [...] I have been considering the objection that certain intrinsic features of experience must be *psychologically relevant* properties apart from their contribution to function, since these are the properties we are or can be aware of” (41–42, emphases mine). So, the thesis of mental-paint transparency is better understood as asserting that when we try to turn our attention to our visual experiences of objects, we can never be aware of those intrinsic “psychologically relevant” (or mental) properties by virtue of which those experiences represent what they do.

in those intrinsic features of a visual experience by virtue of which that experience has, say, the tree as its object, and it is *not* by virtue of the intrinsic properties of the experiencing act that that experience has the tree as its object. The idea that it is by virtue of the intrinsic properties of the experiencing act that an experience is of the object that it is of (or represents what it represents) is inconsistent with the evidently true assumption that the very same (kind of) experiencing act can have different objects on different occasions (e.g., a tree on one, a coffee cup on another, and so on): the very same experiencing act would not be directed towards different objects but would always be directed towards the same object if it were by virtue of its intrinsic properties that it is directed towards an object. Consider the following analogy. I can kick different objects, say, a soccer ball or a basketball. The very same act of my kicking may take different balls as its objects. In case I kick the soccer ball rather than the basketball lying next to it, it is not by virtue of the intrinsic properties of the kicking act that my kicking the soccer ball takes as its object the soccer ball (but not the basketball). Similarly, the visual act of experiencing may take different things as its objects; and in case it takes a particular thing as its object, it is not by virtue of its intrinsic properties that it takes as its object that thing (but not another thing).

The distinction between the thesis of experiencing-act transparency and the thesis of mental-paint transparency enables a suitable taxonomy of some visual phenomena that are typically appealed to by the opponents of the sort of representationalism advocated by Harman. Two such phenomena are blurred vision and double vision. In the case of some typical examples of blurred vision, it seems that blurriness is not presented as a property of the objects of experience but as a property of the experiencing act itself: what introspectively seems to be blurred is not the objects but the experiencing itself. So, blurred vision is better conceived as militating against the thesis of experiencing-act transparency.<sup>9</sup> In the case of some typical examples of double vision, on the other hand, what introspectively appears to be

---

<sup>9</sup> Smith (2008) writes: “Suppose a myopic person were suddenly to start seeing more and more clearly until he ended up with 20/20 vision. This change in experience would not be taken by this person to be a change in the features of the objects seen. It would immediately be taken for what it is: a change in the character of the visual



double seems to be neither the objects of experience nor the experiencing-act itself but could reasonably be taken as those features of the experience by virtue of which the experience represents what it does. So, double vision is better interpreted as militating against the thesis of mental-paint transparency.<sup>10</sup>

In order for Harman's defusing attempt to be successful, both the thesis of experiencing-act transparency and the thesis of mental-paint transparency must be true. If the thesis of experiencing-act transparency is false and the visual experiencing of a particular object has intrinsic properties that can be identified through introspection, then the anti-functional arguments that Harman targets go unscathed by the distinction between the properties of the experienced objects and the properties of *the experiencing act*.<sup>11</sup> And, if the thesis of mental-paint transparency is false and introspective awareness reveals the mental paint by virtue of which experiences are

---

experience itself. Blurriness is not a way that things in the world *seem to be*. It is, however, a feature of experience of which we are usually aware when it is there. The Transparency Thesis is therefore false" (Smith 2008, 201). Smith argues that in the case of blurred vision, objects are "seen blurrily" (Smith 2008, 202), and therefore, Smith's conclusion that the transparency thesis is false is to be understood as the thesis that the thesis of *experiencing-act* transparency is false. (See also Boghossian and Velleman [1989], Pace [2007], and Allen [2013].)

<sup>10</sup> Boghossian and Velleman (1989) write: "If you press the side of one eyeball, you can see this line of type twice without seeing the page as bearing two identical lines of type. Indeed, you cannot even force the resulting experience into representing the existence of two lines, even if you try. Similarly, you can see nearby objects double by focusing on distant objects behind them, and yet you cannot get yourself to see the number of nearby objects as doubling [...] None of these experiences can be adequately described solely in terms of their intentional content. Their description requires *reference to areas of color in a visual field*, areas that split in two [...] without anything's being represented to you as being so" (Boghossian and Velleman 1989, 94, emphasis mine). I suggest that Boghossian and Velleman are to be understood as arguing against the thesis of mental-paint transparency. For a defense of representationalism against objections from such visual "oddities" as double vision and blurred vision, see Tye (2002).

<sup>11</sup> In his response to Block (1996), Harman writes: "What it is like to see something as ahead and to the right is not normally the same as what it is like to hear something as ahead and to the right" (Harman 1996, 76). This strongly suggests that Harman

experiences of objects that they are actually of, then the anti-functionalist arguments in question go unscathed by the distinction between the properties of the experienced objects and the properties arranged in such a way as to constitute *the mental paint*.

My main concern in this paper is the thesis of mental-paint transparency, but I cannot help making a few (and admittedly quick) critical observations, some of which are familiar and more persuasive than others, about the thesis of experiencing-act transparency, the one that in any case occupies, historically speaking, the more prominent position among the two. First, if the experiencing act does not have intrinsic properties that we are aware of or can be revealed by introspection, as the thesis of experiencing-act transparency claims it does not, then it is at least not clear that we can plausibly claim that when we turn our attention to our visual experiences, we are aware of the properties of the experienced *objects*. This is, in effect, Moore's main problem in his famous "The Refutation of Idealism" (1903), the work which still continues to set the stage for philosophical investigations into the transparency of experience. In that work, Moore's putative refutation of idealism rests on the thesis that introspective reflection provides good reason to prefer the act-object model of experience over its competitor (which is what Moore himself calls, somewhat confusingly when viewed from the contemporary perspective, "the content view" [Moore 1903, 447], a forerunner of the adverbial account), and the problem Moore faces is why, if experiences have an act-object structure, we do not typically find

---

is an "intramodal intentionalist" in Byrne's sense: "*Intermodal* intentionalists hold, while *intramodal* intentionalists deny, that the phenomenal difference *between* perceptual modalities—between visual and auditory experiences, for example—is determined by a difference in content" (Byrne 2001, 205). Intramodal intentionalists typically account for the phenomenal difference between perceptual modalities by an appeal to their distinct functional roles in one's overall cognitive economy (see Dretske [1995] and Tye [1995]). I do not wish to claim here that the phenomenal difference between perceptual modalities cannot be adequately accounted for by a functionalist story, but that the anti-functionalist arguments Harman is targeting can be undermined *solely* by making the distinction between properties of the experience of the object and properties of the experience of the object (along with the deliverances of introspection), as Harman intends to do in his (1990), only if the thesis of experiencing-act transparency is true.

the act component in our introspective queries. Moore assumes, rightly I think, that introspection can provide positive support for the act-object model only if both components are introspectively accessible, and the apparent resistance of the act component to introspective access appears to undermine that model. Moore's solution is, as far as I can see, captured by his following remarks: "Yet it [the experiencing act] *can* be distinguished if we look attentively enough, and if we know that there is something to look for" (Moore 1903, 550). So, I take it that for Moore, it is only difficult but not impossible to introspectively attend to the act component of experiences.<sup>12</sup> The point I wish to make is not that what I take to be Moore's solution to the problem is adequate, but that the very problem he realizes and struggles with also afflicts Harman's view about experiences. The problem is that it is not clear that we can plausibly adopt the talk of the "objects" of experiences, as Harman does, without endorsing the act-object model because the only plausible answer to the question of what the object of an experience should be 'combined,' 'supplemented' or '(inter)penetrated' with in order to get the experience of the object appears to be *the experiencing act*. And, it is at least dubious that we can plausibly presume that introspection prefers the act-object model to its rivals (e.g. the adverbial model), as Moore himself is so acutely aware, if the act component is not introspectively accessible.

Secondly, even if we set aside Moore's problem and grant that the idea that experiencing acts are not introspectively accessible looks plausible when such acts are inspected intra-modally (within a perceptual modality), there are good reasons, as various philosophers have pointed out before, to doubt that idea when those acts are inspected inter-modally (across different modalities). This is because it is not clear that we can plausibly distinguish perceptual modalities as different as vision, tactition, and olfaction, without pointing out the intrinsic phenomenological differences between acts that accompany those senses, like seeing, touching, and smelling. There is, it seems, all the phenomenological difference between seeing and feeling a square object; however, given the sameness of the object, it is not clear how we can account for the difference without appealing to the intrinsic

---

<sup>12</sup> For an influential defense of this idea, see Kind (2003, 229).

phenomenological differences between seeing and touching. It is on the basis of such inter-modal considerations Grice (1989), among others, defends the introspective accessibility of the intrinsic features of visual experiencing when he writes:

In addition to the specific differences between visual experiences, signaled by the various property-words employed, there is a generic resemblance signaled by the use of the word “look,” which differentiates visual from nonvisual sense-experience. This resemblance can be noticed and labeled, but perhaps not further described. (Grice 1989, 267)

Furthermore, though Grice’s (mild) skepticism about the prospects of describing the distinctive characteristic of visual experiences appears to be widely shared, some philosophers have gone some significant way towards describing the intrinsic phenomenological differences among various sorts of perceptual modality. For instance, comparing the main forms of perception from what he calls a “purely phenomenological point of view” (Broad 1952, 30), Broad writes:

In its purely phenomenological aspect *seeing* is ostensibly *saltatory*. It seems to leap the spatial gap between the percipient’s body and a remote region of space. Then, again, it is ostensibly *prehensive* of the surfaces of distant bodies as colored and extended, and of external events as color-occurrences *localized* in remote regions of space. In its purely phenomenological aspect *hearing* is ostensibly prehensive, not of bodies, but only of events and processes as occurrences of sound-qualities. It is not ostensibly saltatory, for these events or processes are not heard as localized in remote restricted regions of space. They are heard rather as emanating from remote centers and pervading with diminishing intensity the surrounding space. (Broad 1952, 32)

One might agree or disagree with Broad’s observations here (I, for one, am inclined to agree with a good portion of them), but their availability counts not only against Grice’s (and others’) skepticism about the prospects of providing the relevant descriptions but also, and more importantly for the purposes of this paper, against the thesis of experiencing-act transparency.

This is because if that thesis were true, then Broad would not be able to provide the account above of what “phenomenologically” differentiates seeing from hearing. Or, we can say this at the least: the fact that Broad, for one, has provided an account along the lines above is a formidable challenge to the proponent of the thesis of experiencing-act transparency.

Thirdly, and finally, if the thesis of experiencing-act transparency is true, then it is not clear in virtue of what perceptual experiences are conscious. A visual experience of seeing, say, a red object is a (phenomenally) conscious experience: there is something it is like to have that experience. Now, it cannot be solely in virtue of the redness of that object that the experience in question is conscious given the fact that the redness of that object is not conscious in any sense of the term: there is nothing it is like for that object to be red. The only other candidate that appears to remain and we might appeal to is the act component, the visual experiencing-act: it must at least in part be in virtue of some intrinsic properties of the visual experiencing act that the experience of seeing a red object is conscious. Let us call those intrinsic properties of the visual experiencing act by virtue of which a visual experience is conscious its *conscious properties*.<sup>13</sup> Now, the thesis of experiencing-act transparency claims that the intrinsic properties of the visual experiencing act are not introspectible. So, according to this thesis, the conscious properties of the experiencing act are not introspectible. But this is at least confusing because it does not seem that we can make sense of the idea that introspection is limited with respect to its power to detect conscious properties, that there might be conscious properties that are closed in principle to introspection. And this is in turn because it does not seem that there is any other criterion for the identification of conscious properties than appealing to introspection. This does not mean that introspective awareness is required for consciousness or that it possesses such ideal epistemic virtues as infallibility, incorrigibility, or perfect reliability; however, it means that introspective accessibility is the arbiter by which we can tell conscious properties. Hence the idea that experiencing acts have

---

<sup>13</sup> The notion of conscious properties here is similar to Sundström’s (2018) notion of “consciousness properties:” “*Consciousness properties* are properties that contribute to making up what things are or can be *like* for subjects” (Sundström 2018, 681, emphases original).

some conscious properties that are in principle introspectively inaccessible appears to be barely intelligible.<sup>14</sup>

I now want to turn to the main topic of this paper, the thesis of mental-paint transparency, viz. that mental paint is not introspectively accessible (or that the intrinsic properties by virtue of which visual experiences are of the objects they are of are not introspectively accessible). As we have seen, Harman's defense of this thesis rests, in effect, on the distinction between the properties of the objects of experience and the mental paint (the properties by virtue of which experiences are of those objects they are of), and the deliverances of his introspective queries. I grant, and agree with Harman, that the properties of the objects of experiences are introspectively accessible but I want to deny that mental paint is not introspectively accessible. This is, unlike the main line of opposition to the thesis of mental-paint transparency one typically finds in the literature, not because mental paint involves some introspectively accessible properties that are different from the properties of the objects of experiences but because what I shall call *the identity thesis* is true, viz. that the properties of the objects of experiences *are* the properties by virtue of which those experiences are of the objects they are of (or represent what they represent).<sup>15</sup> The upshot is

---

<sup>14</sup> Compare Goldman's (1993) "rock objection" to higher-order (thought) theories of consciousness: "How could possession of a meta-state confer subjectivity or feeling on a lower-level state that didn't otherwise possess it? Why would being an intentional object or referent of a meta-state confer consciousness on a first-order state? A rock does not become conscious when someone has a belief about it. Why should a first-order psychological state become conscious simply by having a belief about it?" (Goldman 1993, 368) Furthermore, it is worth noting that the intimate tie between being conscious and being introspectible in principle has nothing much to do with Searle's rather controversial thesis that there are no unconscious-in-principle mental states, that "we have no notion of the unconscious except that which is potentially conscious" (Searle 1992, 152).

<sup>15</sup> Two points (one clarificatory, the other qualificatory) need to be made here. First, the identity thesis is *not* the thesis that properties of the object of experience are identical to properties of the experience of an object but the thesis that properties of the object of experience are identical to a subclass of properties of the experience of an object (namely, properties by virtue of which an experience represents what it does). Surely, a given experience might have a property (like occurring on a certain

that Harman's distinction between properties of the objects of experience and mental paint turns out to be a distinction without a difference (while the distinction between properties of the objects of experience and properties of experiencing act need not be and, I believe, is not).

Let us go back to my visual experience of the red coffee cup. The object of my experience is represented as being red, and given that, for Harman, the properties of the object of experience are those properties it is represented as having, the object of my experience *is* red. So, one of the properties of the object of my experience is redness. Now, what are those properties of my experience by virtue of which it is of a red object? I maintain that redness is one of those properties by virtue of which my experience is of a red object because, if the object of my experience were represented not as red but, say, as green, then the object of my experience would be green and hence my experience would not be of a red object. It is by virtue of redness being represented as being instantiated by the object of experience that the experience represents what it does (i.e. a red object). If this strikes you as truistic, then note that this counts in favor of the identity thesis because if the identity thesis is true, then it *is* truistic. Try to conceive of a visual experience whose intentional object is red, and the property by virtue of which the experience is of a red object is not redness but greenness. I predict that you will feel an immediate difficulty in envisaging a scenario with these two features: the idea that the object of the visual experience is red appears to "cancel out" the idea that the property by virtue of which the experience

---

day) that its object does not (or need not) have. Secondly, the identity thesis thus clarified needs a qualification, given that properties by virtue of which an experience represents what it does might involve a property (like being caused by its object) its object does not (or need not) have. The identity thesis suitably qualified may take one of the following forms: it is to be understood either as the thesis that properties of the object of experience are identical to *mental* (or, as Harman puts it, *psychologically relevant*) properties by virtue of which an experience represents what it does (see fn. 8), or as the thesis that properties of the object of experience are *among* properties by virtue of which an experience represents what it does. I will argue that the identity thesis qualified in one of these ways undermines Harman's response to the anti-functionalism. For convenience, however, I will suppress the qualification and take the identity thesis in its unqualified form, since the qualification has no direct bearing on my argument.

is of a red object is greenness, and vice versa. Similar observations can be made about the shape properties of the objects of experiences. The object of my current experience is represented as being coffee-cup shaped, and hence, it *is* coffee-cup shaped. One of its properties is having a coffee-cup shape. Now, what are those properties by virtue of which my experience is of a coffee-cup shaped object? I maintain that being coffee-cup shaped is one of those properties by virtue of which my experience is of a coffee-cup shaped object because, if the object of my experience were represented not as coffee-cup shaped but, say, as rectangular, then the object of my experience would be rectangular and hence my experience would not be of a coffee-cup shaped object. More generally, there is no property that I find in my experience in question of the red coffee cup that is a property of the object of my experience but not a property by virtue of which my experience is of that object. Mental paint qua properties by virtue of which an experience is of the object that is of is, as it were, right before our eyes and as introspectively accessible as properties of the object, and this appears to be so evidently because mental paint is the same as (an array of) properties of the object of experience.

The identification of mental paint with properties of the object of experience should not come as a surprise if a trap that Harman seems to fall into here is avoided. Consider a non-transparent representation, for instance, a painting of a unicorn. “In the case of a painting,” Harman notes, “Eloise can be aware of those features of the painting that are responsible for its being a painting of a unicorn. That is, she can turn her attention to the pattern of the paint on the canvas by virtue of which the painting represents a unicorn” (39). In the case of a painting of a unicorn, we can, Harman maintains, clearly distinguish the vehicle of representation (the painting) from the object of representation (the unicorn), and turn our attention from one to the other. However, in the case of a (hallucinatory) visual experience of a unicorn, we cannot clearly distinguish the two (the mental paint and the unicorn) and turn our attention from one to the other. The conclusion Harman draws from this asymmetry between paintings and visual experiences is that mental paint is not introspectively accessible. However, this is a non-sequitur because it rests on the unwarranted assumption that if the object of representation is introspectively accessible, then



the vehicle of representation can only be introspectively accessible if it can be introspectively distinguished from the object of representation. This assumption is unwarranted because there is nothing in the original definition of mental paint that places such a constraint as introspective distinguishability from the object of representation on the introspective accessibility of mental paint. Mental paint consists, recall, in those properties by virtue of which an experience is of the object that it is of, and the question whether mental paint is introspectively accessible is the question whether *those* properties are introspectively accessible: if they are accessible, mental paint is accessible; if not, not. Accordingly, the fact that the mental paint is not introspectively distinguishable from the properties of the object of experience has no tendency to show that the mental paint is not introspectively accessible. Once the unwarranted assumption in question is discarded, it is clear that just the opposite is indeed true: the fact that the mental paint is not introspectively distinguishable from the properties of the object of experience supports the thesis that the former is as introspectively accessible as the latter and, at one remove, the thesis that the two are identical.

The truth of the identity thesis is what makes visual experiences (and other perceptual experiences), at least to some extent, philosophically troubling and fascinating. It captures, for instance, what various philosophers have, on one clear interpretation, meant by claiming that visual experiences are “self-presenting” states, that they are such that their objects *present themselves* to the subject without any other intermediary objects functioning as representational mediums.<sup>16</sup> An interesting philosophical question

---

<sup>16</sup> For a theory of knowledge that builds on a notion of self-presenting states, see Chisholm (1966). However, note that the notion of self-presenting states Chisholm adopts is, unlike the interpretation I prefer of that notion, more epistemological than ontological. According to Chisholm, a self-presenting state is such that if the subject is in that state, then it is *evident* to her that she is in that state (or, *roughly*, she is justified in believing that she is in that state). I believe but will not argue that one can consistently hold that there are self-presenting states in the ontological sense of the term without there being self-presenting states in Chisholm’s epistemological sense of the term. (Additionally, one might argue, though I will not in this paper, that the connection between the epistemological and ontological notions of self-presenting states is that the fact that their objects present

here is how such self-presenting states are to be accounted for, where the duality of the object of representation and the vehicle of representation (in our case, mental paint) collapses into a peculiar sort of unity. Furthermore, the thesis that mental paint is the same as properties of the object of experience accords well with William James' following eloquent characterization of the history of philosophy of perception:

The whole philosophy of perception from Democritus's time downwards has been just one wrangle over the paradox that what is evidently one reality should be in two places at once, both in outer space and in a person's mind. 'Representative' theories of perception avoid the logical paradox, but on the other hand they violate the reader's sense of life, which knows no intervening mental image but seems to see the room and the book immediately just as they physically exist. (James 1904, 81)

Consider the painting of a unicorn again. The painting of a unicorn does not give rise to a problem similar in structure to the problem (or "the paradox") that we encounter in the case of visual experience. In the case of the painting, "what is evidently one reality" is *not* (and does not appear to be) "in two places at once," both out there in the wilderness as a canvas-independent object and on the canvas: the unicorn is not (represented as being) on the canvas, and the canvas is not the unicorn. And, this is plausibly because the properties of the vehicle of a pictorial representation and the properties of the object of a pictorial representation are different and can be clearly distinguished. However, in the case of visual experience, James points out, what is evidently one reality is (or appears to be) in two places at once, both in outer space and *in the experience*. And, this is plausibly because the properties of the vehicle of a visual representation (that is, mental paint) and the properties of the object of a visual representation

---

themselves without any other intermediary objects (the ontological notion) explains why it is difficult to see how one can be in those states without its being evident to one that one is in those states (the epistemological notion). For an excellent discussion of the notion of self-presenting states, see Lehrer (2002). Lehrer is emphatic that in the case of conscious mental states including visual experiences, the distinction between the vehicle of representation and the object of representation disappears (Lehrer 2002, 422, 426).

cannot be distinguished and indeed are identical. Why does the unicorn, an object distinct from the visual experience of a unicorn, appear to be *in* (in some clear sense) the visual experience of a unicorn (but *not* in the painting of a unicorn)? The answer is that the properties of the vehicle of a visual experience cannot be distinguished from the properties of its object, while the properties of the vehicle of a pictorial representation can be distinguished from the properties of its object.<sup>17</sup> This means that the problem that James thinks defines “the whole philosophy of perception from Democritus’s time downwards” arises because of the indistinguishability or identity of the mental paint and the properties of the objects of a visual experience.

There are three objections I want to address, in an order ascending in force, against the identity thesis and one objection regarding the bearing of the identity thesis on Harman’s defense of functionalism. First, it might be objected that the identity thesis entails the thesis that visual experiences present us with mental paint, which is in turn what only a sense-datum theorist would wish to endorse. So, the objection goes, since the sense-datum theory is false, the identity thesis must also be false.

This objection trades on the ambiguity of the term ‘mental paint.’ The sense in which the sense-datum theorist defends the thesis that visual experiences present us with mental paint is not necessarily the sense in which the identity thesis entails that visual experiences present us with mental paint. The sense-datum theory holds that visual experiences present us with mental paint in the sense that they present us with mental (or internal) objects and properties rather than public (or external) objects and properties. The paint we are presented with in having visual experiences is, on (a traditional version of) that theory, *mental* in the sense that the objects and properties we thereby see (or “sense”) are such things that can only be ‘located’ in the subject’s mind. However, the sense in which the identity thesis entails that visual experiences present us with mental paint is consistent with the thesis that the paint we are presented with in having visual experiences is *not* mental, that the objects and properties we see are such things that can (only) be located in the public (or external) world. On this

---

<sup>17</sup> Thanks to an anonymous reviewer for pressing on this issue.

sense, to say that visual experiences present us with mental paint is to say that visual experiences present us with those properties by virtue of which those experiences represent their objects and properties, *whether those objects and properties themselves be mental or not*. Briefly put, mental paint in the sense that is of concern to us need not be mental in the sense-datum theorist's sense and, accordingly, visual experiences may present us with mental paint in the former sense without presenting us with mental paint in the latter sense.

Secondly, it might be objected that the argument I have presented for the identity thesis works only if, for Harman, the properties an experience represents its object as having are the properties of that object; however, if the properties an experience represents its object as having are the properties of that object, then it is not possible for the experience to be illusionary or hallucinatory (or to *misrepresent* its object in one way or another). Since Harman explicitly allows that possibility (34), as any bona fide representationalist would do, Harman should not be viewed as holding that the properties an experience represents its object as having are the properties of that object.

There are two things I would like to say in response. First, the thesis that the properties a visual experience represents its object as having are the properties of that object does not exclude the possibility of misrepresentation. Suppose that there is a red cup before me on the table, which my visual experience represents as green. In that case, the object of my experience is a green cup, and it is actually because the object of my experience is a green cup that my experience misrepresents what is really there (or is illusory). Secondly, Harman's treatment of hallucination as a form of perceptual misrepresentation supports the thesis that he holds that the properties an experience represents its object as having are the properties of that object. Suppose that it looks to Eloise as if there is a tree before her, while there is no such thing in the environment. In such a case, Harman says, "what Eloise sees before her *is a tree*, whether or not it is a hallucination. That is to say, the content of her visual experience is that she is presented with a tree, not with an idea of a tree" (36, emphasis mine). It is clear that the object of Eloise's experience ("what she sees before her") can *be* a tree, as Harman claims, only because her

experience *represents* it as being a tree, given that there is *ex hypothesi* no tree in the environment.<sup>18</sup>

Thirdly, it might be objected that it does not make sense, or is at least odd, to say that my experience of the redness of the coffee cup is red or to say that my experience of the coffee-cup shape of the coffee cup is coffee-cup shaped: experiences can neither be red nor be coffee-cup shaped. Given that the mental paint consists in the properties by virtue of which my experience is of the object that it is of, then, the objection goes, it seems that claiming that the mental paint is the same as the properties of the object of experience, I am committed to make such absurd or at least odd remarks. Hence, the mental paint cannot be the same as the properties of the object of experience.

I would like to respond to this objection by reiterating the distinction between experiencing act and mental paint. I agree that experiences *qua* experiencing acts cannot have such properties as color or shape, but it is clear that the identity thesis does not imply anything to the contrary, given that mental paint is not experiencing act. Furthermore, just as one can consistently claim that the object of experience is red without claiming that the experience itself is red, one can also consistently claim that the mental paint is red without claiming that the experience itself is red, if the identity thesis is true. However, the basic concern behind the objection might perhaps be clarified by emphasizing that mental paint is the properties *of an experience* by virtue of which the experience is of the object that it is of, and hence, that if mental paint is the same as such properties of the object of experience as redness and being coffee-cup shaped, then one cannot avoid the conclusion that those properties might be (and, in some cases, are) the properties *of an experience*. My reply now is twofold. First, it is not clear

---

<sup>18</sup> The notion of the object of experience Harman has in mind is very similar to the notion of the object of experience Valberg (1992) defines in terms of what is *present* in experience: “By an ‘object of experience’ we shall mean something present in experience: something which is right *there*, available for us to pick out or focus on, and refer to demonstratively” (Valberg 1992, 21–22). Of course, however, Harman rejects the conclusion of what Valberg calls “the problematic reasoning,” that “what is present in experience (present to us, present) is always an internal object; that external objects are never actually present to us” (Valberg 1992, 19).

that we cannot sensibly talk about redness (and being coffee-cup shaped) being a property of an experience, where by “experience,” we mean mental paint but not experiencing act. If the paint in a painting representing a red object might be red, then it is not clear that experiences qua mental paints (but not qua experiencing acts) representing a red object cannot be red. Secondly, even if it is agreed that there is *no* sense in which redness might be a property of an experience, the allegedly problematic consequence follows only on the condition that mental paint is to be defined as the properties *of an experience* by virtue of which the experience is of the object that it is of. This problem disappears simply by devising a technical term such as “a property *featuring in* an experience” and define mental paint as the properties *featuring in* an experience by virtue of which the experience is of the object that it is of, where a property features in an experience just in case that property is either a property of the experience or a property of one of its (same-level) “components” (such as its object). A definition along these lines adequately captures all the relevant contours of the dialectic here without simply begging the question against the identity thesis. (The anti-functionalist, for instance, is to be construed, on this interpretation, as claiming that there are some properties *featuring in* [but *not necessarily of*] an experience but left out by the functionalist story.) This being so, however, Harman’s original distinction between properties of the object of experience and properties of the experience of an object, where mental paint belongs to the latter, has deservedly taken its hold in the literature, and it is best, I believe, to keep it as it is because it is crisp and sharp; however, it must be noted that the sort of *of-ness* deployed in the articulation of mental paint is better taken with a grain of salt.

The final objection I want to consider concerns the bearing of the identity thesis on Harman’s defense of functionalism against the argument from the intrinsic properties of experience. It might be argued that the truth of the identity thesis does not threaten Harman’s defense of functionalism. After all, it might be claimed, Harman holds that, given its concern with the nature of experience but not with the nature of the objects of experience, properties of the object of experience pose no problem for functionalism; and, if mental paint is identical to properties of the object of experience, as the identity thesis claims, then a proper response available to Harman is

simply that as with properties of the object of experience, mental paint poses no problem for functionalism.

However, this objection misapprehends the structure of the dialectic between Harman the functionalist and the anti-functionalist. The anti-functionalist originally argues that introspection reveals that there are some intrinsic properties of experience, properties which, given its exclusive concern with relational features of experience, functionalism is not in a position to account for. Harman's response is that the anti-functionalist argument commits the fallacy of confusing mental paint with properties of the object of experience. According to Harman, introspection reveals properties of the object of experience, which is not to be confused with mental paint, properties of the experience by virtue of which it represents what it does. However, the point is that if the identity thesis is true, then properties of the object of experience are properties of the experience by virtue of which it represents what it does: the identification of mental paint with properties of the object of experience entails that pace Harman, the anti-functionalist argument cannot be rightly accused of illegitimately conflating the properties of the experience of an object with the properties of the object of an experience. In other words, if the identity thesis is true, then among the properties introspecting reveals about a given experience are *its* intrinsic properties, and Harman's response is thereby undermined.<sup>19</sup>

---

<sup>19</sup> The central aim of this paper is to argue that Harman fails to defend *functionalism* against an objection from the intrinsic properties of experience. Still, I wish to make a number of points about how the truth of the identity thesis bears on *representationalism*, the view that the phenomenal character of a given experience ("what it is like" to have that experience) is *exhausted* (Block 1995, 20) by its representational content, especially given that the philosophical focus with respect to experiential transparency has shifted away from functionalism and towards representationalism. There are two broadly distinct alternatives concerning the representationalist conception of mental paint. On one alternative, the sense the representationalist attributes to 'mental paint' is the same as Harman's (i.e. properties by virtue of which an experience represents what it does), and the representationalist argues that the intrinsic qualities that we are aware of when we introspect our experiences are only those properties of the object of experience, which figure in the "content" of experience, but that we are never aware of mental paint (in Harman's sense). If the identity thesis is true, then the representationalist is mistaken to think that we are

I want to stress this point because it is very important. Harman's objection to the anti-functionalists' arguments from the intrinsic qualities of experiences is captured by the idea that when one attends to one's experience of the redness of an apple, one is not aware of an intrinsic quality of the experience but of an intrinsic quality of the apple. Harman argues that since redness is not (and is not experienced as) a quality of the experience, the fact that a functional definition of the visual experience of redness does not capture that quality does not detract from the truth or plausibility of that definition. However, if mental paint by definition consists in properties of the experience by virtue of which it represents what it does, then given that, as I have argued, the identity thesis is true, redness is not only a property of the object of experience but also a property of (or, using the jargon introduced above, featuring in) the experience by virtue of which it represents what it does. So, the anti-functionalists' defending the argument above is to be interpreted as claiming that redness, that very feature of the object of experience, is a feature of my experience that resists a functionalist treatment. The anti-functionalists need not deny and indeed might fully embrace Harman's point that redness is a quality of the object of my experience but still consistently, and plausibly, argue that that point hardly defuses the objection that redness is a quality of (or in) my experience that is left out by a functionalist account. The upshot is that Harman's distinction between properties of the object of experience and mental paint does not help the functionalist because the anti-functionalists cannot be confusing the two, given their identity: Harman's attempt to defuse the anti-functionalists' arguments from the qualities of experiences misfires and is therefore unsuccessful.

---

never aware of mental paint. On another alternative, the sense the representationalist attributes to 'mental paint' is different from Harman's and the representationalist argues that we are never aware of mental paint in *that* (different) sense. What bearing the truth of the identity thesis has on representationalism depends on how that sense is specified, and the possibility that the identity thesis is consistent with representationalism is left open. So, the answer to the question regarding the bearing of the identity thesis on representationalism requires the clarification of the notion of mental paint as the representationalist conceives it, a task which falls beyond the scope of this paper and must await another occasion. (Thanks to an anonymous reviewer for raising this worry.)



## References

- Allen, Keith. 2013. "Blur." *Philosophical Studies* 162 (2): 257–73.  
<https://doi.org/10.1007/s11098-011-9758-6>
- Block, Ned. 1995. "Mental Paint and Mental Latex." *Philosophical Issues* 7: 19–49.  
<https://doi.org/10.2307/1522889>
- Block, Ned. 2010. "Attention and Mental Paint." *Philosophical Issues* 20 (1): 23–63. <https://doi.org/10.1111/j.1533-6077.2010.00177.x>
- Boghossian, Pauk, and David Velleman. 1989. "Color as a Secondary Quality." *Mind* 98 (389): 81–103. <https://doi.org/10.1093/mind/XCVIII.389.81>
- Broad, C. D. 1952. "Some Elementary Reflections on Sense-Perception." *Philosophy* 27 (100): 3–17. <https://doi.org/10.1017/S0031819100019732>
- Byrne, Alex. 2001. "Intentionalism Defended." *The Philosophical Review* 110 (2): 199–240. <https://doi.org/10.2307/2693675>
- Chisholm, Roderick. 1966. *Theory of Knowledge*. Englewood Cliffs: Prentice Hall.
- Dretske, Fred. 1995. *Naturalizing the Mind*. Cambridge: MIT Press.
- Goldman, Alvin. 1993. "Consciousness, Folk Psychology, and Cognitive Science." *Consciousness and Cognition* 2 (4): 364–82.  
<https://doi.org/10.1006/ccog.1993.1030>
- Grice, Paul. 1989. "Some Remarks about the Senses." In *Studies in the Way of Words*, 248–68. Cambridge: Harvard University Press.
- Harman, Gilbert. 1990. "The Intrinsic Quality of Experience." *Philosophical Perspectives* 4: 31–52. <https://doi.org/10.2307/2214186>
- Harman, Gilbert. 1995. "Qualia and Color Concepts." *Philosophical Issues* 7: 75–79. <https://doi.org/10.2307/1522893>
- James, William. 1904. "Does 'Consciousness' Exist?" *The Journal of Philosophy, Psychology, and Scientific Methods* 1 (18): 477–91.
- Kind, Amy. 2003. "What's So Transparent about Transparency?" *Philosophical Studies* 115 (3): 225–44. <https://doi.org/10.1023/A:1025124607332>
- Lehrer, Keith. 2002. "Self-Presentation, Representation, and the Self." *Philosophy and Phenomenological Research* 64 (2): 412–30.  
<https://doi.org/10.1111/j.1933-1592.2002.tb00012.x>
- Moore, George Edward 1903. "The Refutation of Idealism." *Mind* 12 (48): 433–53.  
<https://doi.org/10.1093/mind/XII.4.433>
- Pace, Michael. 2007. "Blurred Vision and the Transparency of Experience." *Pacific Philosophical Quarterly* 88 (3): 283–309. <https://doi.org/10.1111/j.1468-0114.2007.00296.x>
- Searle, John. 1992. *The Rediscovery of the Mind*. Cambridge: MIT Press.
- Smith, A. D. 2008. "Translucent Experiences." *Philosophical Studies* 140 (2): 197–212. <https://doi.org/10.1007/s11098-007-9137-5>

- 
- Sundström, Pär. 2018. “How Physicalists Can—and Cannot—Explain the Seeming “Absurdity” of Physicalism.” *Philosophy and Phenomenological Research* 97 (3): 681–703. <https://doi.org/10.1111/phpr.12394>
- Tye, Michael. 1992. “Visual Qualia and Visual Content.” In *The Contents of Experience*, edited by Tim Crane, 158–76. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511554582.008>
- Tye, Michael. 1995. *Ten Problems of Consciousness*. Cambridge: MIT Press.
- Tye, Michael. 2002. “Blurry Images, Double Vision, and Other Oddities.” In *Consciousness, Color, and Content*, 69–97. Cambridge: MIT Press.
- Valberg, J. J. 1992. “The Puzzle of Experience.” In *The Contents of Experience*, edited by Tim Crane, 18–46. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511554582.003>