**The transparency of experience and the neuroscience of attention[[1]](#footnote-2)\***

(Forthcoming in *Synthese*. https://doi.org/10.1007/s11229-019-02366-8)

Assaf Weksler (Department of Philosophy & Psyphas Program in Psychology and Philosophy, University of Haifa, Haifa, Israel. Email: assaf.weksler@gmail.com)

Hilla Jacobson (Department of Philosophy & Department of Cognitive Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel. Email: hilla.jacobson@mail.huji.ac.il)

Zohar Z. Bronfman (The Cohn Institute for the History and Philosophy of Science and Ideas & School of Psychology, Tel Aviv University, Tel Aviv, Israel. Email: zoharbronfman@gmail.com)

It seems that you can direct your attention to the representational content of your experience. For example, you can attend to the color that your experience represents an object as having. Can you turn your attention towards intrinsic properties of your experience that carry its representational content, i.e., to “mental paint?” According to the thesis of transparency, you cannot. The controversy about transparency is central to the debate between representationalists, who hold that phenomenal character is determined by representational content (Byrne 2001; Byrne & Tye 2006; Dretske 1995; Harman 1990; Mendelovici 2014; Tye 1995, 2002, 2014), and phenomenists, who hold that phenomenal character is an intrinsic property of experience that can vary independently of representational content (Block 2003; Burge 2003; Kind 2003, 2008; Loar 2003; Papineau 2014; Prinz 2012). So far, arguments for and against transparency were conducted from the armchair, relying mainly on introspective observations (Block 2003; Burge 2003; Harman 1990; Kind 2003, 2008; Loar 2003; Tye 1995, 2002, 2014). In this paper, by examining neuroscientific accounts of attention, we argue that it is not possible to attend to mental paint; in other words, we argue in favor of transparency.

In Section 1.1, we introduce the key notions of mental paint, vehicle properties and representational content by reviewing the debate between phenomenism and representationalism. This is a necessary background for the thesis of transparency that we spell out in Section 1.2. In Section 2 we survey various neuroscientific accounts of first-order (FO) sensory attention, on which attending is a matter of modulating the neuronal basis of sensory representations. For each account we argue that it not applicable to mental paint. In Section 3, we discuss higher-order (HO) accounts of attention to mental paint. In Section 3.1, we consider HO *cognitive* attention to mental paint, which is a kind of thinking about mental paint, arguing that it is irrelevant to transparency. In Section 3.2, we discuss HO *sensory* attention to mental paint. On this suggestion, there are HO perceptions (HOPs) of mental paint, which are attentionally modulated. We argue that this suggestion has apparently incoherent consequences. Sensory (both FO and HO) and (HO) cognitive forms of attention appear to exhaust the space of possible views of attention to mental paint. Thus, the overall argument can be presented in the form of a trilemma: attention to mental paint is either FO sensory, HO cognitive, or HO sensory attention. Each of these options have serious problems. The result is an argument by elimination in favor of the transparency thesis. In Section 4, we argue that extant philosophical theories of attention do not provide a third conception of attention, over and above the sensory and cognitive varieties.

1. **Background**
   1. **Representational content, vehicle properties, and phenomenal character**

Suppose you perceptually experience the flowers of an almond tree. Your experience represents the flowers' petals as white. In other words, your experience has the representational content, “ahead are white petals.” Additionally, there is something it is like to have your experience, involving white\*[[2]](#footnote-3) phenomenal character. According to phenomenism (Block 2003; Kind 2003, 2008; Loar 2003; Papineau 2014; Prinz 2012), these two aspects, namely representational content and phenomenal character, can vary independently. On this view, representational content is extrinsic – i.e., determined by causal relations with the environment (see below) – whereas phenomenal character is intrinsic. Block (2003) defines two different kinds of intrinsic phenomenal character, to which he refers as “mental paint” and “mental oil.” Mental paint is defined as carrying the representational content of experience. As Papineau (2014) puts it, mental paint properties “are the simple intrinsic properties whose instantiation in fact represents to subjects that things are thus and so […]” (p. 21), in virtue of said causal relations. In contrast, mental oil does not represent anything to subjects – it does not carry the representational content of experience. The view we target in this paper (clarified in Section 1.2) is that it is possible to shift attention *from the representational content of experience to the intrinsic phenomenal property that carries it – i.e., to mental paint*. Since mental oil does not carry the representational content of experience, we set it aside.

Let us introduce a simplified neuroscientific take on phenomenism, which we will use throughout the paper. Aside from the simplifying assumptions we add (we describe their divergence from reality below), the phenomenistic model we present is not novel: it is explicitly endorsed by leading phenomenists such as Block (2003), Loar (2003), Papineau (2014) and Prinz (2012). Let us assume that in area V4 of the visual cortex there is a single neuron that can represent whiteness. Let's assume that this neuron underlies our conscious visual experience of whiteness. It can represent whiteness in virtue of *tracking* whiteness (see Mendelovici and Bourget 2014 for a review of tracking accounts of intentionality; such accounts are assumed by the aforementioned phenomenists). Tracking is an "externalistic" relation, since it holds between the neuron and properties instantiated in the external world. Without loss of generality, we will assume that “tracking” means causal co-variation in normal conditions. That is, the neuron tracks whiteness in the sense that, normally, it fires at a high frequency in response to whiteness but not to other colors. This implies that the representational property in question, that of representing whiteness, is extrinsic. Now, when looking at the almond tree, the neuron fires at a high frequency, thereby representing whiteness. In other words, its representational content includes whiteness. This representational content of the neuron determines the representational content of the experience.

The neuron’s firing frequency is an intrinsic property of the neuron. We will call it "a neuronal vehicle property." As described in the preceding paragraph, neuronal vehicle properties, in combination with the aforementioned externalistic tracking relation, determine the representational content the neuron has at a given moment (equivalently, they determine the representational properties of the neuron). We will sometimes say for short that the neuronal vehicle properties “carry” the representational content.

On the simplified neuroscientific take on phenomenism we now consider, some of these intrinsic, neuronal vehicle properties—but not the extrinsic, tracking properties—determine the phenomenal character of experience. They determine, for example, that one enjoys white\* phenomenal character rather than black\*. This determination is entirely “local.” That is, externalistic, tracking, representational factors do not contribute to it (Block 2003; Papineau 2014; Prinz 2012). Consequently, the phenomenal character in question counts as "mental paint."

It is crucial to observe that on this phenomenist model, a single neuronal vehicle property serves a *dual role*: first, it determines phenomenal character (mental paint) by itself (i.e., regardless of the tracking relation). Second, in combination with the tracking relation, the vehicle property determines the representational content of the neuron, which determines the representational content of the experience (equivalently, it determines the representational property of the experience). The reason why a single neuronal vehicle property – rather than two separate vehicle properties – plays both roles can be traced back to the definition of mental paint. On that definition, mental paint is an intrinsic phenomenal character that, via tracking relations, represents to subjects a worldly property. In other words, mental paint carries the representational content of experience. Mental paint tracks a worldly property in virtue of the fact that its (mental paint’s) neuronal basis – the neuronal vehicle property that determines it – tracks this worldly property. Thus, this neuronal vehicle property serves the dual role of determining mental paint and (via tracking relations) determining which property this mental paint represents to the subject. In other words, the representational content that this neuronal vehicle property carries determines the representational content that mental paint carries.

Our model is simplified since it involves a *single* neuron in a *specific* brain area (V4) and its firing *frequency*. In reality, the neuronal vehicle of an experience of whiteness is much more complex. For example, according to Prinz (2012), the vehicle is a population of neurons, and it represents whiteness via a combination of population coding and temporal coding. While Prinz holds that the relevant population is in V4, others (e.g., Dehaene & Naccache 2001) would claim that it is distributed across various brain regions. We use the simplified model in order to make the discussion below easy to follow; our arguments generalize straightforwardly to the more complex models.

So far, we have described phenomenism. Let us turn to its competitor, representationalism (Byrne 2001; Byrne & Tye 2006; Dretske 1995; Harman 1990; Mendelovici 2014; Tye 1995, 2002, 2014). On this view, phenomenal character is determined by the representational content of experience. Return to the almond tree example. Representationalist grant that you enjoy white\* phenomenal character when you look at the tree. But they hold that this phenomenal character is determined by the fact that your experience represents the petals as white: that a neuron in V4 represents whiteness (which determines that the experience represents whiteness) determines that the experience has white\* phenomenal character, rather than black\*, say.

In sum, both camps in the dispute grant that experience has both representational content and phenomenal character. The dispute concerns their relationship. Phenomenists hold that phenomenal character is mental paint – an intrinsic, conscious property distinct from representational content. Mental paint is entirely determined by intrinsic, vehicle properties of the neuron underlying it, regardless of tracking relations. Representationalists hold that phenomenal character is determined by the representational content of the experience, which is itself determined by the representational content of the neuron underlying it. In turn, this representational content is determined by a combination of neuronal vehicle properties and tracking relations.

* 1. **The thesis of transparency**

The transparency thesis concerns attention in ordinary perceptual experiences (not, e.g., the experience of phosphenes or afterimages). Philosophers argue that it challenges phenomenism and supports representationalism (Harman 1990; Tye 1995, 2002, 2014). There are two versions of the thesis in the literature, one phenomenological, the other ontological. The first is a thesis about what attention *seems* to be directed to, upon phenomenological reflection. The second is a thesis about what attention is *in fact* directed to. Typically, philosophers support the ontological version on the basis of the phenomenological version, combined with an additional premise, namely that introspection does not involve massive error (Tye 2002, p. 139). In contrast, we will tackle the ontological version *directly* on grounds of theoretical considerations. For this reason, we set aside the phenomenological version, and will henceforth use “transparency” only in the ontological sense.

According to the thesis of transparency, when you introspect your perceptual experience, your attention always “goes through it” to objects in the world and their properties (objects and properties the experience represents). In other words, your attention can only be directed to the representational content of your experience. For example, looking at the almond tree, you focus your attention on the white color that the petals appear to have (i.e., the white color that your experience represents the petals as having). Next, you try to turn your attention “inwards,” to white\* mental paint. According to the thesis of transparency, you cannot do so.

Transparency is supposed to be evidence for representationalism. The basic thought is that if you enjoy mental paint, you should be able to attend to it. But you can't (as per transparency), so there is no mental paint (see Kind 2003, 2008, 2010; Loar 2003; Tye 2014).

Ned Block (2003), Amy Kind (2003, 2008) and Brian Loar (2003), who hold that mental paint exists, deny the transparency thesis. They grant that *usually* when you introspect your experience, your attention "goes through" to its representational content – i.e., it 'ignores' mental paint and 'focuses' on what it represents; but they think it is nevertheless *possible* to turn attention "inwards" to mental paint (Block 2003, pp. 173, 178-9; Kind 2003, p. 230, 2008, p. 287; Loar 2003, pp. 77, 94). Kind and Loar emphasize that attending to mental paint is *difficult*.

In the remainder of this paper we will assume, with Block, Kind and Loar, that our attention is typically directed to the representational content of experience, and argue, against them, that it is not possible to direct attention to mental paint. From this the thesis of transparency follows. Note that we do not aim to refute phenomenism, vindicate representationalism, or even defend the argument from transparency. We only purport to support the transparency thesis itself. Because the opponents of transparency that we target (namely Block, Kind and Loar) assume that attention is typically directed to the representational content of experience, it is legitimate for us to assume it in our argument as well. However, we would like to emphasize that one can resist the move from the impossibility of attentional shifts (between representational content and mental paint) to transparency by denying this assumption[[3]](#footnote-4).

1. **Attention as modulation of first-order sensory representations**

In this section we examine accounts of first-order (FO) sensory attention. These are standard neuroscientific accounts of attention, on which attending is a matter of modulating the neuronal basis of FO sensory representations. It is important to note that most (if not all) of these accounts are not in competition, because they concern different findings about how attention works. Hence, they can be taken to concern different aspects of attention. We argue that these accounts of attention are inapplicable to mental paint. Note that these accounts are of attention *itself*, not of *orienting* (controlling) attention (for a philosophical discussion of this distinction see Watzl 2017, Ch. 6-7).

In the remainder of this section, unless otherwise noted, talk of “attention” will refer to *FO sensory attention*.

The arguments we develop in this section conform roughly to the following schema, which we clarify below:

1. A single (active) neuron N underlies a given visual experience. This neuron has both representational content and a vehicle property. The vehicle property determines mental paint by itself and, in combination with externalistic, tracking factors, the vehicle property determines the representational content of the neuron (which determines the representational content of experience). (premise)
2. On a given scientific sense of attention S, to attend to the *representational* *content* of the experience is to modulate in way W a feature of type F of the neuron underlying this representational content, namely N. (premise)
3. Suppose, *for the sake of* *reductio*, that we can attend to the mental paint of the experience, in sense S, i.e., in the sense of modulating in way W a feature of type F of the neuron underlying this mental paint.
4. We can attend to the mental paint of the experience in the sense of modulating in way W a feature of type F of the neuron N. (from 1, 3)
5. There are no two features of type F of neuron N, one associated with the representational content of experience and another with mental paint. (premise)
6. Therefore, attending to the representational content of experience is *ipso facto* attending to mental paint. (from 2, 4, 5)
7. But, according to phenomenism, attention is usually directed to the representational content of experience, and *not* to mental paint. (premise)
8. There is a contradiction between 6 and 7.
9. Thus, premise 3 is false (argument by reductio), that is, attention in sense S is not applicable to mental paint.

Let us explain each step of the argument. Premise 1 describes the simplified phenomenistic model from Section 1.1. As we have explained, apart from the simplifying assumptions (intended to make the discussion easier to follow), this model is already accepted by leading phenomenists. Premise 2 is based on the neuroscience of attention. All the neuroscientific accounts of attention we survey are of *ordinary*, everyday attention, which all parties (in the transparency debate in question) assume to be attention to the *representational content of experience*. Each account of attention provides values for S, W and F. For example, in Section 2.1, we discuss the account of attending as RF (receptive field) shrinking. When inserted into the argument schema, S becomes “attending as RF shrinking”, F becomes “RF” and W becomes “shrinking.” Premise 3 simply applies a given neuroscientific account of attention to mental paint. Interim conclusion 4 might seem sufficient to derive a contradiction, but it is not. Consider again RF shrinkage. Interim conclusion 4 says that attention to mental paint is a matter of shrinking the RF of N. This does not yet generate a contradiction because there might be two separate RFs for N, one associated with the representational content of experience and another associated with mental paint. This possibility is rejected by premise 5 (when the value of F is RF), which is defended separately for each value of F we survey below. Interim conclusion 6 is self-explanatory. Premise 7 was explained in Section 1.2.

Considering premise 3, one might object that it is *a priori* implausible that attending to mental paint is a matter of modulating its neuronal basis, i.e., modulating an *FO* state. For, attention to mental paint is *introspective*, and so it should be an *HO* attention directed at the FO state. The objector would conclude that we do not need a complex, empirically-based argument (the rest of Section 2) to rule this FO approach out, and we can directly move to discussions of HO accounts of attention to mental paint (Section 3). The objection is wrongheaded. Suppose you are looking at the ocean. You attend to the blueness and judge that it is instantiated by the ocean – an exteroceptive judgment. This is an instance of FO sensory attention. The blueness you are attending to is a component of the representational content of your experience. While continuing to attend to the blueness, in an FO sensory manner, you can also judge that it *appears* to be instantiated by the ocean, or equivalently, that *your experience* represents the ocean as blue, or that *the representational content* of your experience involves blueness. These are introspective judgments. Thus, the same FO sensory attention to blueness, the same "highlighting" of it, is a basis for both exteroceptive and introspective judgments (Tye 2002, p, 139). Thus, there is nothing *a priori* implausible in the suggestion that attention to mental paint, which is a basis for introspective judgments about it, is FO sensory attention.

* 1. **Shrinking RF**

Neurons in the visual cortex have a receptive field (RF), which is the area in the retina they respond to. Consider Figure 1. Suppose a single neuron in the visual cortex receives input (via neuronal intermediaries) from the circled area. This neuron, let us suppose, is “tuned” to vertical lines, so that it fires at 60 s/s (spikes per second) when there is a vertical line in its RF and at 10 s/s when there is a horizontal line there. In Figure 1A, the neuron’s RF includes both a vertical and a horizontal line. The neuron responds by firing at 35 s/s, which is the average of 60 s/s and 10 s/s (Reynolds, Chelazzi, and Desimone 1999). According to the RF shrinkage account of attention, to attend to the representational content of the experience that this neuron underlies is to shrink the neuron’s RF (Moran & Desimone 1985; Womelsdorf et al. 2008; for a philosophical review see Wu 2014, Section 2.4.2). For example, attending to the location of the left line is a matter of shrinking the RF of the neuron, as shown in Figure 1B. That is, the neuron now receives input only from the left line. Consequently, it fires at 60 s/s. This makes the neuronal activity more informative regarding the left line. Of course, the drawing in Figure 1 is an oversimplification. For an actual single neuron recording of RF shrinkage see Womelsdorf et al. (2008).

The preceding account is of attention to the representational content of experience. Is this account applicable to mental paint?

|  |
| --- |
|  |
| **Figure 1** An example of RF shrinkage. In state A, the RF of a certain neurons includes both lines. In state B, the RF shrinks and includes only the left line. |

The vehicle property of the neuron determines mental paint, according to phenomenism. On the present model, attention to the representational content of a given experience is a matter of shrinking the RF of the neuron whose vehicle property carries this representational content, call this neuron N. We are trying to apply this model to mental paint. So, attention to mental paint should be a matter of shrinking the RF of the neuron whose vehicle property determines it, which is the very same neuron N. Now, even if there are two (or more) RFs for the neuron, it is not the case that *one is for its representational content* (which determines the representational content of experience) *and another is for its intrinsic, vehicle property* (which determines mental paint). Here is why. The RF of a neuron is the area in the retina the neuron responds to. What could it mean to say that there is a second RF? It could only mean that the neuron responds to an additional area in the retina, beyond the first RF. But if that is the case, then this second RF is, like the first RF, *of the neuron simpliciter*: it does not make sense to say that the second RF is *of* the vehicle property, whereas the first is *of* the representational content.

Thus, even if there are two (or more) RFs for the neuron, it is not the case that attention to the representational content of experience shrinks one of them, whereas attention to mental paint shrinks the other. So, if attending to mental paint were a matter of shrinking a neuron's RF(s), then it would shrink the same RF(s) that attention to the representational content of experience shrinks. But then attending to the representational content of experience would *ipso facto* be attending to mental paint. This result is inconsistent with the claim that usually our attention is directed to the representational content of experience *and not* to mental paint. Thus, the claim that we can attend to mental paint is incompatible with the neuroscientific understanding of attention in terms of RF shrinkage.

* 1. **Raising signal-to-noise ratio**

Recall that we apply roughly the same argument schema for each neuroscientific account of attention. We have seen in Section 2.1 how the argument works for RF shrinkage. For the remaining accounts, we will focus mainly on issues specific to each and will not repeat familiar stages of the argument schema.

Neurons often fire spontaneously. That spontaneous activity is referred to as “noise.” Aside from spontaneous firing, neurons also fire in response to their “preferred” input according to their tuning curve. That response is called “signal.” Consider the grey line in Figure 2. The amplitude, 40 s/s, is the firing rate of a neuron responding to its preferred stimulus, namely this - - orientation. The asymptote, 18 s/s, is the noise (it is there regardless of the stimuli). This noise is also present in the preferred stimulus condition, of course. The higher the signal-to-noise ratio, the more accurate the processing of the information is.

On the account of attending as increasing signal-to-noise ratio, to attend to the representational content of the experience that this neuron underlies is to increase the neuron’s signal-to-noise ratio (Noudoost et al. 2010; for a philosophical discussion see Wu 2014, Section 2.4.1). Consider the black line from Figure 2. This is the same neuron now attentionally modulated. Its amplitude is 57 s/s, meaning the signal is better, while the asymptote (the noise) remains roughly 18 s/s.

|  |
| --- |
|  |
| **Figure 2** An illustration of attention as increasing the signal-to-noise ratio of a neuron. The grey line depicts the firing frequency of the neuron as it responds to various unattended stimuli. The black line depicts the firing frequency of the same neuron as it responds to various attended stimuli. Reprinted with permission from McAdams & Maunsell (1999): Society for Neuroscience. |

This account of attention as increasing the signal-to-noise ratio of an (active) neuron is an account of attention to the representational content of experience. Is this account applicable to mental paint?

Recall that mental paint is determined by a neuronal vehicle property, i.e., a neuron’s firing at a certain rate. As explained before, this neuronal vehicle property, in combination with externalistic (tracking) factors, determines the representational content of experience. On the present account, attending to the representational content of experience is a matter of increasing the signal-to-noise ratio of this neuron. Thus, attending to mental paint, on the present account, should be a matter of increasing the signal-to-noise ratio *of the same neuron*. For this to work, there should be *two* signal-to-noise ratios for the same neuron, one associated with its representational content (which determines the representational content of experience) and one with its intrinsic, vehicle property (which determines mental paint), regardless of the neuron’s representational content. The former signal-to-noise ratio should be (at least partly) associated with the stimuli that the neuron tracks, whereas the latter signal-to-noise ratio should be associated with the neuron’s intrinsic, vehicle property, *regardless of the stimuli that the neuron tracks*. But the latter notion is incoherent: without the stimuli it becomes impossible to distinguish between *signal*, namely firing rates that are caused by preferred stimuli, and *noise*, namely firing rates that are caused by some other factor. Thus, the neuron does not have two signal-to-noise ratios, one for its representational content, and another for its intrinsic vehicle property. From this point the argument proceeds in accordance with the argument schema and reaches a contradiction. Thus, the claim that we can attend to mental paint is incompatible with the neuroscientific understanding of attention in terms of increasing signal-to-noise ratio.

* 1. **Shielding from decay**

Suppose a neuron, N, in V4 is tuned to the color green. You look at a green circle and consequently N fires at a high rate. Then the stimulus disappears. Consequently, the activity of N gradually decays. On the account of attending as shielding from decay, to attend to the green circle is to slow down this decay. To illustrate, consider the following experiment (Figure 3).

|  |
| --- |
|  |
| **Figure 3** A retro cue experiment. Subjects first see a series of colored circles. After an interval, a cue is presented that either correctly predicts the location of the subsequent probe (valid cue) or fails to predict it (invalid cue). When the probe appears, the subjects need to make a same/different judgment. Reprinted with permission from Shepherdson et al. (2018): American Psychological Association. |

Subjects look at colored circles arranged in a circle. One second after stimulus offset, a cue ("retro cue", RC) is presented. After an additional second or two a probe color appears where the cue was (valid cue) or in some other location (invalid cue), and subjects are asked to make a same/different judgment, comparing the probe with the previously seen color at that location. The typical results in experiments of this sort are that performance after the cue-onset remains constant for a couple of seconds if the cue is valid, but drops rapidly if the cue is invalid (Figure 4). This means that the neuron underlying the sensory representation of the color at the cued location – i.e., the neuron underlying the attended representational content of experience – is maintained in an active state, shielded from decay, whereas the neurons underlying the representations of colors at different locations decay rapidly (Gazzaley & Nobre 2012; Lepsien & Nobre 2006).

|  |
| --- |
|  |
| **Figure 4** Results of a typical retro cue experiment. The arrow depicts the point in time in which the retro cue appears. The lines signify performance as a function of the interval between the offset of the memory array and the onset of the probe. If the cue is valid (green line), performance (after the appearance of the cue) is constant – the representation does not decay. If the cue is invalid (red line), performance drops rapidly – the representation decays. Reprinted with permission from Souza & Oberauer (2016): Springer Nature. |

The present account is of attention to the representational content of experience. Does this account apply to mental paint?

Consider an (active) neuron underlying an experience. The intrinsic, vehicle property of the neuron (e.g., its firing rate) determines mental paint, according to phenomenism, and it also carries representational content. Thus, one shields the representational content of the neuron—and hence of the experience—from decay *by* shielding this vehicle property from decay. So, attending to the representational content (of experience) is a matter of shielding it from decay, but it inevitably involves shielding said vehicle property (of the neuron) from decay as well. Now suppose for the sake of *reductio* that attending to mental paint amounts to shielding the vehicle property that determines it from decay. It follows that attending to the representational content of experience is *ipso facto* attending to mental paint. Via steps familiar from the argument schema, we reach a contradiction. Thus, the claim that we can attend to mental paint is incompatible with the neuroscientific understanding of attention in terms of shielding a representation from decay.

* 1. **Suppressing distractors**

Another central function of attention concerns ignoring, or suppressing, task-irrelevant distractors. Consider for example the study by de Fockert et al. (2001). Participants see written names superimposed on faces. They need to rapidly report whether the name is of a singer or a politician. The faces are distractors – some faces are incongruent with the names, and the subjects need to disregard them. It is difficult to do so: the faces are automatically processed and categorized, and hence the subjects need to suppress the representation of faces, in a top-down manner.

The phenomenon of suppressing distractors has been studied by neuroscience. Consider a study by Rissman et al. (2009). In Figure 5, the blue line depicts the activity in a brain area representing scenes, during a task that requires remembering faces and *ignoring* distractor scenes. On the low load condition said activity is suppressed, relative to high load condition (high load hampers this suppression of distractors). On the present account of attention, suppressing the distractor scenes *is* attending to the task-relevant stimuli, namely faces.

The present account of attention—as suppressing the neural activity of distractors—is an account of attention to the representational content of experience. Is this account applicable to mental paint?

|  |
| --- |
|  |
| **Figure 5** Example of distractor suppression as studied by neuroscience. The lines depict activity of representations of scenes in the visual cortex, in different tasks: remembering faces and ignoring scenes (blue line), passively viewing scenes and faces (yellow line), and remembering scenes while ignoring faces (red lines). All tasks were done while either remembering ordered digits (low verbal load) or remembering random digits (high verbal load). The crucial finding is captured by the different blue lines in the low load vs. high load conditions. Reprinted with permission from Rissman et al. (2009): Elsevier. |

On the distractor suppression account of attention, attending to mental paint amounts to suppressing the representational content of a certain (active) neuron (which determines the representational content of experience), without suppressing its intrinsic, vehicle property (which determines mental paint). But that is impossible: one suppresses the representational content of a neuron *by* suppressing the activity of its vehicle property, which carries that representational content. For example, prior to the suppression, the neuron fires at 40 s/s, and as a result of suppression, it fires at 20 s/s. Its representational content is now "weaker." But this is achieved *by* downregulating the intrinsic property of the neuron, the vehicle property, namely the *firing frequency*. Thus, it is not possible to suppress the representational content without suppressing the vehicle property that carries it. The claim that we can attend to mental paint is, therefore, incompatible with the neuroscientific understanding of attention in terms of suppression of neural activity of distractors.

* 1. **Allocating resources**

Neural representations necessitate biological resources and are affected by the amount of resources available to them. A resource could be an ATP energy molecule, a neurotransmitter, a sodium ion, etc. According to the resource-allocation account of attention (Bays et al. 2009; Bays 2015), at least some of the resources are scarce and limited, and mental representations compete for them. The available resources must be optimally allocated between the competing representations to allow for efficient processing. Attending is therefore seen as a matter of allocating the resources to certain representations over others, in order to meet task demands.

This account is of attention to the representational content of experience. Is this account applicable to mental paint?

Applying this account of attention to mental paint amounts to the idea of allocating resources to the neuron underlying it. But attending to the representational content of experience already involves allocating resources to the same neuron. Moreover, it is implausible to hold that some of the resources are “for” the representational content of experience (or of the neuron) whereas others are “for” the mental paint (or the intrinsic, neuronal vehicle property that determines it). Resources are allocated to the neuron *simpliciter*. As a consequence, firing rates increase, for example, which is a change in the vehicle property, and these firing rates carry information, or have representational content. Thus, attending to the representational content of experience allocates the same resources to the same neuron as attending to mental paint. From this point we reach a contradiction in the familiar way. Thus, the claim that we can attend to mental paint is incompatible with the neuroscientific understanding of attention in terms of resource allocation.

* 1. **Interim summary**

When we try to apply neuroscientific accounts of first-order (FO) sensory attention to mental paint, we reach a contradiction (for each account). This provides reason to hold that we can't sensorily attend to mental paint, in a first-order sense. But perhaps there is a different, higher-order (HO) conception of attention that is compatible with the idea of attention to mental paint (and with neuroscience)? We address this issue in the next section.

1. **Higher-order accounts of attention to mental paint**
   1. **Higher-order cognitive attention**

We have argued that one cannot FO sensorily attend to mental paint. In the next subsection we will argue that one cannot HO sensorily attend to mental paint. But perhaps one can HO *cognitively* attend to it. What is cognitive attention? In general, cognitively attending to X is some sort of *thinking* about X. For example, in his discussion of transparency, Tye (2014) suggests that cognitive attention is *focused thinking* as opposed to absent-minded thinking. In a trivial sense, we can think (a HO thought) about mental paint, since we debate about it. Phenomenists, for example, think that there is a phenomenal property that remains constant despite shifts in representational content, and representationalists deny that. Both are thinking about mental paint, either affirming or denying its existence. And they can do so in a focused way, i.e. without thinking about something else at the same time. It follows that we can cognitively attend (in a HO way) to mental paint (ibid.)! To emphasize, we can attend to mental paint in that sense, even if it is never instantiated, just like we can attend to the property of being a wizard (i.e., thinking about it in a focused way) even though it is never instantiated. This characterization of cognitive attention deliberately ignores the topic of recognitional phenomenal concepts (Loar 2003). We postpone discussion of such concepts to Section 4.

In the present section we will argue that (HO) cognitive attention is *irrelevant* to transparency (Tye, 2014, defends a similar claim, except he *assumes* from the outset that transparency is true). Our argument is that, first, transparency concerns attention in a sense that allows direct identification of its target. Second, cognitive attention by itself (i.e., without sensory attention) does not allow direct identification of its target.

Attention to mental paint, in a sense that is useful to phenomenists, should allow a subject to directly identify it, assuming the subject possesses the relevant concept, and the mental paint type clearly falls under that concept. For example, it should allow a subject who possesses the concept of blueness\* to directly identify that the mental paint she is enjoying, when looking at the ocean at daytime, is blue\*. By “direct” identification we mean that it does not pass through identification of something else. For example, it is not an inference based on identifying that the experience represents blueness, combined with the theoretical assumption that experiences that represent blueness have blue\* mental paint. While a phenomenist can adopt the view that mental paint is a theoretical posit, if she does so, then arguing against transparency becomes pointless for her. Indeed, Shoemaker (1994) is a phenomenist who holds that mental paint is a theoretical posit, and consequently accepts the thesis of transparency (for discussion see Block 2003, p. 173). The reason why phenomenists such as Block, Kind and Loar present arguments against transparency is precisely because they think that mental paint is *not* a theoretical posit – it is something that we can identify directly (provided we possess the relevant concept, etc. We will omit this qualification henceforth).[[4]](#footnote-5) Thus, for them, transparency means that we cannot attend to mental paint, *in a sense of attention that allows direct identification thereof*. This is the sense of transparency that they argue against, and that we try to defend in the present paper. The existence of a kind of attention to mental paint that does not enable direct identification thereof is, therefore, irrelevant to transparency.

We shall now argue that if you cannot directly identify an item based on FO sensory attention, then cognitive attention (whether FO or HO) will not help you directly identify it. In other words, cognitive attention cannot *replace* FO sensory attention, as enabling direct identification of its target. Call this the *irreplaceability thesis*.

To argue for this thesis, we consider the phenomenon of visual crowding (Block 2013; Intriligator & Cavanagh 2001) as a representative example of a situation in which sensory attention cannot be directed to an item and consequently one cannot directly identify that item on its basis. Consider Figure 6. If you fixate on the red minus sign on the left in each row you cannot directly identify the middle item on the right. The middle item is “crowded” by its flankers. The reason is that the window of FO sensory attention is too coarse-grained: you cannot focus on the middle item (Block 2013).

In this situation, you cannot, on the basis of FO sensory attention, directly identify the middle item. But can cognitive attention help? The answer is negative. Thinking in a focused way that there is a middle item there, wondering what it is, will not help you directly identify that it is (e.g.,) a hamburger. Also, there is no reason to think that a more complex thinking routine would help. Sure, thinking can help you identify the item *indirectly*. For example, by thinking that the pattern “pretzel, hamburger, pizza” occurred before, and by recognizing that what you are seeing has color and texture similar to those three items, you could *infer* that the middle item is a hamburger. But indirect identification is not what we are after. Thus, it seems that cognitive attention cannot replaceFO sensory attention, as enabling direct identification of its target. The irreplaceability thesis appears to be true.

|  |
| --- |
|  |
| **Figure 6** An example of crowding. When looking at the red minus sign on the left in each row, you cannot identify the middle item on the right-hand side. Reprinted with permission from Pelli (2008): Elsevier. Used in Block (2013). |

An objector might claim that the crowding case is special, and that we should not generalize from it, because several studies suggest that it is possible to directly identify an item without FO sensory attention, on the basis of a different kind of attention. For example, one such (representative) study (Reddy et al. 2007) found that participants can directly identify the gender of a face at the periphery even though their focal attention – a form of FO sensory attention – is entirely consumed by a demanding primary task at fixation (for a review of similar studies see van Boxtel et al. 2010). More specifically, when participants *intend* to identify the faces in advance, i.e., when the faces are task-relevant, they can directly identify them without focally attending to them. Reddy et al. (2007) call this an effect of "top-down task-based attention." This form of attention appears to be cognitive, or at least cognitive-like, since intention, like thought, is propositional and non-sensory. Thus, this finding looks like evidence that cognitive attention can replace FO sensory attention, as enabling direct identification of its target.

This is a mistake. The neural basis of "top-down task-based attention" consists of top-down *modulations of FO sensory representations in the visual cortex*, of the sort discussed in Section 2 (Reddy et al. 2007). These modulations do not amount to *cognitive* attention, since cognitive attention is supposed to be an *alternative* to the notion of attention as modulation of sensory representations. After all, the whole point of appealing to cognitive attention in the first place was to avoid the problems from Section 2, which are problems for the suggestion that attending to mental paint is a matter of modulating FO sensory representations. In other words, both focal attention and top-down task-based attention are forms of FO sensory attention. Note that directly identifying the gender of a face (like directly identifying a background scene) requires very little FO sensory attention. This is why one can do it while also performing an (FO sensory) attentionally demanding primary task.

It thus appears that the irreplaceability thesis is true: cognitive attention cannot replace FO sensory attention, as enabling direct identification of its target. Given that we cannot FO-sensorily attend to mental paint, HO cognitive attention to mental paint apparently cannot enable direct identification thereof. It seems, therefore, that the existence of cognitive attention to mental paint is irrelevant to the thesis of transparency.

One might try to resist this conclusion by saying that since the empirical evidence we have presented concerns *exteroceptive* identification, the irreplaceability thesis it supports does not concern *introspective* identification, and therefore it is inapplicable to mental paint. We grant that the studies we have presented concern exteroceptive identification, but, ultimately, they are studies about *FO sensory attention*. In Section 2 we have explained that FO sensory attention is a basis for introspective judgments about the representational content of experience (as well as for exteroceptive judgements about external objects). So while the studies we have surveyed in the present section concern identification of properties of external objects, they trivially apply also to *introspective identification* of properties that the experience represent external objects as having. Thus, the evidence we have presented supports an irreplaceability thesis with respect to introspection and not only with respect to exteroception.

Let us repeat the dialectic of the present subsection, only this time emphasizing *introspective* identification: FO sensory attention highlights a property that figures in the representational content of experience, and so allows us to directly identify it introspectively. According to the irreplaceability thesis, if we cannot FO-sensorily attend to such a property, then HO cognitively attending to it will not help us directly identify it introspectively. We apply the same reasoning to mental paint. Since we cannot highlight it via FO sensory attention (Section 2), HO cognitively attending to it will not help us directly identify it introspectively.

An opponent to the transparency thesis might now suggest that while the irreplaceability thesis applies to introspective identification *of the representational content of experience*, it does not apply to introspective identification *of mental paint*. It is hard to think of an independent (i.e., not *ad hoc*) motivation for thinking so. Remember: the same intrinsic neuronal vehicle property underlies both representational content and mental paint. Whereas representational content is determined in part by externalistic factors, mental paint is determined entirely by said neuronal vehicle property. It is hard to see how this difference might be relevant to the irreplaceability thesis. Consider the other main difference, namely that representational content involves properties the experience represents external objects as having, whereas mental paint involves properties that are in fact instantiated by the experience. This difference too does not look relevant to the irreplaceability thesis.

* 1. **Higher-order sensory attention**

One might propose a higher-order perception (HOP) account of attention to mental paint. The idea is that we can perceive mental paint via an "inner sense", and the neuronal basis of this higher-order perception can be attentionally modulated in the ways surveyed in Section 2. This amounts to higher-order (HO) sensory attention to mental paint.

Against the existence of inner sense (for sensing experiences), Peter Carruthers has argued, roughly, that it must be as computationally complex as outer sense, and that it is implausible that natural selection has produced such a complex system twice over – once for outer sense and once for inner sense (Carruthers 2000, p. 213). Moreover, the leading advocate of HOP in the past two decades, William Lycan, has recently found this argument compelling (Sauret & Lycan 2014). The upshot is that there is already an apparently compelling argument against inner sense—and a fortiori against HO sensory attention to mental paint—in the literature.

In any case, it seems that the notion of HO sensory attention to mental paint has problematic consequences, regardless of computational complexity issues. Consider for example signal-to-noise ratio increase. Paying HO sensory attention to mental paint is supposed to be a matter of increasing the signal-to-noise ratio of the neuron that underlies a HOP directed at this mental paint (we are again assuming it is a single neuron for simplicity). Recall that according to opponents of transparency, attending to mental paint is typically difficult, whereas attending to representational content is typically easy. Given the present proposal, this implies that increasing the signal-to-noise ratio of the neuron underlying said HOP is difficult, whereas increasing the signal-to-noise ratio of the neuron underlying the (FO) experience itself is easy. Consequently, the information you get about mental paint via HO sensory attention will sometimes be quite noisy, in comparison with the information you get about the representational content of the experience via FO sensory attention. Thus, scenarios of the following type would sometimes occur, if there were such a thing as HO sensory attention to mental paint. You look at a burgundy-colored backpack. You FO-sensorily attend to the color-involving representational content of your experience. Consequently, the signal-to-noise ratio of the neuron underlying it is excellent, and so it seems to you, introspectively, that your experience represents the backpack as burgundy. Next, you HO-sensorily attend to the mental paint of your experience. As we have explained, the information you thereby get about the mental paint will be quite noisy, at least sometimes. Because of the noise, your mental paint might introspectively seem to you to be different than it is. For example, the burgundy\* mental paint you are enjoying might introspectively seem to you to be brown\*. However, it is difficult to make sense of such a situation. If it introspectively seems to you that your experience instantiates brown\* mental paint—and hence brown\* phenomenal character—then it should seem to you, introspectively, that your experience represents the backpack as brown. Yet, as we have explained, it introspectively seems to you that your experience represents the backpack as burgundy! The situation looks incoherent. Thus, the notion of HO sensory attention to mental paint appears to have an incoherent consequence.

This concludes the trilemma we present to the proponent of the view that our attention is usually directed to the representational content of experience, yet that it is possible to direct it to mental paint. Attention to mental paint is either first-order or second-order attention. Neuroscience rules out the first option. The second option includes higher-order cognitive attention and higher-order sensory attention. Higher-order cognitive attention to mental paint (in the absence of first-order sensory attention to it) apparently does not enable direct identification thereof. The notion of higher-order sensory attention to mental paint has an apparently incoherent prediction. Therefore, via elimination we get support for the conclusion that we cannot attend to mental paint in a sense that allows direct identification thereof; in other words, we get support for transparency.

1. **Alternative conceptions of attention?**

Wu (2014) claims that attention is selection (of a perceptual or cognitive state) for action. On Watzl’s (2017) view, “attention organizes a subject’s mental life so that some of its aspects are prioritized relative to others” (p. 72). On Mole’s (2011) view, attention is “cognitive unison:” it occurs (roughly) when cognitive processes that could serve some action are not busy serving a different action—in this case the action is performed attentively. All these views are abstract – the philosophers in question abstract away from the specific neuroscientific details to create a broad personal-level picture of attention.

Because these conceptions of attention are abstract, they don't undermine our trilemma. For, our trilemma concerns accounts of attention that are “one level down” from these conceptions. For example, on Wu’s view, to attend to mental paint would be to select it for action, where the action is (say) reporting on it. This approach does not tell us *how* mental paint is selected: it could be done via (FO or HO) sensory attention or by (HO) cognitive attention. On Watzl’s view, to attend to mental paint would be to prioritize it relative to other mental items. But how? Again, it could be via (FO or HO) sensory attention or (HO) cognitive attention. On Mole’s view, to attend to mental paint would be to perform an action related to mental paint, e.g., reporting it, while the cognitive processes that could serve it are not busy serving a different action. But which processes allow us to report on mental paint? Do they involve what we have called (FO or HO) sensory attention or (HO) cognitive attention? This question is left open by Mole’s view.

Put differently, to determine *whether* we can attend to mental paint in Wu, Watzl, or Mole’s senses, we need to look one level down, namely at (FO and HO) sensory and (HO) cognitive attention, which seems to cover the logical space of possible accounts (on that level), and which leads to the problems we surveyed in sections 2 and 3. To avoid these problems, one needs to suggest a third sense of attention *on the same level* as sensory and cognitive attention. In other words, one needs to show that, contrary to appearance, the options we have surveyed do not exhaust the logical space of possible views on that level.

Note that similar considerations apply to the notion of recognitional phenomenal concepts directed at mental paint (Loar 2003). Such concepts (if they exist) are constituted by an ability to directly recognize (identify) mental paint in introspection, which is based on attending to mental paint. This characterization of recognitional phenomenal concepts appeals to attention in an abstract way: it is not specified whether the attention in question is FO sensory, HO sensory, or HO cognitive. Thus, the notion of a recognitional phenomenal concept does not offer a third conception of attention that could undermine our trilemma.

1. **Conclusion**

We have argued that the neuroscientific accounts of sensory attention are incompatible with first-order sensory attention to mental paint (Section 2). We further argued that cognitive attention to mental paint (in a higher-order manner) is too weak to enable direct identification of mental paint (in the absence of first-order sensory attention to mental paint), and hence it is irrelevant to transparency (Section 3.1). We then showed that the notion of higher-order sensory attention to mental paint has an apparently incoherent implication (Section 3.2). These conceptions of attention, namely first-order sensory, higher-order sensory, and higher-order cognitive, appear to exhaust the space of possible accounts of attention to mental paint. This amounts to an argument by elimination against the Block-Kind-Loar commitment to attentional shifts between the representational content of experience and mental paint (in a sense of “attention” that is relevant to transparency). The impossibility of such attentional shifts, combined with the assumption that our attention is usually directed at the representational content of experience (an assumption that Block, Kind and Loar accept, but which can be challenged), implies the transparency thesis (Section 1). Extant philosophical discussions of attention do not provide a third conception of attention, over and above the sensory and cognitive varieties (Section 4).

**Bibliography**

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of vision*, *9*(10), 7-7.

Bays, P. M. (2015). Spikes not slots: noise in neural populations limits working memory. *Trends in cognitive sciences*, *19*(8), 431-438.

Block, N. (2003). Mental Paint. In M. Hahn & B. Ramberg (Eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge, Mass.: MIT Press.

Bourget, D., & Mendelovici, A. (2014). Tracking representationalism. In Andrew Bailey (ed.) *Philosophy of Mind: The Key Thinkers*. Continuum.

Burge, T. (2003). Qualia and Intentional Content: Reply to Block. In M. Hahn & B. Ramberg (Eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge, Mass.: MIT Press.

Byrne, A. (2001). Intentionalism defended. *The Philosophical Review*, *110*(2), 199-240.

Byrne, A., & Tye, M. (2006). Qualia ain’t in the head. *Noûs*, *40*(2), 241-255.

Carruthers, P. (2000). *Phenomenal consciousness: A naturalistic theory*. Cambridge University Press.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*(5509), 1803-1806.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, *79*(1-2), 1-37.

Dretske, F. (1995). Naturalizing the Mind. Cambridge, MA: MIT Press.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends in cognitive sciences*, *16*(2), 129-135.

Harman, G. (1990). The intrinsic quality of experience. *Philosophical perspectives*, *4*, 31-52.

Intriligator, J., and P. Cavanagh. (2001) ‘‘The Spatial Resolution of Visual Attention.’’ *Cognitive Psychology 43*(3): 171–216.

Kind, A. (2003). What’s so transparent about transparency?. *Philosophical Studies*, *115*(3), 225-244.

Kind, A. (2008). How to Believe in Qualia. In E. Wright (Ed.), *The Case for Qualia*. Cambridge, Mass.: MIT Press.

Kind, A. (2010). Transparency and representationalist theories of consciousness. *Philosophy Compass*, *5*(10), 902-913.

Lepsien, J., & Nobre, A. C. (2006). Attentional modulation of object representations in working memory. *Cerebral cortex*, *17*(9), 2072-2083.

Loar, B. (2003). Transparent experience and the availability of qualia. In Q. Smith & A. Jokic (Eds.), *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.

Mendelovici, A. (2014). Pure intentionalism about moods and emotions. In U. Kriegel (Ed.), *Current*

*controversies in philosophy of mind*. New York: Routledge.

Mole, C. (2011). *Attention is cognitive unison: An essay in philosophical psychology*. Oxford University Press.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782-784.

McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, *19*(1), 431-441.

Noudoost, B., Chang, M. H., Steinmetz, N. A., & Moore, T. (2010). Top-down control of visual attention. *Current opinion in neurobiology*, 20(2), 183-190.

Papineau, D. (2014). Sensory Experience and Representational Properties. *Proceedings of the Aristotelian Society*, 114(1), 1-33

Pelli, D. (2008). Crowding: A cortical constraint on object recognition.’’ *Current opinion in neurobiology 18*(4), 445–51.

Prinz, J. J. (2012). *The Conscious Brain*. New York: Oxford University Press.

Reddy, L., Moradi, F., & Koch, C. (2007). Top–down biases win against focal attention in the fusiform face area. *Neuroimage*, *38*(4), 730-739.

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, 19(5), 1736-1753.

Rissman, J., Gazzaley, A., & D’esposito, M. (2009). The effect of non-visual working memory load on top-down modulation of visual processing. *Neuropsychologia*, 47(7), 1637-1646.

Sauret, W., & Lycan, W. G. (2014). Attention and internal monitoring: A Farewell to HOP. *Analysis, 74*(3), 363-370.

Shepherdson, P., Oberauer, K., & Souza, A. S. (2018). Working memory load and the retro-cue effect: A diffusion model account. *Journal of Experimental Psychology: Human Perception and Performance, 44*(2), 286-310.

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, *78*(7), 1839-1860.

Tye, M. (1995). *Ten problems of consciousness: A representational theory of the phenomenal mind*. Cambridge, MA: MIT Press.

Tye, M. (2002). Representationalism and the Transparency of Experience. *Noûs*, *36*(1), 137-151.

Tye, M. (2014). Transparency, qualia realism and representationalism. *Philosophical Studies*, *170*(1), 39-57.

Van Boxtel, J. J., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: on sufficiency and necessity. *Frontiers in psychology* *1*, 217.

Watzl, S. (2017). *Structuring mind: The nature of attention and how it shapes consciousness*. Oxford University Press.

Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *Journal of Neuroscience*, *28*(36), 8934-8944.

Wu, W. (2014) *Attention*, Routledge.

1. \* We would like to thank the audience at the Philosophy Department’s colloquium, University of Haifa, and at the colloquium of the Program for Consciousness Research, Tel Aviv University. Special thanks also to Michael Antony, Arnon Cahen, Baruch Eitam, Brie Gertler, Arnon Keren, Dominique Lamy, Ran Lanzet, Ariel Meirav, Jorge Morales, Liad Mudrik, Gil Sagi, Yaron Senderowicz, Marius Usher, and Yaffa Yeshurun for their helpful comments. This research was funded by the Humanities Fund. [↑](#footnote-ref-2)
2. It is customary to use asterisks to denote phenomenal character types, as in “white\*” (see Papineau 2014). [↑](#footnote-ref-3)
3. An anonymous reviewer suggested a consideration against this assumption along the following lines. The assumption is supposed to apply to hallucinations. The representational content of a hallucination of a yellow lemon does not include the yellowness of objects in the subject’s vicinity (i.e., instances of yellowness), since nothing there need be yellow. It therefore seems that the representational content should include the abstract property of yellowness, even though it is uninstantiated. But then the assumption that we typically attend to the representational content of experience appears to imply that, when hallucinating a yellow lemon, we attend to the abstract property of yellowness even though it is uninstiantiated. This latter idea seems very peculiar. [↑](#footnote-ref-4)
4. Couldn’t Block, Kind or Loar have a weaker sense of "attention" in mind? Specifically, perhaps they mean that attention to mental paint allows us to ask of it, “what is *that?*”, yet it does not allow us to directly identify it as blue\*, even in simple cases such as looking at the ocean at daytime. It is difficult to make sense of such an option. If we possess the concept of blueness\* yet we can’t tell that the mental paint we currently enjoy is blue\* even in the most clear-cut cases, then we don’t seem to notice it at all! [↑](#footnote-ref-5)