Is Representationalism Committed to Colour Physicalism?

DANIEL MARIO WEGER
Goethe-University Frankfurt am Main, Germany

The circularity problem states that the representationalist about phenomenal consciousness gives a circular explanation if she adopts the classic view about secondary qualities, such as colours, that characterises them as dispositions to produce experiences with a specific phenomenal character. Since colour primitivism faces severe difficulties, it seems that colour physicalism is the only viable option for the representationalist. I will argue that the representationalist is not committed to colour physicalism because she can adopt an anti-realist theory of colour. My diagnosis is that the alleged commitment to colour physicalism rests upon the acceptance of colour realism which is due to the approval of externalist versions of representationalism, such as tracking representationalism. I will argue that the representationalist can deal with the circularity problem by adopting figurative projectivism, which holds that colours are contingently non-instantiated properties that only figure in the representational contents of colour experiences.

Keywords: Representationalism about phenomenal consciousness; secondary qualities; circularity problem; colour physicalism; colour projectivism.

1. *Introduction*

Representationalism about phenomenal consciousness holds that the phenomenal character of an experience can be explained in terms of its representational content. Well-known representationalists such as Dretske (1995), Lycan (1996), and Tye (1995, 2000) prefer *strong representationalism,* which has it that phenomenal character is just one and

the same as representational content of a specific sort.[1] The significant advantage of the identity claim is that it comes with an account of what phenomenal consciousness is by its nature, whereas *weak representationalism*, construed as the thesis that phenomenal character supervenes on representational content, fails to provide such an account. My focus in this paper will be on strong representationalism because the problem I will be dealing with only afflicts strong versions of representationalism—but more on this later on.

The starting point of the question I will deal with in this paper is that representationalism[2] faces a significant problem with secondary qualities such as colours.[3] Since representationalism explains the phenomenal character of an experience in terms of its representational content, it is not compatible with the classic dispositionalist view about secondary qualities that construes them as dispositions to produce experiences with a specific phenomenal character in normal observers under standard conditions. This is what I will call *the circularity problem* for representationalism. Hence, it is generally accepted that the representationalist must adhere to a theory about colours that does not characterise colours in terms of the phenomenal character of the experiences they are apt to produce in observers. Since colour primitivism is at odds with what empirical science tells us about colour vision and the surfaces of perceived objects, colour physicalism seems to be the only viable option. Thus, representationalism is commonly held to be committed to colour physicalism, and representationalists have been making considerable efforts to vindicate colour physicalism. Yet, colour physicalism itself faces severe objections and is therefore not undisputed. Being committed to a colour theory widely believed to be false, the representationalist finds herself in a very unpleasant situation.

In this paper, I will examine whether representationalism is committed to colour physicalism in the first place. I will give a negative answer to this question and argue that the representationalist can adopt figurative projectivism instead because representationalism, in general, is compatible with an anti-realist theory of colour. Moreover, I will show that only externalist versions of representationalism, such as tracking representationalism, favoured by Dretske, Lycan, and Tye, need to stick to colour realism, i.e., a view that colours are instantiated in material objects.

---

[1] The qualification 'of a specific sort' is needed because phenomenal character cannot be held to be identical to representational content *tout court*. Obviously, there are mental states with representational content that lack phenomenal character, e.g., standing states like beliefs and wishes or non-conscious occurrent states like those in subliminal perception or early sensory information processing.

[2] From here, "representationalism" is used to refer only to strong representationalism if not stated otherwise.

[3] Since colours are treated as paradigmatic for secondary qualities, I will restrict my argument to the case of colours. Nevertheless, the points made in this paper carry over to other secondary qualities such as smells, sounds, tastes etc.

I will proceed as follows: In section 2, I will elaborate on the circularity problem to clarify why representationalism is incompatible with a specific construal of secondary qualities like colours. Section 3 will state the problems bestowing colour primitivism and explain why representationalism seems thus committed to colour physicalism. In section 4, I will present the objections against colour physicalism and depict an argument against representationalism that emerges from the preceding considerations. In section 5, I will examine the efforts to save colour physicalism undertaken by representationalists and argue that they fail. In section 6, I will show how the representationalist might resist the commitment to colour physicalism by adopting anti-realism about colours. Moreover, I will explain why the commitment to colour physicalism rests upon the approval of an externalist version of representationalism, such as tracking representationalism. In section 7, I will examine anti-realist theories of colour and argue that figurative projectivism is compatible with representationalism and offers a promising alternative to deal with the circularity problem. Section 8 will clarify what makes figurative projectivism attractive and how some prima facie problems might be attenuated. Finally, I will give a short outlook on the ensuing consequences for the prospects of representationalism in section 9.

2. *The circularity problem for representationalism*

At the heart of what I will deal with in this paper is what I will call the *circularity problem* for representationalism. Michael Tye (1995: 144) depicts the setting as follows:

> On the face of it, colors and other "secondary qualities" (smells, tastes, and sounds, for example) pose a special difficulty for the theory I have been developing. If these qualities are subjective, or defined in part by their phenomenal character, then what it is like to undergo the experiences of such qualities cannot itself be understood in terms of the experiences' representing them. That would create an immediate vicious circle.

This statement suggests that secondary qualities, such as colours, pose a particular threat to representationalism. Moreover, the potential problem facing representationalism is described as a case of circular reasoning. Thus, we need to look at two issues: What is it about secondary qualities that makes them a problem for representationalism? And what is the vicious circle that threatens representationalism?

What does it mean that secondary qualities are "subjective"? To start with, primary qualities like shape, size, and motion are commonly held to be properties that are intrinsic to the objects that possess them and, therefore, observer-independent. This means that these are qualities that objects have irrespective of whether they are possibly perceived or not. Hence, primary qualities can be characterised without appealing to how they might affect observers. In contrast, secondary qualities are usually construed as qualities defined in terms of subjective responses,

i.e., how they might affect perceiving subjects. Therefore, what characterises a secondary quality as the quality it is, are the responses this quality is apt to produce in perceivers. To the extent that the quality's nature is thus dependent on what responses it possibly causes in a perceiving subject, it is subjective in a sense.

With this in mind, we can also make sense of the clause that secondary qualities are "defined in part by their phenomenal character" if we assume that the subjective responses they are apt to produce in perceivers are essentially characterised by their phenomenal character. Accordingly, we can conceive of secondary qualities as defined in terms of the phenomenal character of the experiences they are apt to produce in the subjects that perceive them. When applied to colours, we receive the view that colours are characterised by their being disposed to produce visual experiences with a specific phenomenal character. For example, something is red just in case it is apt to produce sensations with a reddish phenomenal character. However, something needs to be added to the present account since the fact that objects might produce different responses in different observers and under different conditions is not compatible with the common-sense intuition that each object possesses only one real colour. This problem is usually fixed by adding the two qualifiers of *normal observers* and *standard conditions* to determine the colour property of a specific object. Thereby, we receive the following characterisation of the property of being red:

> X is red $=_{\text{def}}$ X is disposed to produce experiences with reddish phenomenal character in normal observers under standard conditions.

Such a view amounts to giving a dispositionalist account of colours because it defines colours as dispositions to produce visual experiences with a specific phenomenal character in a particular class of observers under certain conditions.

Now, we can see what problem arises for representationalism. As already indicated by Tye, representationalism faces the threat of running into a vicious circle. This problem is clearly brought out in the following passage of William Lycan (2019):

> [O]ne could not (without circularity) explicate phenomenal greenness in terms of represented real-world public colour and then turn around and construe the latter real physical greenness as a mere disposition to produce sensations of phenomenal greenness, or in any other way that presupposed phenomenal greenness.

On the one hand, representationalism has it that, say, an experience with greenish phenomenal character can be explained in terms of its representing that something is green. On the other hand, the dispositionalist theory of colour states that something is green just in case it is disposed to produce experiences with greenish phenomenal character in normal observers under standard conditions. Combining these two claims obviously gives a circular account because phenomenal character is explained in terms of representational content, and the proper-

ties that figure in the representational content of a relevant experience are explained in terms of the phenomenal character of the experiences they are apt to produce. This circularity is fatal, for it undermines the representationalist's aspirations to give an informative account of phenomenal consciousness.

Note that the circularity problem only concerns strong representationalism, which claims that phenomenal character is identical with representational content of a specific sort. In contrast, weak representationalism might adopt the view about secondary qualities presented above because it solely holds that phenomenal character supervenes on representational content. This implies that there can be no change in phenomenal character without a corresponding change in representational content. However, this does not require defining phenomenal character in terms of representational content—as strong representationalism does—and, therefore, no circularity is threatening. Nevertheless, it does not seem to be the right move to advert to weak representationalism in the face of the circularity problem since the notion of supervenience is not an explanatory one but serves as a starting point for further investigation into the relation between phenomenal character and representational content rather than being a substantial theory about phenomenal character.

So, suppose the representationalist wants to avoid the circularity problem. In that case, she must give an account of colours that does not define them in terms of the phenomenal character of the experiences they are apt to produce in perceivers. In short, a dispositionalist theory of colour, as presented here, is not an option for her.[4] To put it straight, the circularity problem brings about the following constraint, (C), for a representationalist theory about phenomenal consciousness:

(C) If representationalism is true, then colours must not be defined in terms of the phenomenal character of the experiences they are apt to produce in perceivers.

Now, it is time to look at which theories of colour satisfy the requirement stated in (C). This will be the topic of the next section.

3. *Why representationalism seems committed to colour physicalism?*

Up to this point, representationalism has not yet encountered any substantial difficulty. To be precise, the circularity problem only constrains the representationalist's choice regarding the metaphysics of colours. So, let us now look at the options she has. The two most influ-

---

[4] The very idea of a dispositionalist theory of colour does not entail a characterisation of colours in terms of phenomenal character. It is possible, for example, to define colours as dispositions to appear or look a certain way. However, such versions of dispositionalism face the problem of delivering a circular account (see Boghossian and Velleman (1997) and McGinn (1996); for a response see Byrne and Hilbert (2011)).

ential accounts of colour satisfying (C) are colour primitivism[5] and colour physicalism.[6] Colour primitivism holds that colours are simple and intrinsic properties sui generis, whereas colour physicalism claims that colours can be identified with physical properties of material objects. While primitivism conceives colours as non-analysable, non-relational, and non-reducible properties, physicalism assumes that colours can be reduced to physical properties of some kind.

Let us first consider colour primitivism. It enjoys a lot of prima facie plausibility because it matches our common-sense intuitions about colours. According to colour primitivism, colours just are what they phenomenologically appear to be: qualities that populate the surfaces of the things we visually perceive and are involved in bringing about colour experiences. Although colour primitivists hold that colours correlate with the physical properties of surfaces in some way, they reject that the first can be reduced to the latter. The primary motivation for this view is that the essence of colours is fully revealed in visual experience (Byrne and Hilbert 2007; Johnston 1992).

While this view is intuitively plausible, it faces several serious objections. Here, I will focus on the most pressing ones: First, colour primitivism flies in the face of what science tells us: Our best scientific theories, including those concerned with colour vision, suggest that material objects are not coloured since we can explain how colour experience comes about without having to allude to primitive colour properties. It is sufficient to advert to the physical properties of light, the perceived objects, the perceiver's visual system, and the lighting conditions to comprehensively explain how colour perception works (Gow 2014: 809; Maund 2018; Rubenstein 2018). Thus, properties like those postulated by the primitivist are explanatorily idle. Suppose the primitivist nevertheless holds that material objects instantiate primitive colour properties. In that case, she faces the following dilemma: Either she claims—contra what empirical sciences suggest—that primitive colour properties are involved in the production of colour experiences, or she adopts the view that colours are causally inert epiphenomena. While the first option comes with the cost of embracing causal overdetermination, the second one contradicts the common-sense assumption that the colours of objects are involved in the production of colour experiences (Gow 2014: 809; Hardin 1988: 61; Byrne and Hilbert 2007: 82–85). Therefore, neither option is plausible.

Another forceful objection against colour primitivism has it that the same external world object can look different concerning its colour on different occasions and to different perceivers. This results in the following dilemma: Either we accept that it has more than one colour,

---

[5] This view is held by Campbell (1997), Hacker (1987) and McGinn (1996), for example.

[6] Well-known defenders of this view are Armstrong (1997) Byrne and Hilbert (1997, 2003), Smart (1997) and Tye (1995, 2000).

or we need some non-arbitrary way to decide which one is "the right" colour of the object. Taking the first option flies in the face of our everyday assumption that objects only have one colour and that they do not change their colour every once in a while. The problem with the second option is that there is no non-arbitrary way to determine the conditions under which the object's actual colour is revealed (Hardin 1988: 80).

Altogether, these objections suggest that colour primitivism is in bad shape and that the representationalist should not adopt it. And now, we can see how the circularity problem for representationalism, together with the shortcomings of colour primitivism, ultimately leads to the claim that representationalism is committed to colour physicalism: The need to satisfy the constraint stated in (C), which is a result of the circularity problem, combined with the fact that colour primitivism fails leads to the claim that colour physicalism is the only viable option for the representationalist. However, the representationalist might remain unshaken. If she is willing to accept colour physicalism, she still does not face any more profound problems. Only if there were compelling arguments against colour physicalism would the representationalist find herself in an unpleasant situation. But unfortunately, this is precisely the case, as I will show in the next section.

4. *The problems with colour physicalism*

Colour physicalism claims that colours can be reduced to physical properties. One way to accommodate this claim is to hold that colours are identical to spectral reflectances of surfaces. The spectral reflectances of a surface is its disposition to reflect and absorb a certain amount of the incident light at every wavelength of the visible spectrum (Byrne and Hilbert 1997, 2003; Tye 1995, 2000). Yet, this view is susceptible to several objections. First, it is phenomenologically inadequate because colours do not look like surface spectral reflectances in visual perception.[7] If I visually experience a red object, it does not look to me as if the object instantiated such-and-such a spectral reflectance. Instead, it seems that the perceived object instantiates a specific qualitative property at its surface that is also had by ripe tomatoes and fire engines. So, the nature of colours as given in experience is radically different from what colour physicalism tells us (Averill and Hazlett 2011; Campbell 1997; Mendelovici 2018).

Second, colour physicalism cannot account for the alleged truth of claims about similarity relations between colours and their structural features, such as "Orange is more similar to red than to green" or "Red is a unique hue, whereas orange is a binary one" because there is nothing about surface reflectances that renders these statements true (Hardin 1988; Pautz 2006).

---

[7] A similar objection has been raised against colour dispositionalism in its realist version by McGinn (1996).

Third, the problem of metamers brings major trouble to colour physicalism. The starting point for this objection is the empirical fact that objects with different surface spectral reflectances can look the same for a specific observer under certain lighting conditions. Therefore, colours cannot be identified with spectral reflectances (Hardin 1988). A usual response on behalf of the colour physicalist is to identify colours with sets or disjunctions of spectral reflectances properties (Byrne and Hilbert 1997, 2003). Yet, this proposal comes to nothing because the only thing that keeps together the elements of a specific set is their aptness to look the same or produce visual experiences with the same phenomenal character (Gow 2014: 806).[8] This, however, means that colour physicalism would give up its aspiration to define colours without recourse to possible subjective responses.

Overall, the considerations just presented give us strong enough reason to accept the claim that colour physicalism is quite implausible. Taken together with the assumption that representationalism is committed to colour physicalism, this yields a disastrous conclusion for the representationalist. Now, there are two reactions on behalf of the representationalist. On the one hand, she might refuse the claim that colour physicalism is false and defend it against the abovementioned objections. On the other hand, she might argue that representationalism is not committed to colour physicalism in the first place by either rebutting the circularity problem and its consequences or by showing that representationalism might be combined with another theory of colour. So far, representationalists have usually taken the first route and tried to vindicate colour physicalism, for example, Dretske (1995) and Tye (1995, 2000). In the next section, I will examine their efforts to save colour physicalism and show that they fail.

5. *Defending colour physicalism*

In *Naturalizing the Mind*, Dretske (1995: 88–93) defends the view that colours are objective properties. Though he does not fully embrace the idea that colours are identical to spectral reflectances of surfaces, he acknowledges that the latter might play a significant role in characterising what kind of objective properties colours are. Nevertheless, his position should be considered a version of colour physicalism because he holds that the objective properties with which colours are identical can be characterised in broadly physical terms. In this context, he addresses the problem of metamers and appeals to the fact that the visual system of humans was naturally selected under some specific circumstances because it enabled humans to identify the colours of objects under these very circumstances and, thus, helped them to flourish and survive. According to Dretske, metamerism results when the

---

[8] This point is even acknowledged by Dretske (1995: 89–90) who otherwise defends colour physicalism. See also the section 5 of this paper.

visual system of humans operates under conditions for which it was not originally selected. So, for him, metamerism is a case of perceiving colours under conditions that deviate from selection conditions and, therefore, misrepresentation is just what we should expect. But this, Dretske continues, in no way implies that colour experiences do not represent objective properties.

This line of response is not appropriate because the core of the problem of metamers is that two or more surfaces with different spectral reflectances look the same under some specific lighting conditions. In contrast, Dretske considers cases of objects with different objective properties—spectral reflectances, for example—looking similar under different conditions. This, however, misses the point stated by the problem of metamers. Moreover, there is no reason to think that metamerism could not occur under what Dretske calls selection conditions. To defend his position, Dretske would need to show that metamerism can only happen when selection conditions do not obtain. His claim that metamerism always involves illusion is extremely hard to swallow (Shrock 2017: 141–142). Finally, the assumption that there must be some objective property represented in colour experience seems to be nothing more than wishful thinking since it is doubtful that colour vision even has the biological function of detecting physical properties of objects. As Ross (2000: 123–124) remarks, it is much more adequate to account for colour vision's biological function in ecological terms.

Let us turn to another well-known representationalist who tackles the circularity problem by trying to vindicate colour physicalism. In *Ten Problems of Consciousness*, Tye (1995: 146–148) defends a view that identifies colours with ordered triples of spectral reflectances, each of which covers a specific band of wavelengths corresponding to the wavelength ranges to which the three different types of cones in the human eye are each sensitive to. Given this proposal, Tye wants to account for the problem of metamers by holding that metamers have similar triples of spectral reflectances. This allows for objects with different spectral reflectances to have the same colour as long as their spectral reflectance properties are similar enough concerning the relevant bands of wavelengths. Moreover, Tye thinks that his proposal can also account for the similarity relations between colours because the relations between the triples of spectral reflectance mirror the structure of the hue circle.

While this is an interesting proposal that could be empirically assessed, it is purely speculative that metamers have similar triples of spectral reflectance. Anyway, I think that even empirical evidence will not do it. This response is unsatisfying because it provides an anthropocentric account as far as the selection of the relevant wavelength bands solely rests upon a contingent fact about the visual systems of humans. So, whether the proposal can eventually be empirically corroborated or not, it fails as an objective theory of colour because it defines colours

relative to the structure of the visual system of a specific species. Yet, it is conceivable that the visual system of another species has a different number of cones or that its cones respond to different wavelength bands, for example. Nevertheless, these animals might have experiences with bluish, reddish, and yellowish phenomenal character. How should we then decide what triples or n-tuples of spectral reflectance colours are? On the one hand, it seems arbitrary to define colours relative to the visual system of a specific species. And on the other hand, taking all possible combinations of n-tuples of spectral reflectance into account delivers a disjunctive characterisation of colours that is eclectic rather than objective.

In *Consciousness, Colour, and Content,* Tye offers another proposal that incorporates insights from the opponent-process theory to deal with the problem of metamerism (2000: 159–161) and the fact that colour physicalism cannot account for the alleged truth of statements about similarity relations between colour (2000: 162–165). This proposal is prima facie plausible because it provides a characterisation of colours that is insofar objective and observer-independent as activity patterns in the opponent process channels are objectively discernible and quantifiable. However, a severe problem with this proposal becomes apparent upon closer inspection. Why would we accept the claim that some colour is identical to a set of conditions or properties that cause a specific activity pattern in opponent-process channels? I suspect that the inclination to accept such a claim rests on the assumption that some particular activity pattern in opponent-process channels produces experiences with a specific phenomenal character. But this results in a dilemma for Tye: Either he accepts this assumption and gives an account that indirectly characterises colours in terms of the phenomenal character of the experiences they are apt to bring about. Or he rejects this assumption and claims that there is no relevant connection between activity in opponent-process channels and the phenomenal character of colour experience. But in the latter case, it is utterly mysterious why activity in opponent-process channels should then be an appropriate candidate for the characterisation of colours at all or a better one than any other neuronal activity that also bears no relevant connection to the phenomenal character of colour experience.

Finally, Tye (2000: 150) claims that the phenomenon of colour constancy shows us that the colour of an object is not to be identified with the wavelength of light it reflects in specific lighting conditions. Since the problem of metamers has it that objects with differing spectral reflectances can look the same under certain lighting conditions, the problem of metamers does not challenge colour physicalism that identifies colours with spectral reflectances, or so he thinks. However, this line of reasoning is not convincing because the problem of metamers has it that objects with different spectral reflectances can look the same under some specific illumination conditions. In contrast, colour constancy considers the same object and, thus, the same spectral reflec-

tance under different illumination conditions. Hence, Tyes remarks, true as they may be, do nothing to alleviate the problem of metamerism, which remains as pressing for colour physicalism as ever.

All things considered, the efforts made by Dretske and Tye to defend colour physicalism do not look very promising, and the situation is even worse for the representationalist.[9,10] If representationalism is committed to colour physicalism and there are no persuasive responses to the objections against colour physicalism, it seems that representationalism is fighting a lost cause. Thus, arguing that representationalism is not committed to colour physicalism seems to be the only way out. I will opt for this alternative route and show how this can be accomplished in the following sections.

6. *Physicalism or primitivism about colours? A false choice*

As I have shown in the last section, it seems a desperate move to stick to colour physicalism and defend it against the objections put forth against it. Hence, the representationalist must either rebut the circularity problem and its consequences or show that representationalism might be combined with another theory of colour. As far as the circularity argument is concerned, I do not see any way out for the representationalist but to accept it and the constraint that comes with it.[11] Any attempt to reject the circularity argument is, I think, doomed to failure. Thus, I will take (C) for granted, as do all the representationalists involved in the debate. So, the only remaining option is to defend another theory of colour. Of course, it would, in principle, also be possible to defend colour primitivism against the objections discussed in section 3. But since colour primitivism's problems weigh heavy and colour primitivism runs contrary to the physicalist convictions usually held by representationalists, I will not consider this option.

But what other theory of colour could the representationalist turn to? Upon closer examination, it becomes clear that we have only consid-

[9] For further criticism of the strategies deployed in Tye (2000), see Hardin (2003). For further, albeit quite similar, proposals to defend colour physicalism, see Bradley and Tye (2001).

[10] Another objection against colour physicalism I have not discussed here concerns its dispositional aspect and the resulting problem that colour properties as understood by the colour physicalist have no causal powers. For a response, see Tye (2000: 161–162), and for a critical assessment thereof, see Wright (2003: 520–521).

[11] It is possible to refuse the constraint imposed by the circularity problem by holding that the properties that are represented in colour experience are different from colours. However, such an account would need to say what these properties are that we represent in colour experience and that are different from colours, how they relate to colours and why we mistakenly take them to be colours. Moreover, these properties would need to be characterized without reference on the phenomenal character of colour experience. Otherwise, the problem of giving a circular account would arise again.

ered realist theories of colour so far, i.e., theories assuming that colours are properties that are instantiated in material objects. However, there is no need to accept colour realism in the first place. Neither is representationalism as such committed to colour realism nor is it necessary to accept colour realism to give an adequate response to the circularity argument. Remind that the constraint stated in (C) is all that follows from the circularity problem. And for a theory about colours to satisfy this requirement, it is enough if it does not define colours in terms of subjective responses. Whether colours are properties instantiated in material objects is a separate question. It is open to the representationalist whether to give a positive or negative answer to it when faced with the circularity problem. Thus, the representationalist might willingly adopt a theory of colour that is anti-realist in spirit in the absence of independent reasons for assuming that colours are properties of material objects. So, my diagnosis is that the claim that representationalism is committed to colour physicalism is based on the implicit—and, as I will argue shortly, unjustified—acceptance of colour realism. Therefore, assuming that the choice for the representationalist is between colour physicalism and colour primitivism is entirely misguided.

Before saying something about how to spell out the proposal of adopting anti-realism about colours, it is worth considering the motivation for representationalists such as Tye and Dretske to assume colour realism. It seems natural to them to accept colour realism because they favour tracking representationalism, an externalist version of representationalism. This strand combines the representationalist idea with the claim that mental states obtain their representational content in virtue of their tracking features in the subject's environment. Tracking is cashed out either as a matter of having the function to provide information about the subject's environment (Dretske 1995) or as a matter of being related to the subject's environment in an appropriate way, for example, standing to it in a specific causal relation such as causal covariance under optimal conditions (Tye 1995, 2000). Accordingly, the tracking theory requires that the represented properties be instantiated in material objects, at least in content endowing conditions such as the conditions of evolutionary selection or optimal conditions (Mendelovici 2013). Therefore, tracking representationalism is to be considered externalist in spirit.[12] Obviously, this leads tracking representationalists to claim that colours are properties instantiated in material objects and, consequently, to accept colour realism. Since tracking representationalists are usually drawn to naturalistic metaphysics, they adopt and defend colour physicalism, as shown in the last section.

However, since there is no need to accept the tracking theory in the first place, the representationalist is free to adopt an anti-realist theory

[12] According to Gow (2017), externalist representationalism in general is committed to colour physicalism. She arrives at this conclusion by similar considerations as the ones invoked in this paper.

about colours. Thus, my argument is that representationalism is not committed to colour physicalism, although this comes at the cost of rejecting externalist versions of representationalism such as tracking representationalism.

7. *Anti-realist theories of colours*

The next step for the representationalist is to find out which anti-realist colour theories are on offer and whether they are compatible with representationalism.[13] There are two major candidates: Eliminativism[14] and projectivism. First, eliminativism holds that colours do not exist. According to this view, there are no colours at all, not even uninstantiated ones. Second, projectivism claims that we project the experienced colours onto the objects we perceive due to our having colour experiences but that the perceived objects themselves are not coloured. However, projectivism comes in two versions, literal and figurative projectivism, and they differ significantly (Shoemaker 1990, 1997). Literal projectivism claims that colours are properties that are instantiated in visual fields or other mental entities similar to sense-data (Boghossian and Velleman 1997). In contrast, figurative projectivism has it that colours are properties that are not instantiated at all, neither in material objects nor in visual fields (Maund 2006; Pautz 2006; Wright 2003).

Now, which form of anti-realism is compatible with representationalism? Obviously, eliminativism is no viable option for the representationalist because it denies the very existence of colours—properties that feature in the representationalist's explanation of the phenomenal character of colour experiences. Moreover, literal projectivism is no good option for the representationalist either because it presupposes visual fields or other sense-data-like entities, which conflicts with the representationalist's aim of giving a physicalistically respectable explanation of phenomenal consciousness. Besides these metaphysical worries, literal projectivism is not compatible with representationalism because it holds that colours are modifications of our experiences or mental entities rather than properties that figure in the representational contents of colour experiences. Figurative projectivism, however, is compatible with representationalism because it both assumes the existence of colour properties—as opposed to eliminativism—and it does

---

[13] Wright (2003) has already argued that representationalism is compatible with the denial of colour realism. However, he does not start his discussion from the circularity argument and does not present anti-realism about colours as an adequate response to the circularity problem. Another major difference is that he claims that externalist versions of strong representationalism are compatible with colour projectivism, whereas I deny this.

[14] The term "eliminativism" is used ambiguously in the debate about the metaphysics of colours. Sometimes, it serves as a label for what I call anti-realist theories, in other cases it is only used to refer to the view that colours do not exist, full stop. I will only use the term "eliminativism" in the latter sense here.

not presuppose the existence of questionable mental entities like visual fields or sense-data—in contrast to literal projectivism.

Though, does figurative projectivism not finally collapse into colour eliminativism since it holds that colours are nowhere instantiated? Not at all. Figurative projectivism assumes that colour properties do, in fact, exist because it claims that they figure in the representational contents of colour experiences. However, it is only a contingent matter of fact that colours are not instantiated in our world. According to figurative projectivism, colours might nevertheless be instantiated in some possible world, e.g., an Edenic world where we would be directly acquainted with the objects around us and their intrinsic qualities (Chalmers 2006). But this is just to say that figurative projectivism is open to the possibility that some possible worlds are different from our actual world concerning the instantiation of colour properties. It does not bear on the central tenet of figurative projectivism that we only mistakenly project colours onto the surfaces of perceived objects when having colour experiences in our actual world.

But what about figurative projectivism as a theory of colour that satisfies the constraint stated in (C)? To begin with, figurative projectivism only tells us something about where colour properties are instantiated or, instead, that they are instantiated neither in material objects nor mental entities. However, it does not tell us anything about the metaphysical nature of colours. It holds that colours are not identical to any kind of properties of material objects since it opposes the very idea that material world objects instantiate colour properties. But what positive claim about the metaphysical nature of colours might be made by the figurative projectivist?

The fact that figurative projectivism holds that we mistakenly project colours onto the perceived objects and that colours are only contingently not instantiated in our world suggests that it is most naturally combined with a primitivist account of the metaphysical nature of colours claiming that colours are simple and intrinsic properties sui generis. Yet, this does not mean that figurative projectivism ultimately collapses into colour primitivism. The primitivist approach described above is realist in spirit, whereas figurative projectivism is anti-realist. They converge in what they say about the metaphysical nature of colours. Both views reject identifying colours with physical properties of external world objects or with dispositions to cause visual experiences with a specific phenomenal character but construe them as simple and intrinsic properties sui generis. Nevertheless, they diverge in what they say about the instantiation of colour properties. To be precise, both primitivism and figurative projectivism deny that colours are properties of our colour experiences themselves. But according to primitivism, colours are had by external world objects, while figurative projectivism holds that nothing instantiates colour properties in our actual world.

8. *Assessing figurative projectivism*

Before dealing with the alleged problems of figurative projectivism, let us first look at how it fares compared to colour physicalism and colour primitivism.[15] In line with what colour physicalism says, figurative projectivism has it that colour experiences are typically caused by the physical properties of the material objects we perceive. However, in contrast to colour physicalism, it denies that colours can be in some way identified with the physical properties of external world objects. Therefore, it disagrees with colour physicalism on whether material objects possess colour properties. This means that figurative projectivism incorporates the advantages of colour physicalism—its compatibility with empirical findings of colour vision—while avoiding its pitfalls—the problems due to identifying colours with the physical properties of material objects.

In accordance with primitivism, figurative projectivism holds that colours are simple, intrinsic, and non-reducible properties. Yet, these two views differ regarding the instantiation of colours. While primitivism purports that material objects have colour properties so construed, figurative projectivism claims that colour properties are nowhere instantiated neither in material objects nor in perceivers or mental entities like sense-data and visual fields. Thus, it is sometimes even held that figurative projectivism, as presented here, is just an anti-realist version of primitivism. So, while figurative projectivism adopts the part of primitivism that fits our common-sense notion of colours—its account of the metaphysical nature of colours –, it does not inherit the problems that come with the claim that colour properties as construed by the primitivist are instantiated in material objects –the unpalatable consequences of colour primitivism regarding what science tells us about colour vision and the surface properties of material objects.

Now, let us examine the difficulties that are supposed to come along with figurative projectivism. As set out in the last section, figurative projectivism states that our colour experiences mistakenly represent objects as coloured, even in the case of successfully perceiving an object. Accordingly, figurative projectivism is committed to the claim that all our colour experiences are non-veridical. However, this flies in the face of our common-sense intuitions about colour perception. It seems appropriate to accept a principle of charity that assumes that not all our colour experiences are blatantly false but that they are more or less correct most of the time (Shoemaker 1997).[16] Moreover, we presume that our judgments about the colours of objects are more or less accurate most of the time and that our discriminations of the colours of objects guide our successful behaviour towards our environment. But figurative projectivism must deny all this. Since figurative projectivism

---

[15] Some of the points made here are similar to those in Wright (2003: 522).

[16] In contrast, Boghossian and Velleman (1997) hold that it is wrong to appeal to a principle of charity in the case of colour experience.

holds that all of our colour experiences are non-veridical, it comes with the unpleasing consequence that our attributions of colour properties to material objects in our everyday use of colour concepts and terms are wrong, given that we thereby express the contents of our colour experiences. And in addition to that, it does not seem capable of giving us an adequate explanation of how colour perception may serve as a guide for successful behaviour towards our environment when colour perception gets things wrong all the time.

First, let us consider the point about colour perception's alleged uselessness due to its being non-veridical. The assumption upon which this objection is implicitly based is that colour perception can only serve its use in guiding successful behaviour towards our environment if it represents accurately. However, plausible as this claim might seem prima facie, it becomes apparent that it is entirely misguided upon closer consideration. To see this, let us assume that colour experience is misrepresenting all the time. Of course, this implies that colour experience is always non-veridical. Still, it does not preclude colour experience from misrepresenting *systematically*, i.e., that there is some pattern in the way we mistakenly misrepresent objects as having colours. Mendelovici (2013: 422) claims that systematic misrepresentation involves reliability because it "is getting things wrong in the same way all the time." Thus, while being non-veridical, our colour experiences might nevertheless be reliable if they misrepresent systematically. Furthermore, Mendelovici stresses that veridicality must be kept separate from reliability, and it is the latter that secures guidance of successful behaviour. Thus, colour experience might nonetheless serve as a guide to successful behaviour despite misrepresenting our environment if it does so systematically and is therefore reliable.[17]

But what about figurative projectivism and colour experience as systematic misrepresentation? As mentioned above, figurative projectivism holds that colour experience is, at least in the case of perception, caused by the physical properties of the perceived objects. Therefore, as per figurative projectivism, it is plausible to assume that the way colour experience misrepresents objects as having colour properties correlates with the physical properties causing the relevant colour experiences. And this is just what systematic misrepresentation amounts to: a specific type of representation is tokened in similar conditions on various occasions having a content that is never satisfied (Mendelovici 2013: 423). So, while figurative projectivism has it that colour experience is always non-veridical, it can account for the fact that colour perception is useful by claiming that colour experience is reliable because it misrepresents systematically.

---

[17] A similar point is made in Gow (2016, 2019), who emphasizes that success in not dependent on accuracy. More general, this way of reasoning is usually embraced by proponents of figurative projectivism, see Maund (2006), Pautz (2006) and Wright (2003), for example.

This brings us to the objection that figurative projectivism has the implausible consequence of our everyday colour discourse being flawed because we mistakenly attribute colour properties to material objects. As Boghossian and Velleman (1997: 99) put it, there is no problem in our reports about the colours of objects asserting falsehoods. While it may be true that our everyday colour talk is, strictly speaking, thoroughly false, it can nevertheless be useful in serving the purpose of communication or planning our actions. Boghossian and Velleman point out that even though we often assert falsehoods in everyday talk, e.g., when claiming that the sun rises, there is no need to revise our way of talking about objects and their properties in the light of new (scientific) evidence. Just as we can successfully communicate by asserting that the sun rises and plan our actions based on the belief that the sun rises—though this is false, strictly speaking—we can do so in the case of colours as well. Again, the reason for this is that our colour experience is systematically misrepresenting and, thus, the contents expressed by our assertions about the colours of objects are also systematically false. And since we all systematically misrepresent the objects in our environment with colours that they do not possess, we can successfully communicate with each other even though our colour attributions are false, strictly speaking.

Another worry concerning figurative projectivism issued by Tye (2000: 166) is that it is unclear what it would take for colour experience to be veridical, assuming that figurative projectivism is true. As mentioned before, figurative projectivism has it that it is only a contingent matter of fact that the material objects in our world are not coloured. However, there may be a world where material objects do, in fact, possess the colour properties they are represented as having. In such circumstances, our colour experiences were veridical. This idea can be cashed out, for example, by conceiving of colours as Edenic properties that are instantiated in an Edenic world (Chalmers 2006).

9. *The prospects of representationalism*

In this closing section, I want to briefly summarise what has been said so far and look at the prospects of representationalism. The circularity problem is that representationalism cannot adopt any theory about colours that characterises them as dispositions to produce experiences with a specific phenomenal character. Since colour primitivism faces severe difficulties, it is usually held that the representationalist can only satisfy the constraint imposed by the circularity problem by accepting colour physicalism. Therefore, representationalism seems committed to colour physicalism. However, colour physicalism is not undisputed and defending it from the objections raised against it does not turn out to be fruitful, as I have shown.

But is representationalism committed to colour physicalism at all? I have argued that the answer to this question is "no." This is because

representationalism is not committed to a realist theory of colour that holds that colours are instantiated in material objects. The motivation to adopt a realist theory of colour, so my diagnosis, is based upon defending an externalist version of representationalism, such as tracking representationalism. Yet, since representationalism, in general, can be defended without accepting an externalist theory about mental representation, there is no need to stick to colour realism in the first place. Therefore, the representationalist can deal with the circularity problem by adopting an anti-realist theory of colour. However, this comes at the cost of renouncing externalist versions of representationalism. As far as the choice among anti-realist theories of colour is concerned, the representationalist should opt for figurative projectivism instead of eliminativism and literal projectivism because only figurative projectivism satisfies the requirements of a representationalist theory of phenomenal consciousness. We can thus conclude that the circularity problem does not pose an essential threat to representationalism because representationalism, in general, is not committed to colour physicalism. Only externalist versions such as tracking representationalism are. It is now up to representationalists to develop an updated version of the theory that provides an account of mental representation that leaves externalist commitments behind and shows how an internalist version of representationalism, in combination with figurative projectivism, can be made to work out to give an adequate explanation of colour experiences.

References

Armstrong, D. M. 1997. "Smart and the secondary qualities." In A. Byrne and D. R. Hilbert (eds.). *Readings on Color. The Philosophy of Color*. Cambridge: MIT Press, 33–46.

Averill, E. W. and Hazlett, A. 2011. "Color objectivism and color projectivism." *Philosophical Psychology* 24 (6): 751–765.

Boghossian, P. A. and Velleman, J. D. 1997. "Colour as a Secondary Quality." In A. Byrne and D. R. Hilbert (eds.). *Readings on Color. The Philosophy of Color*. Cambridge: MIT Press, 81–103.

Bradley, P. and Tye, M. 2001. "Of Colors, Kestrels, Caterpillars, and Leaves." *The Journal of Philosophy* 98 (9): 469–487.

Byrne, A. and Hilbert, D. R. 1997. "Colors and Reflectances." In A. Byrne and D. R. Hilbert (eds.). *Readings on Color. The Philosophy of Color*. Cambridge: MIT Press, 263–288.

____ 2003. "Color realism and color science." *Behavioral and Brain Sciences* 26 (1): 3–64.

____ 2007. "Color primitivism." *Erkenntnis* 66 (1–2): 73–105.

____ 2011. "Are colors secondary qualities?" In L. Nolan (ed.). *Primary and Secondary Qualities: The Historical and Ongoing Debate*. Oxford: Oxford University Press, 339–361.

Campbell, J. 1997. "A Simple View of Color." In A. Byrne and D. R. Hilbert (eds.). *Readings on Color. The Philosophy of Color*. Cambridge: MIT Press, 177–190.

Chalmers, D. J. 2006. "Perception and the Fall from Eden." In T. S. Gendler and J. Hawthorne (eds.). *Perceptual experience*. Oxford: Clarendon Press, 49–125.

Dretske, F. I. 1995. *Naturalizing the Mind*. Cambridge: MIT Press.

Gow, L. 2014. "Colour." *Philosophy Compass* 9 (11): 803–813.

____ 2016. "The Limitations of Perceptual Transparency." *The Philosophical Quarterly* 66 (265): 723–744.

____ 2017. "Colour hallucination: A new problem for externalist representationalism." *Analysis* 77 (4): 695–704.

____ 2019. "Everything is clear: All perceptual experiences are transparent." *European Journal of Philosophy* 27 (2): 412–425.

Hacker, P. M. S. 1987. *Appearance and reality*: *A philosophical investigation into perception and perceptual qualities*. Oxford: Basil Blackwell.

Hardin, C. L. 1988. *Color for philosophers*: *Unweaving the rainbow*. Indianapolis: Hackett.

____ 2003. "A Spectral Reflectance Doth Not A Color Make." *Journal of Philosophy* 100 (4): 191–202.

Johnston, M. 1992. "How to speak of the colors." *Philosophical Studies* 68 (3): 221–263.

Lycan, W. G. 1996. *Consciousness and experience*. Cambridge: MIT Press.

____ 2019. "Representational Theories of Consciousness." In Zalta, E. N. (ed.) The Stanford Encyclopedia of Philosophy. Online available: https://plato.stanford.edu/entries/consciousness-representational/. Last accessed on: 25 Aug 2021.

Maund, B. 2006. "The Illusory Theory of Colours. An Anti-Realist Theory." *Dialectica* 60 (3): 245–268.

____ 2018. "Color." In Zalta, E. N. (ed.) *The Stanford Encyclopedia of Philosophy*. Online available: https://plato.stanford.edu/archives/sum2018/entries/color/. Last accessed on: 1 Nov 2018.

McGinn, C. 1996. "Another Look at Color." *The Journal of Philosophy* 93 (11): 537–553.

Mendelovici, A. 2013. "Reliable misrepresentation and tracking theories of mental representation." *Philosophical Studies* 165 (2): 421–443.

____ 2018. *The Phenomenal Basis of Intentionality*. New York: Oxford University Press.

Pautz, A. 2006. "Can the physicalist explain colour structure in terms of colour experience?" *Australasian Journal of Philosophy* 84 (4): 535–564.

Ross, P. W. 2000. "The Relativity Of Color." *Synthese* 123 (1): 105–129.

Rubenstein, E. M. 2018. "Color." In *Internet Encyclopedia of Philosophy*. Online available: https://www.iep.utm.edu/color/. Last accessed on: 29 Oct 2018.

Shoemaker, S. 1990. "Qualities and Qualia: What's in the Mind?" *Philosophy and Phenomenological Research* 50: 109–131.

____ 1997. "Phenomenal Character." In A. Byrne and D. R. Hilbert (eds.). *Readings on Color*. *The Philosophy of Color*. Cambridge: MIT Press, 227–245.

Shrock, C. A. 2017. *Thomas Reid and the Problem of Secondary Qualities*. Edinburgh: Edinburgh University Press.

Smart, J. J. C. 1997. "On some criticisms of a physicalist theory of colors." In A. Byrne and D. R. Hilbert (eds.). *Readings on Color. The Philosophy of Color*. Cambridge: MIT Press, 1–32.

Tye, M. 1995. *Ten Problems of Consciousness*: *A Representational Theory of the Phenomenal Mind*. Cambridge: MIT Press.

____ 2000. *Consciousness, color, and content*. Cambridge: MIT Press.

Wright, W. 2003. "Projectivist representationalism and color." *Philosophical Psychology* 16 (4): 515–533.