



Seeing colours unconsciously

Paweł Jakub Zięba¹

Received: 25 February 2021 / Accepted: 2 April 2022
© The Author(s) 2022

Abstract

According to unconscious perception hypothesis (UP), mental states of the same fundamental kind as ordinary conscious seeing can occur unconsciously. The proponents of UP often support it with empirical evidence for a more specific hypothesis, according to which colours can be seen unconsciously (UPC). However, UPC is a general claim that admits of many interpretations. The main aim of this paper is to determine which of them is the most plausible. To this end, I investigate how adopting various conceptions of colour and perceptual phenomenal character affects UPC's resilience to objections. This brings me to the conclusion that the most plausible reading of UPC is the one according to which the phenomenal character of colour perception (i) is constituted by colours qua primitive mind-independent qualities of the environment and (ii) is not essentially tied to consciousness. My conclusion not only identifies the most plausible interpretation of UPC, but also highlights and supports an unorthodox version of the relational theory of perception, which is a perfectly viable yet so far overlooked stance in the debate about unconscious perception.

Keywords Unconscious perception · Colour perception · Phenomenal overflow · Consciousness

1 Introduction

According to unconscious perception hypothesis,
(UP) Mental states of the same fundamental kind as ordinary conscious seeing can occur unconsciously.

✉ Paweł Jakub Zięba
pj.zieba@uj.edu.pl

¹ Jagiellonian University, Cracow, Poland

A significant amount of empirical evidence for UP suggests that a more specific hypothesis is true:

(UPC) Colours can be seen unconsciously¹.

UPC is a general claim that admits of many interpretations. The precise meaning of UPC depends on what conceptions of colour and perceptual phenomenal character are assumed. Some such conceptions simply rule out UPC, and among those that allow it some render it more plausible than others.

Since UPC can be specified in many ways, it is reasonable to ask which reading of UPC is the most plausible. The main aim of this paper is to answer this question. To this end, I consider how adopting various conceptions of colour and perceptual phenomenal character affects UPC's resilience to objections. This brings me to the conclusion that the most plausible reading of UPC combines (a) colour primitivism (i.e. the view that colours are *sui generis* primitive properties) with (b) the view that the phenomenal character of colour perception² (i) is constituted by colours qua mind-independent qualities of the environment and (ii) is not essentially tied to consciousness.

In Sect. 2, I set forth a selection of empirical results often mentioned in support of UPC. Section 3 considers two standard objections against UP (which apply to UPC as well) and finds them inconclusive. Sections 2 and 3 motivate pursuing the main research question of the paper by providing some initial replies to the sceptic about UP.

Section 4 deals with a third objection that targets UPC specifically. I argue that this objection loses force when UPC is understood through the lens of colour primitivism. This section has two purposes. The first one is to defend UPC. While inconclusiveness of the three objections does not prove UPC, it justifies investigating the ramifications of this hypothesis. The second purpose of Sect. 4 is to determine which conception of colour renders UPC the most plausible (i.e. the most resilient to objections). I conclude that the primitivist reading is the most plausible, and for this reason work *only* with this (i.e. primitivist) reading in the remainder of the paper.

Section 5 outlines four general conceptions of perceptual phenomenal character, each of which can be specified in various ways. This typology is supposed to cover the most popular views in the literature, but it is not intended to be exhaustive. Scrutinizing what can and cannot be said about UPC on each of these accounts leads me to the conclusion that conceptions countenancing unconscious phenomenal character (i.e. inequivalence theories) are more suitable to accommodate the primitivist reading of UPC than conceptions according to which phenomenal character is essentially conscious (i.e. equivalence theories). Consequently, by the end of Sect. 5, the choice of the most plausible reading of UPC is narrowed down to two options.

¹ UP does not entail UPC, i.e. showing that UPC is false would not suffice to undermine UP. UP is thus consistent with the claim that some things can be seen only consciously. On the flip side, the stronger evidence for UPC is gathered, the more plausible UP becomes.

² Phenomenal character is typically introduced as an aspect of 'perceptual experience'. The term 'perceptual experience' usually refers to conscious instances of perception and hallucination. Since this paper is exclusively concerned with perception understood as a successful sensory encounter with the world, by 'perceptual experience' I always mean conscious perception, never hallucination.

Section 6 provides a reason (not decisive, but substantial) to prefer the externalist version of the inequivalence conception over the internalist one. The reason is that the former avoids a problematic dilemma faced by the latter. As a result, I arrive at the conclusion that the most plausible reading of UPC is the one according to which the phenomenal character of colour perception (i) is constituted by colours qua primitive mind-independent qualities of the environment and (ii) is not essentially tied to consciousness. This conclusion not only identifies the most plausible interpretation of UPC, but also suggests that the proponents of UPC should consider adopting an unorthodox version of the relational theory of perception, which is a perfectly viable yet so far overlooked stance in the debate about unconscious perception.

Section 7 is a brief conclusion.

2 Empirical evidence

This section is a brief survey of empirical results suggesting that colours can be seen unconsciously. They are often mentioned as the strongest available evidence for UP (see e.g. Block 2016; Peters et al. 2017).

The first batch of evidence for UPC comes from studies using transcranial magnetic stimulation of primary visual cortex. In the (Boyer, Harrison, and Ro 2005) study, this method was used to evoke temporary blindsight in the subjects. Despite having their awareness of the target stimuli (*inter alia* colours) suppressed, they performed with above-chance accuracy in forced choice discrimination tasks. These results were replicated in (Railo et al. 2012), although the researchers do not rule out that the above-chance performance was enabled by residual awareness of the stimulus' colour.

Evidence for UPC was also found in metacontrast studies. Metacontrast is a kind of priming where the prime and the target stimulus appear at different locations. Typically, the target is a small dot, whereas the prime is an annulus surrounding the dot. The subjects are first presented with the annulus for a very short amount of time, then the annulus disappears, and the dot is presented. The exposition of the annulus is too short for the subjects to become consciously aware of it. And yet the subjects recognise the colour of the dot more quickly when the annulus and the dot are of the same colour (Ro et al. 2009).

In another metacontrast study (Norman et al. 2014), it was found that this effect occurs only if the prime and the target have the same reflectance, and even if they are presented to the subject in different illuminations. It was also observed that the effect does not occur when the prime and the target have different reflectances but their respective illuminations are such that the same light hits the retina in both cases. This indicates a dissociation between perceptual consciousness and colour constancy. The latter occurs when a change in illumination does not prevent the subject from perceiving a single colour. In the study at hand, the annulus was perceived to be of the same colour as the dot even though the annulus and the dot were illuminated differently. Since the reflectance of the annulus has influenced the subjects' reaction times despite not being consciously accessible to them, the researchers conclude that colour constancy does not require consciousness.

Perhaps the most evocative evidence for UPC comes from studies using dichoptic chromatic masking. In the original study (Moutoussis and Zeki 2002), which was later replicated a number of times and in various ways (Fahrenfort et al. 2012; Fogelson et al. 2014; Schuriger et al. 2010; 2015), each eye of the subject received a distinct stimulus, either a red face (or house) on a green background, or a green face (or house) on a red background. When the colours and shapes of the stimuli were complementary, they were synthesized, and the subject consciously experienced a yellow square. While the possibility of residual conscious perception of the faces (or houses) was ruled out in additional tests (e.g. the participants could not guess what is shown to them even if they were promised money for a correct answer), it was found that the exposition of the stimuli activated brain areas associated with recognition of faces (or houses). In control conditions, where the face (or the house) was presented to both eyes and the subjects were consciously aware of the stimuli, the activity in the same areas was found. The only difference was that in the control condition the activity was stronger. The researchers conclude that, in the experimental condition, conscious perception of yellow squares was accompanied with unconscious perception of red/green faces (or houses) on green/red backgrounds.

It is worth noting at this point that one's views about colour and perceptual character determine not only (a) what one is going to think about UPC in general, but also (b) which specific findings are going to strike one as the most compelling (assuming that one already considers UPC a genuine possibility). If one conceives of colours in terms of how we are aware of them when we see them consciously, one is likely to find (Moutoussis and Zeki 2002) and its replications particularly impressive, just like me. If, on the other hand, one associates colour perception with perceptual discrimination (see e.g. Rosenthal 2015) one is going to be more impressed by (Norman et al. 2014).

To illustrate, consider a case described by Grimes (Grimes 1996). A saccade is an eye movement from one fixation to another, occurring two or three times a second. During the fraction of a second when a saccade is executed, the eye is functionally blind. What is commonly known as 'change blindness', occurs when a visual stimulus changes during a saccade, and consists in the inability of the subject to detect a difference in the stimulus that resulted from that change. In one of examples of 'change blindness' described by Grimes, a parrot occupying over 25% of a picture changed its colour from brilliant green to brilliant red when the subject's eyes were executing a saccade. Despite looking at the picture the whole time, 18% of the subjects failed to detect the resulting difference in the parrot's colour.

This finding can be invoked in support of UPC, but only if the following two claims are true: (a) a failure to notice/detect something is a good indication of lack of consciousness of that thing; (b) seeing a difference is 'seeing' in the same sense as seeing ordinary objects and properties in the environment (e.g. apples, moustaches, or colours). If (a-b) are true, it might be argued that the parrot case described by Grimes instantiates unconscious perception of a difference in colour. But, as Dretske (Dretske 2004) cogently argues, (a-b) are questionable. On Dretske's preferred explanation, the parrot case involves two subsequent conscious colour perceptions plus lack of recognition that the colour is different on the two occasions (the lack of recognition is due to lack of perception of a change of the parrot's colour caused by the

fact that the colour changed during a saccade). While I agree with Dretske's interpretation, in the following sections I argue that UPC does not hinge on (a-b).

There is much more evidence that could be adduced in favour of UPC. Here I just gave a couple of examples. For a more extensive review of experimental paradigms providing support for UPC, see (Skrzypulec 2021). Skrzypulec also raises a worry concerning the move from all that evidence to UPC. Later in the paper (Sect. 4.3) I discuss that worry and offer a way of dispelling it.

3 Two initial objections

Currently available evidence does not establish UPC beyond reasonable doubt. Even though UPC is certainly one of the most plausible explanations of what happens in the described experimental situations, as it stands, it remains an optional and controversial hypothesis. The sceptic regarding UP is likely to raise the following objections:

(O1) The putative cases of unconscious perception are not really instances of perception, at least not in the relevant sense of 'perception'.

(O2) Even if the putative cases of unconscious perception are instances of perception, the subjects might be residually and/or transiently conscious of the stimuli.

Let us examine both objections in turn.

3.1 Is it really perception?

The current debate about UP proceeds on the assumption that the definition of perception offered by Burge (Burge 2010, 397–99) is correct³. According to that definition, perception is an objective sensory representation by the individual. The main bone of contention is whether any of the putative cases of unconscious perception meets each of the conditions mentioned in the definition. Both sides agree that if a mental state fails to meet Burge's conditions, it is not a genuine instance of perception.

However, a doubt can be cast on UP by questioning this common assumption. One may argue that Burge's definition does not lend itself to distinguish genuine perceptions from non-perceptual states. If so, the putative cases of unconscious perception may not be genuine cases of perception even if they meet Burge's conditions. For example, Taylor (Taylor 2020) argues that the debate mistakenly presupposes 'perceptual kind essentialism', i.e. the view that there exists a set of properties the possession of which is necessary and sufficient for a mental episode to be a member of the kind 'perception'. According to Taylor, this sort of essentialism is at odds with the current consensus in philosophy of science that essentialism about biological and psychological kinds is fundamentally wrong. After examining a number of potential examples of unconscious perception, he concludes that 'all of the putative cases of unconscious perception are more plausibly seen as cases where it is indeterminate whether the mental episode in question is perceptual' (Taylor 2020, 13).

³ Not all disputants endorse Burge's view. The definition has been assumed for the sake of the argument (Block and Phillips 2017, 185–86; Phillips 2016, 420).

But the debate about essentialism is hardly settled. In recent years, various authors have defended biological natural kind essentialism (Austin 2019; Devitt 2021), argued that essentialism about natural kinds is established by science and semantics (Nimtz 2021), and suggested that the homeostatic property cluster view (Taylor's preferred alternative to essentialism) is compatible with essentialism (Bird 2018). What is more, French and Phillips (Phillips 2018b; French and Phillips forthcoming) have recently argued that perception should be considered a manifest kind as opposed to scientific kind (see Sect. 4.3 and 5.2), whereas Burge (Burge 2010, 389–90) writes that 'although seeing is in a sense a natural kind, it is a hybrid kind' because 'a psychological state [...] depends for being a seeing on entities and causal relations beyond the psychology of the individual.'

Even if essentialism about biological and psychological kinds eventually turns out untenable, Taylor's conclusion is still not mandatory. This is because philosophy of perception is a branch of metaphysics, not philosophy of science. It is far from obvious that anti-essentialism in philosophy of biology or philosophy of psychology mandates anti-essentialism in metaphysics of perception. Of course, a philosopher of perception with particularly strong naturalist inclinations might insist that metaphysics of perception should be subordinated to empirical science (see e.g. Burge 2005). On that view, if scientists endorse anti-essentialism about perception, metaphysicians should follow suit. Nevertheless, a philosopher of perception may as well think outside the box of naturalism. While accepting the naturalist desideratum that a theory of perception should fit and predict empirical data, they may also have a reason to think that their theory of perception should be inconsistent with, say, Berkeleyan idealism, or Humean scepticism (cf. Martin 2006, 355; S. Overgaard 2011, 10). That would impose certain limitations (i.e. necessary and sufficient conditions) on what perception can be. Crucially, such limitations have nothing to do with the purpose and procedure of biology, psychology, or any other empirical enterprise. If so, the subject matter of philosophy of perception is not identical to the subject matter of empirical science of perception, albeit the two partially overlap⁴. From this perspective, essentialism in metaphysics of perception is consistent with anti-essentialism in philosophy of science. Even if empirical evidence renders the status of UP indeterminate, there are other criteria that might help us settle the issue. For instance, the relational theory of perception requires that a mental state must instantiate the relevant sort of perceptual relation in order to count as genuine perception.

Suppose that Burge's definition does constitute an adequate criterion for determining whether a mental state is perceptual. If so, one can raise O1 against UP by calling into question the objectivity of unconscious perception, and/or the involvement of sensory representation in unconscious perception, and/or the attribution of the putative unconscious perception to the individual. The critics of UP usually direct O1

⁴ My suggestion is that philosophy and science pertain to distinct aspects of a single explanandum, and thereby deliver distinct yet complementary explanations of it. I agree with Burge that seeing is a hybrid kind, but I also think that the criteria for membership in this hybrid kind cannot be established without making some metaphysical assumptions. This is different from saying that philosophy is concerned with perception *qua* 'manifest kind', as opposed to perception *qua* 'scientific kind' investigated by science (Phillips 2018b, 478). The latter approach seems to imply that the explananda of philosophy and science are entirely distinct.

at the latter. They argue that unconscious perceptual processing is not an instance of perception because there is no good reason to attribute it to a person. While conscious perceptual experience is a personal state, unconscious perceptual processing is merely a sub-personal state. This is because the former guides action, whereas the latter merely elicits behavioural reactions akin to conditioned responses (cf. Block and Phillips 2017; Phillips 2018b).

But this challenge rests on dubious assumptions. Empirical psychology teaches us that people commonly fail at identifying the causes of their own decisions and actions (see e.g. Nisbett and Wilson 1977; Schwitzgebel 2008; 2012). If introspective accessibility were necessary for guiding action, conscious accessibility would arguably be necessary for that as well. Since the former is not necessary, one may doubt whether the latter is. While this doubt gains some support from evidence about blindsight (Danckert and Rossetti 2005), there is also evidence suggesting that deliberating is often more successful when it is not conscious (Dijksterhuis 2004; 2006; Dijksterhuis and van Olden 2006; Mealor and Dienes 2012), and there are compelling reasons to think that unconscious processing plays as important a role in control and decision making as conscious experience does (see e.g. Haggard and Libet 2001; Clark 2007; Suhler and Churchland 2009; Wu 2013; Shepherd and Mylopoulos 2021)⁵. Hence it is implausible that being consciously available to the subject is requisite for guiding action.

Phillips (Phillips 2018b, 501) argues that unconscious influences on action are better viewed as subpersonal. However, it is questionable whether the distinction between personal and subpersonal *levels of explanation* yields a corresponding distinction between personal and subpersonal *states* (Drayson 2012), and there are no clear obstacles to cross-level explanations (Bermúdez 2000; Wong 2014).

The detractor may admit this, yet nonetheless insist that the term ‘perception’ is ambiguous when used in reference to an explanandum. Depending on the context, ‘perception’ refers to either perceptual experience or perceptual processing. Even the proponents of UP notice that this sometimes makes philosophers of perception talk past each other (Nanay 2017, 2–3). Notwithstanding, it does not follow that the two explananda must be pursued separately, nor that their respective explanations cannot be informed by each other. Although the primary focus of experiments cited in Sect. 2 is ‘perception’ *qua* ‘perceptual processing of chromatic information’, the results of those experiments can still contribute to explaining ‘perception’ *qua* ‘perceptual experience’. And while perceptual processing and perceptual experience often function as distinct explananda, it does not follow that they cannot be regarded as aspects of a single explanandum, namely perception in general.

My response to O1 does not rely on any necessary or sufficient conditions a mental state has to meet in order to count as perception. Instead, it shows that there is no undisputed theoretical basis for thinking that the putative cases of unconscious perception could not possibly be of the same fundamental kind as ordinary conscious seeing.

⁵ This is not to say that unconscious processing plays the same role as conscious experience.

3.2 Is it really unconscious?

O2 expresses a standard methodological worry that applies to all evidence for unconscious perception: it cannot be ruled out that the observed brain activity and behavioural reactions indicate residual and/or transient conscious experience of the stimulus, not unconscious perception.

In empirical experiments, the key terms (e.g. ‘consciousness’, ‘perception’) are defined operationally, i.e. in terms of the ways in which the subjects behave in specific experimental circumstances (see e.g. Yang et al. 2014). When the subject reports that they have not seen the stimulus despite performing with above-chance accuracy in tasks requiring perception of the stimulus, it is concluded that unconscious perception has occurred. However, this sort of evidence for UP can be contaminated with conservative response bias: the stimulus is in fact consciously accessible to the subjects, but the fact that it barely stands out from the noise disposes them to report that they have not seen anything (when asked a ‘yes/no’ question) or that the stimulus has not changed (‘same/different’ question) (Peters and Lau 2015; Phillips 2016; 2018b).

The researchers do their best to prevent this. For example, instead of asking the subjects simple ‘yes/no’ questions, they use graded scales that allow the subjects to express more precisely their degree of confidence about what they did and did not experience. Since unconscious perception is indexed exclusively to the lowest rating on the scale, the availability of intermediate ratings decreases the probability that a weak glimpse of the stimulus occurs without being reported (see e.g. Ramsøy and Overgaard 2004; M. Overgaard and Grünbaum 2011).

Although effectiveness of such measures is admittedly limited (Irvine 2012), for all we know, the problem may not occur in every single study. And since methods of suppressing consciousness are systematically improved, the denial of UP based on potential suppression failure hinges on the outcome of future research.

Another preventive measure consists in asking the subjects to identify the category of the stimulus, and then to wager for monetary rewards on the accuracy of that identification. If the accuracy is at chance while brain scans indicate activity in the relevant areas, it seems reasonable to conclude that the stimulus is perceived unconsciously (see e.g. Schurger et al. 2010).

This brings us back to O1. While the subject’s inability to discriminate the stimulus arguably settles that the stimulus is not consciously available, it may also be taken as suggesting that the stimulus is not perceived at all. For the mere registration of the stimulus by the visual system, evidenced by a brain activity or a reaction time variance, is merely a state of the visual system, not a state of the subject. Still, O1 is inconclusive (see Sect. 3.1).

The last thing to note about O2 is that the charge of suppression failure is much more probable on the assumption of phenomenal overflow (a hypothesis according to which conscious availability exceeds reportability). If overflow is true, the conservative response bias constitutes a compelling alternative to UP. If overflow was shown false, however, this alternative explanation would be much less credible. I say more about overflow and its relation to UP in Sect. 6.

To sum up, while the worry expressed by O2 is serious, it is also inconclusive. As things currently stand, O2 does not suffice to rebut UP⁶.

4 Is it really colour?

This section concerns an objection that targets UPC specifically:

(O3) Even if the cases presented in Sect. 2 are instances of unconscious perception, what is perceived is not colour, at least not in the relevant sense of ‘colour’.

In what follows, I outline a conception of colour that underlies O3 and explain how that conception might be used to undermine UPC (Sect. 4.1). Then two mutually exclusive replies to O3 are discussed. One reply rejects the conception of colour that underlies O3, the other accepts that conception as harmless to UPC. I argue that the former reply is not mandatory (Sect. 4.2) and recommend the latter (Sect. 4.3). Before concluding (Sect. 4.5), I defend my recommendation from criticism ensuing from Rosenthal’s influential remarks about colour and unconscious perception (Sect. 4.4).

4.1 How to think about colours?

O3 presupposes a conception of colour that undermines UPC. Whatever that conception might be, it discredits the evidence for dissociation between colour and consciousness. At minimum, the conception meets the following requirement:

(R) Colours ought to be conceived of in terms of how we are aware of them when we see them consciously⁷.

It seems uncontroversial to say that, from the first-person perspective, colours appear to be simple, homogenous properties. It is perhaps more contentious to say that colours appear to be non-physical. And yet as soon as the apparent simplicity of colours is taken at face value, they no longer seem to fit into the physicalist description of the world.

If these remarks are correct, R leads to primitivism, i.e. the view that colours are simple, *sui generis* properties, not identical to physical properties. O3 thus tells us to conceive of colours as simple properties because this is how they appear when consciously experienced.

Before explaining how primitivism can be used to undermine UPC, I need to address two controversies surrounding R. The latter is no doubt a controversial claim, but not as controversial as it is sometimes taken to be.

First, R might suggest that the nature of colour is exhausted by the qualitative character that determines the what-it-is-like-ness of conscious colour experience. If R indeed had this consequence, it would presumably undermine any attempt to explain

⁶ For more reasons why O1 and O2 are inconclusive, see (Berger and Mylopoulos 2019).

⁷ As I understand it here, R expresses the view that colour experience is non-conceptual. It conceives of colours in terms of how they appear to the subject irrespective of what colour concepts are in the subject’s possession, what the subject knows about colours, or what the subject can or cannot learn about them from perceptual experience.

the qualitative character of colour experience in naturalist or physicalist terms (cf. Rosenthal 2015, 33–34).

Nevertheless, R *does not entail* that colours have no reality independent of how we are aware of them. R is only committed to the claim that the qualitative character of colour revealed in conscious colour experience is an essential component of the colour's nature. This is neutral as to whether colours have any aspects other than the ones we are aware of when we perceive colours consciously, aspects that might admit of a naturalist or physicalist explanation (cf. Allen 2016, Chap. 7; Byrne and Hilbert 2007).

Second, conceiving of colours as we are aware of them seems to validate the claim that colours are intrinsically conscious, which in turn suggests that visual consciousness affords the perceiver privileged epistemic access to them. It follows that undetectable colour inversion is possible. Since every perceiver has such privileged access only to their own colour experiences, it is possible that my colour experiences are radically different from yours (where you see green, I see red, etc.). Because this consequence of R renders any third-person investigation of colours infeasible, we should reject R (see e.g. Rosenthal 2005c, 143–44; 2010, 380).

However, one can consistently accept R and deny the claims about privileged access and spectrum inversion. For R *does not* entail that colours are intrinsically conscious. While conceiving of colours as we are aware of them does lead to primitivism, the latter is compatible with the claim that colours *qua* phenomenal qualities exist independently of consciousness. For example, some primitivists argue that colours are mind-independent (Campbell 1993; Kalderon 2007; Allen 2016), which means that colours are also consciousness-independent⁸. In fact, even if primitivism were combined with mind-dependence of colours, it still would not follow that colours are invariably conscious (cf. Section 5.3).

Having the two controversies out of the way, we can consider how primitivism can be used to undermine UPC. There are at least three different routes from primitivism to O3.

The first route starts with a widespread intuition that there is an inextricable link between colours and consciousness. The intuition supports subjectivism, i.e. the view that colours are primitive properties of conscious experiences. On this view, the term 'colour' picks phenomenal qualities, i.e. that stuff the neuroscientist Mary has discovered for the first time after leaving her black-and-white apartment (Jackson 1986). If subjectivism is true, colours are invariably conscious, and UPC is false. Insofar as evidence for UPC does not deliver a reason to think that the subjects have unconsciously seen colours *qua* phenomenal qualities, it fails to support the claim that colours can be seen unconsciously. If O3 is correct, evidence for UPC can be explained away by saying that certain physical features of the environment, which usually correlate with (and most likely enable) conscious colour experience, admit of being unconsciously registered and processed by the visual system⁹.

⁸ This is not to say that these thinkers endorse UPC. I am only suggesting that their take on primitivism can be consistently combined with UPC. See also Sect. 5.2 and 5.4.

⁹ The subjectivist might adduce the (Breitmeyer, Ro, and Singhal 2004) metacontrast experiment, in which white disks were found to act more like green primes than like blue primes when seen uncon-

The second route from primitivism to O3 goes *via* epiphenomenalism. If colours are non-physical (as primitivism tells us), and all causes are physical causes (as physicalism requires), then colours are causally inefficacious. If epiphenomenalism is true, empirical evidence for UPC is unattainable. If colours do not have any causal effect upon the physical, observation of behaviour or brain activity cannot deliver any evidence of colour perception. For epiphenomenalism entails that no observable behaviour or brain activity could possibly be caused by colours. On this view, evidence mentioned in Sect. 2 cannot be interpreted with reference to colours. Even if this does not disprove UPC, it renders the hypothesis unmotivated.

The third way of substantiating O3 goes through eliminativism. The latter ensues from the combination of R and an error theory about the qualitative features that visual experience seems to reveal. If the concept ‘colour’ refers to qualitative features distinct from anything physical (primitivism), and there are no properties other than physical properties (physicalism), then visual experience is misleading, and ‘colour’ is an empty category. But colours have to exist if they are to be seen unconsciously; UPC presupposes colour realism. No colours, no UPC.

Without R, none of the three arguments for O3 gets off the ground. But R has a substantial justification: any interpretation of UPC that fails to meet this requirement can be accused of confounding (a) that which the concept ‘colour’ really picks (i.e. phenomenal qualities) with (b) enabling conditions of colour experience (e.g. reflectance of the perceived surface, activity in the relevant brain areas). On top of that,

‘The connections reported in text books between wave-length properties of light and colours, and between the ability to see colours and eye structures are established by relying, at some point or other, on judgements of colour based on things looking one or another colour. But that presupposes we can independently trust some judgements about colour based on things looking a certain colour. If that’s no longer a given, it is obscure, to say the least, how information from science might be marshalled to support colour judgements.’ (Jackson 2019, 821).

Hence any response to O3 has to take R into consideration. As far as I can see, the UPC-theorist has two mutually exclusive types of response to choose from: *overriding* and *undercutting*. The former rejects R by telling a revisionary diagnostic story about it. The latter embraces R as harmless to UPC (cf. Pritchard 2016, 16–17).

4.2 The *overriding* response to O3

The rejection of R blocks each of the three arguments for O3. If colours are not conceived as we are aware of them, there is no obstacle to conceive of them as reducible

sciously, and more like blue primes than like green primes when seen consciously. The authors take this as evidence that unconscious perception only tracks wavelength features of the reflecting surface, and thereby does not make colour *qua* phenomenal quality perceptually available (see also Lamme 2015, 10). However, there is an alternative interpretation. Consider colour pluralism, i.e. the view that every object has more than one colour all over at the same time. Which of the object’s colours is perceived on a given occasion turns on the circumstances of perception, including the duration of perception, and the state of the perceiver. On this view, the difference between unconscious and conscious trials can be explained by suggesting that unconscious and conscious perceptions rendered different colours *qua* phenomenal qualities of the disk perceptually available to the subjects. Colour pluralism has both primitivist (see e.g. Kalderon 2007) and reductionist (see e.g. Mizrahi 2006) espousers.

to physical properties. If colours are physical, there is no reason to think that they are intrinsically conscious, nor that they are causally impotent, nor that they should be eliminated. This is an *overriding* response, in that it concedes to the objector that R is inconsistent with UPC but discards R on independent grounds.

While at first glance the *overriding* response seems compelling, there are good reasons not to respond in this way.

First, as we have seen in the previous section, R does not entail that colours have no reality outside of what we are aware of when we see them consciously, nor does it render colours invariably conscious. Consequently, R does not entail the possibility of spectrum inversion. And even if it did, pointing this out would not persuade the objector, as the latter explicitly uses undetectable colour inversion to motivate their view. It is incumbent on the reductionist to explain in physical terms *why* colours appear the way they do when they are seen consciously. This may be infeasible. Unless the relevant explanation is provided, the objector will discard as *ad hoc* the hypothesis that unconscious perceptual processing of chromatic information about the environment is an instance of colour perception¹⁰.

Second, the move from R to epiphenomenalism is not straightforward either. Conceiving of colours as primitive does not necessarily render them causally inefficacious. To reach that conclusion, the objector has to supplement their argument with certain optional and controversial view of causation. Granted, if colours are distinct from physical properties, their causal influence on the physical cannot be explained in terms of mechanistic conception of causation. But mechanicism is not the only game in town. As Allen (Allen 2016, 102–4) has recently argued, causal powers of colours are best understood in terms of difference-making, not in terms of mechanistic processes. As long as this proposal is not refuted, the mechanistic approach to colour causation is not compulsory¹¹.

Third, the claim that colours are not physical does not entail that they do not exist (*pace* eliminativism). The entailment is true if everything that exists is either physical or reducible to the physical, which is disputable. Inasmuch as eliminativism and epiphenomenalism are radically revisionary with respect to visual phenomenology, they are arguably best viewed as the last resort, i.e. one should abstain from considering them until all less revisionary options are shown untenable. A compelling anti-reductionist alternative is offered in Sect. 5.4.

In view of the foregoing, conceiving of colours as we are aware of them is consistent with UPC. Insofar as the UPC-theorist is willing to (i) endorse primitivism and (ii) deny that colours are intrinsically conscious, they can give an *undercutting* response to the objection, namely to embrace R as consistent with UPC.

¹⁰ Although the possibility of spectrum inversion is usually mentioned in support of subjectivism or epiphenomenalism, it could be used to motivate eliminativism as well. The relevant argument would consist in condemning colour realism for validating such untoward possibilities.

¹¹ In a similar vein, one might object that primitivism compromises mind-independence of colours, since the objectivist version of primitivism is committed to a counterintuitive claim that parts can cause their wholes (Kalderon 2011, 255). And yet closer inspection reveals that such possibility is perfectly coherent, and causal relations of this sort are quite common (Friend 2019).

4.3 The *undercutting* response

It may seem at this point that choosing between *overriding* and *undercutting* strategies reduces to picking a side in the debate between primitivism and reductionism in metaphysics of colour¹². If it did, it would be a tough choice because no side in that debate clearly outweighs the other. As long as the primitivism vs. reductionism controversy remains unresolved, both types of response to O3 are available.

An anonymous referee suggested that UPC can be defended from O3 just by denying that primitivism undermines UPC, without endorsing either primitivism or reductionism (even provisionally). Yes, but this would leave UPC underspecified. Recall that the objective of the present discussion is not only to defend UPC from O3, but also to determine which reading of UPC is the most plausible. To achieve the latter goal, I have to choose between the primitivist and the reductionist reading of UPC.

For two reasons, I think that the *undercutting* strategy is preferable. The first reason is that the *undercutting* strategy is less costly. In fact, since it comes down to a simple observation that certain claims fail to follow straightforwardly from primitivism, the UPC-theorist gets it basically for free. The *overriding* response, in contrast, requires showing that colours reduce to the physical, which is a very hard thing to do and, for all we know, might be infeasible. Importantly, this is not to say that reductionism cannot accommodate UPC, or to pretend that primitivism does not carry its own burden of objections. My point is only that the *overriding* strategy makes UPC dependent on reductionism, whereas in fact UPC does not depend on reductionism at all. For this reason, as things now stand, the *undercutting* strategy puts UPC in a better dialectical position. The proponents of UPC can and should grant to the objectors that primitivism is true, even if only temporarily¹³. Should reductionism ultimately prevail, the UPC-theorists can retreat to reductionism. But as far as the plausibility of UPC is concerned, and as long as reductionism remains merely one of the options in play, it would be precipitant to assume reductionism in advance.

No doubt this argument for the *undercutting* strategy will be questioned. For example, one might argue that problems of primitivism countervail (or exceed) the problems of reductionism, so the cost of primitivist UPC is the same as (or higher than) the cost of reductionist UPC. But that would be missing the point of my suggestion, as the advantage of primitivist UPC over reductionist UPC is purely dialectical. Primitivist UPC is less costly not because primitivism is more plausible than reductionism, but because O3, arguably the most serious objection to UPC, assumes

¹² Pautz (Pautz 2009) notes that both primitivism and reductionism have response-independent and response-dependent versions. On the response-independent versions of both views, perceived objects have colours independently of being perceived. This is consistent with (but does not entail) UPC. On the response-dependent versions, colours are dispositions of otherwise uncoloured mind-independent objects to cause colour experiences in perceivers. Thus understood, response-dependent views are inconsistent with UPC because they define colours in terms of conscious experience. To accommodate UPC, the response-dependence theorist would have to modify their position by stipulating that colours are dispositions of otherwise uncoloured mind-independent objects to cause the production of colours in perceivers and that colours are not invariably conscious.

¹³ Unless they have other theoretical commitments, not related to UPC, that compel them to endorse reductionism. If my arguments in this section are correct, UPC itself is not such a commitment.

primitivism. Hence adopting primitivism affords the UPC-theorist the cheapest and most effective response to that objection.

Consider a claim *Z*, neutral with respect to the primitivism vs. reductionism divide, yet dependent on UPC. According to my recommendation of the *undercutting* approach, UPC gives the *Z*-theorist a reason to prefer primitivism over reductionism. For embracing primitivist UPC allows the *Z*-theorist to overcome O3 (and thereby to secure UPC as a motivation for *Z*) with virtually *zero* effort. In this regard, primitivism renders UPC more plausible than reductionism.

The second reason to opt for the *undercutting* strategy comes from the observation that the ramifications of UPC for the debate about UP can vary depending on how UPC is specified. In particular, primitivist UPC is more problematic for the sceptic about UP than reductionist UPC. I base this diagnosis on two examples.

First, consider an argument recently put forward by Phillips and French (Phillips 2018b; see also French and Phillips forthcoming):

(P1) Perception in its ordinary sense is essentially tied to its phenomenal nature, which manifests itself in perceptual experience.

(P2) Unconscious perception lacks that phenomenal nature.

(C) Unconscious perception is not perception in the ordinary sense.

This argument does not work against primitivist UPC, because the latter entails that the phenomenal nature of perception can occur outside of consciousness. Reductionist UPC, in contrast, is susceptible to the challenge expressed by the argument because it allows Phillips and French to insist that unconscious colour perception lacks the phenomenal nature characteristic to perception in its ordinary sense.

Second, according to Skrzypulec (Skrzypulec 2021), for all that current evidence for UPC tells us, the content of unconscious perception of a colour might be significantly impoverished in comparison to the content of conscious perception of that colour. Assuming reductionist UPC, one could use this observation to buttress the position of Phillips and French, or at least argue that UPC is only partially true. Contrast this with the interpretation of UPC in terms of objectivist colour primitivism and Pure Relationalism (i.e. the view that perception is just a relation between the subject and the perceived item (Stoneham 2008)). On this interpretation, consciousness has no impact on how much of the environment one perceives; it can only influence how much of what one perceives can be used for cognition and action. It follows that there is no difference between conscious and unconscious colour perception with respect to what properties are perceived.

The foregoing provides some substantial reasons to think that the primitivist reading of UPC is more resilient to objections than the reductionist reading, which suggests that the *undercutting* strategy is preferable over the *overriding* one.

To challenge this verdict, one would have to raise some objections against UPC that reach the primitivist reading but fall short of the reductionist interpretation (importantly, those would have to be objections directed specifically against UPC, not just objections to primitivism in general). Alternatively, one would have to reject my criterion of resilience to objections, offer a different one, and then argue that that other criterion renders reductionist UPC preferable over primitivist UPC.

4.4 Defending the *undercutting* response

Since my proposal presumes R, it can be questioned on the same grounds as the latter. To reiterate, the rationale behind R is that it blocks the charge of confounding (a) that which the concept ‘colour’ really picks (i.e. phenomenal qualities) with (b) enabling conditions of colour experience. The reductionist will of course disagree. For example, they might say something like this:

‘What intuitively resists description in mathematical terms is the qualitative aspect of physical color properties *as we are conscious of it*. When these properties are not consciously perceived, there is nothing about them that resists mathematical description. We have no intuitive sense that something is lost if we describe the color properties objects have independently of being seen in terms of surface reflectance properties and the like.’ (Rosenthal 2005b, 160–61, emphasis in original).

According to Rosenthal, there is no reason to assume that colours are as we are aware of them when they are seen unconsciously. If so, it is also unjustified to think that colours elude mathematical treatment. But this is not intended to suggest that colours simply lack the qualitative nature that conscious perception seems to reveal. Instead, Rosenthal contends that:

‘Qualitative states occur without being conscious, without, that is, there being anything at all it’s like to be in those states. So there is no reason to suppose that, when there is something it’s like to be in a qualitative state, what it’s like to be in it reveals all its qualitative character.’ (Rosenthal 2005b, 172).

Drawing on these two remarks by Rosenthal, the reductionist can argue that R is unmotivated (i.e. conceiving of colours as we are aware of them is unjustified), and on this basis urge that the *overriding* strategy is the only legitimate way to defend UPC from O3.

However, this case against R fails. To see why, notice first that both of Rosenthal’s remarks presuppose that the qualitative character of colours does not determine what it is like to consciously perceive colours. For if the qualitative character of colours *does* determine what it is like to consciously perceive colours, (i) it is *not true* that when colours are not consciously perceived there is nothing about them that resists mathematical treatment (i.e. the first remark is false), and (ii) the claim that colour perception can occur unconsciously *does not entail* that the qualitative character of conscious colour perception is not revealed by what it is like to undergo that perception (i.e. the second sentence in the second quoted remark fails to follow from the first sentence in that remark).

Now, the crux of the matter is that the presupposition behind Rosenthal’s remarks begs the question against R. O3 says that we should conceive of colours as we are aware of them (R), which is tantamount to saying that we should conceive of colours in terms of what it is like to consciously experience them. As the proponents of R

understand it, this means that colours themselves determine what it is like to see colours, which is antithetical to the presupposition behind Rosenthal's remarks¹⁴.

The presupposition Rosenthal makes in the two quoted remarks ensues from his HOT theory of consciousness¹⁵. According to that theory, the 'what-it-is-like' of a conscious perception of a colour is determined by a higher-order thought the perceiver has about that perception (Rosenthal 2005a, 186). When evaluated from this perspective, R may indeed seem unjustified, at least insofar as the nature of colour is supposed to be independent of anyone's thoughts about colour.

Crucially, however, this is not how the objector understands R. For them, the 'what-it-is-like' of a perception is determined by the qualitative aspects of that perception (Block 2011b, 425; Rosenthal 2011, 434)¹⁶. This comparison indicates a clash between the HOT theory (at least Rosenthal's version of it) and primitivism. Contrary to the HOT theory, primitivism conceives of colours in terms of what it is like to consciously experience them. That being so, simply assuming the HOT theory begs the question against primitivism.

The comparison also explains why the charge that R construes colours as invariably conscious is a *non sequitur*. As we have seen above, primitivism is compatible with, but does not entail, the claim that colours are intrinsically conscious. Conceiving of colours in terms of what it is like to consciously experience them presupposes that consciousness is a source of knowledge about colours¹⁷. Still, it does not follow that colours are inextricably bound to consciousness. It would follow if the HOT theory were true, but that theory is not mandatory.

In order to impose Rosenthal's understanding of 'what-it-is-like' on the primitivist, the reductionist would have to first prove that the HOT theory is indispensable. We have now come back to where we started. This is not a place to assess the HOT theory. But when the *overriding* strategy is motivated by Rosenthal's view, UPC inherits the burden of objections to the HOT theory. No such problem occurs when O3 is answered in the *undercutting* way. The *undercutting* strategy is thus less costly than the *overriding* one. So my recommendation remains untouched. That said,

¹⁴ This is not inconsistent with UPC. The R-theorist can accommodate UPC by suggesting that colours realize their potential to determine what it is like to consciously perceive them only when they are consciously perceived.

¹⁵ The diagnosis is also motivated by Rosenthal's homomorphism theory of mental qualities (a.k.a. Quality Space Theory, QST). I explain the latter briefly in Sect. 5.3. While the HOT theory and the homomorphism theory are strictly speaking distinct ideas, they arguably converge in Rosenthal's renunciation of R. Note also that it is irrelevant for present purposes whether the truth of the presupposition behind the quoted remarks depends on QST and/or HOT theory; all that matters is that the presupposition is optional and controversial.

¹⁶ Rosenthal maintains that (a) what it is like to see a colour is distinct from (b) the qualitative aspect of that seeing. Block disagrees; on his view, (a) is constituted by (b). And yet both thinkers endorse UPC (see e.g. Block 2016; Rosenthal 2005c). Because UPC follows straightforwardly from Rosenthal's account, the critic of UPC is clearly going to side with Block in his dispute with Rosenthal. But the fact that Block can consistently endorse UPC and reject Rosenthal's view demonstrates that UPC *does not* depend on Rosenthal's HOT theory of consciousness.

¹⁷ This is not to say that visual consciousness is the only source of knowledge about colours, nor that knowledge about colours acquired by conscious seeing is exhaustive (cf. Allen 2016, Chap. 7; Byrne and Hilbert 2007).

Rosenthal's view is an important reference point for present considerations, and I will come back to it in Sect. 5.3.

4.5 Mid-conclusions

What I have said so far justifies two conclusions. First, since UPC is a coherent and substantially motivated hypothesis, and none of the three main objections against it is decisive, its philosophical ramifications cannot be ignored. Hence what I shall say in the next section cannot be discarded as mere speculation. Second, while O3 does not undermine UPC, an important lesson to be learned from O3 is that, as things now stand, the plausibility of UPC increases when colours are conceived in terms of how we consciously experience them. Inasmuch as that conception of colour favours primitivism, the latter is more congruent with UPC than reductionism.

5 Phenomenal qualities and consciousness

Textbook definitions describe the phenomenal character of a perceptual experience as a set of properties determining what it is like to have that experience from the first-person perspective. Colour is a paradigmatic example of such a property. For example, what it is like to see blue is at least partially determined by the colour blue. Various theories of phenomenal character specify this general thought in different ways. For present purposes, it is useful to categorize them based on how they respond to the following questions:

(Q1) Is the claim 'a perception P has a phenomenal character' equivalent to the claim 'a perception P is conscious'?

(Q2) Is the phenomenal character of perception constituted by mind-dependent properties, or is it constituted by mind-independent properties?¹⁸

Table 1 A classification of theories of the phenomenal character of perception

	Q1: equivalence	Q1: no equivalence
Q2: mind-dependent	equivalence internalism	inequivalence internalism
Q2: mind-independent	equivalence externalism	inequivalence externalism

¹⁸ For simplicity, I set aside the question of whether phenomenal character is cognitively penetrable. According to the cognitive penetration hypothesis, the phenomenal character of perception can be and often is modified by higher-order mental states of the perceiver. It is controversial whether the phenomenal character of colour perception is penetrable in this sense (MacPherson 2012; Brogaard and Gatzia 2017). If it is, Q2 can be reformulated as follows: is the phenomenal character of perception constituted *exclusively* by mind-dependent properties, or is it *also* constituted by mind-independent properties? Alternatively, it might be suggested that Q2 concerns only the genuinely perceptual (i.e. presentational) layer of phenomenology, leaving aside the 'cognitive' layer that might be superimposed on the perceptual layer due to cognitive penetration.

Consequently, conceptions of phenomenal character divide into four types (see Table 1):

The goal of this Section is twofold: (i) to show that inequivalence views are better suited to accommodate UPC than equivalence views (i.e. that saying ‘no’ to Q1 renders UPC more plausible); (ii) to highlight the possibility of inequivalence externalism, a perfectly viable position that hasn’t been acknowledged so far in the debate about unconscious perception.

There are three important things to note about the typology in Table 1. First, it is not exhaustive. Since O3 can be responded in two different ways (*undercutting* and *overriding*), each of the four conceptions in Table 1 has two versions: one that accepts R and one that denies it. Therefore, the UPC-theorist has at least eight options to choose from. However, because of the reasons given in Sect. 4, I shall proceed on the assumption that O3 is responded in the *undercutting* manner, thereby confining the choice to four options. In Sect. 5.1–5.4 below, I consider some specific examples of these four general conceptions in order to determine which conception of phenomenal character yields the most plausible reading of UPC. This survey is not intended to be exhaustive. To keep the discussion within manageable bounds, I focus on the possibilities that are either already defended in the literature or can be viewed as natural developments of theories that are already present in the literature.

Second, by assuming that O3 should be dealt with in the *undercutting* way, I assume that what it is like to consciously perceive colours is determined by colours qua primitive *sui generis* properties. On this view, the qualitative character of a colour (i) is an essential component of that colour’s nature, and (ii) determines what it is like to perceive that colour (i.e. the phenomenal character of a conscious experience of that colour) when it is perceived consciously. Whether the qualitative character of colour is invariably conscious turns on the answer to Q1 (i.e. the question of whether the phenomenal character of perception is conscious by definition).

The assumption that R is true (which is an element of the *undercutting* strategy) effectively rules out Rosenthal’s view discussed in Sect. 4.4, according to which what it is like to consciously perceive a colour is determined by something extraneous to the qualitative character of that colour (e.g. by higher-order thoughts one has about that perception). But, as I show in Sect. 5.3–5.4, this does not mean that the qualitative character of colour is invariably conscious. The primitivist can respond negatively to Q1 by suggesting that the qualitative character of a colour has a potential to determine what it is like to perceive that colour, a potential that is only realized when that colour is perceived consciously. On this view, consciousness does not determine what it is like to see colours; it enables colours to determine what it is like to see them¹⁹.

Third, note that Q2 does not ask whether colour objectivism (i.e. the view that colours are mind-independent) is true or not. This is because Q2 concerns the constituents of the phenomenal character of perception, not colours as such. One can con-

¹⁹ This accords with the transparency thesis, according to which ‘experience reveals only the mind-independent objects, qualities and relations that one learns about through perception’ (Martin 2002, 378). While this claim is usually based on introspective grounds, recently it has received a powerful empirical validation (Weksler, Jacobson, and Bronfman 2021).

sistently hold that (i) colour objectivism is true and (ii) colours do not constitute the phenomenal character of colour perception (see e.g. Byrne and Hilbert 2003; Jackson 2019). The representational theory of perception construes phenomenal character as consistent with non-existence of what is represented. If colours *qua* objective properties of the environment are *represented* by the phenomenal character of colour perception, they do not constitute that phenomenal character. For colours to qualify as the constituents of the phenomenal character of colour perception, it has to be true that the phenomenal character of perception of a colour is *inherited* from the qualitative character of that colour (cf. Campbell 1993, 268). Hence the internalism vs. externalism distinction in Table 1 maps onto representationalism vs. relationalism distinction in metaphysics of perception.

Nevertheless, choosing between inequivalence internalism and inequivalence externalism does not boil down to picking a side in the representationalism vs. relationalism debate. In Sect. 6, I shall present an argument for inequivalence externalism that does not hinge on relationalism's standard motivations. Meanwhile, let us test each of the four accounts in Table 1 against the background of UPC. While the latter admits of being incorporated into all four accounts, some of them are much better suited to accommodate it than others.

5.1 Equivalence internalism

Equivalence internalism responds affirmatively to Q1: to say that a perception P has a phenomenal character is equivalent to saying that P is conscious. There is no such thing as unconscious perception with phenomenal character. As to Q2, the equivalence internalist maintains that the phenomenal character of perceptual experience is produced in the mind when the cognitive system is under a suitable causal influence of the perceived object.

The equivalence claim entails that a perception acquires its phenomenal character when the content of that perception becomes consciously available to the perceiver. This suggests that the same mechanism in the mind is responsible for a perception being conscious and for that perception having a phenomenal character. Alternatively, the equivalence internalist may argue that consciousness and phenomenal character are realized by two distinct mechanisms, both of which have to be concurrently activated for the phenomenal character to occur (Prinz 2012, 126–45). This also precludes the possibility of unconscious phenomenal character.

Whatever the equivalence internalist thinks of UP, they deny that unconscious perception has phenomenal character. They may allow for unconscious seeing of colours *qua* reflectance properties of the perceived scene, but they would most likely deny that colours *qua* primitive qualities can be seen unconsciously. This is because combinations of colour primitivism with equivalence internalism typically construe colours as something subjective (i.e. mind-dependent), e.g. as features of conscious perceptual experience. On this view, it makes no sense to say that the stuff Mary has seen for the first time after leaving her black-and-white room can be seen unconsciously. Equivalence internalism of this sort is therefore in conflict with primitivist UPC. Contrary to the latter, it holds that colours conceived of as we are aware of them when we see them consciously cannot occur outside of consciousness.

Equivalence internalism could be consistently combined with objectivist primitivism. On this view, colours are primitive mind-independent qualities of the environment (call them Q1) that are not perceived directly (i.e. presented), but represented in conscious perception by invariably conscious primitive qualities produced in the mind (call them Q2). This view can accommodate unconscious colour perception as unconscious representation of Q1 that does not involve Q2. However, it is unclear to me how this view could be motivated. Accommodating UPC does not seem to be a good motivation, not least because the view in question is susceptible to both the argument presented by Phillips and French, and to the worry raised by Skrzypulec (see Sect. 4.3). For these reasons, I set this view aside.

The equivalence internalist can countenance UPC only with the proviso that seeing colours unconsciously lacks phenomenal character. But, assuming that the combination of equivalence internalism with objectivist primitivism is not really an option, this take on UPC falls prey to O3 (see Sect. 4). When it comes to primitivist UPC, the equivalence internalist will most likely try to discard it by raising objections O1–3 against it. As we have seen in Sects. 3–4, those objections are inconclusive.

Alternatively, the equivalence internalist might agree with Nelkin (Nelkin 1989) that there is more than one way in which a mental state can be conscious. If ‘consciousness’ is a polysemous word, then perhaps putative instances of unconscious perception are unconscious in one sense, and conscious in another. A mental state like this can be considered conscious in the sense that it has a phenomenal character, which makes it a mental representation carrying some non-propositional, ‘imagistic’ content. It is nonetheless unconscious in the sense of not being an object of what Nelkin calls ‘second-order consciousness’. The latter is a mental representation with propositional content that enables direct, non-inferential access to first-order mental states. If this is correct, the truth of UPC depends on how the term ‘unconscious’ is specified.

Nelkin’s proposal seems particularly compelling in the context of a conjecture that at least some of currently popular theories of consciousness fail to share an explanandum, i.e. instead of offering competing explanations of a single phenomenon, they explain a number of distinct phenomena. Although that conjecture is not unreasonable, Nelkin’s idea does not help the equivalence internalist to accommodate UPC.

To see why, consider the well-known distinction coined by Block (Block 1995) between phenomenal and access consciousness. The former refers to the phenomenal character of conscious experience (the properties that determine what it is like to have an experience), the latter picks the availability of some mental content for reasoning and guiding action. It is easy to see that Nelkin’s first-order vs. second-order consciousness distinction overlaps, roughly at least, with Block’s phenomenal vs. access consciousness distinction. If so, implementing Nelkin’s view would amount to suggesting that unconscious perceivers are phenomenally conscious of the stimuli, yet fail to report it because they are not access-conscious of what is presented to them. This, however, does not do anything over and above O2, which is a legitimate yet inconclusive objection to UPC.

5.2 Equivalence externalism

Equivalence externalists also respond affirmatively to Q1, i.e. they reject unconscious phenomenal character. But contrary to equivalence internalists, they regard the phenomenal character of perception as composed of mind-independent properties.

A noteworthy example of equivalence externalism is orthodox relationalism, a version of the relational theory of perception. The latter construes the phenomenal character of perceptual experience as at least partially constituted by mind-independent properties of the perceived scene. This leads to the objectivist version of colour primitivism. If ‘the qualitative character of a colour-experience is inherited from the qualitative character of the colour’ (Campbell 1993, 268), there is no easy way around the claim that colours are primitive properties. Hence relationalists tend to view colours as primitive mind-independent properties.

Apart from endorsing the standard commitments of the relational theory of perception, the orthodox relationalist is likely to emphasize that that theory has been originally introduced to account for conscious perceptual experience, not for perception *per se*. Alternatively, they might insist that only conscious perceptual experience is perception in the ordinary sense, because our most basic grasp of the concept ‘perception’ is essentially tied to the way conscious perceptual experience appears from the first-person perspective (Phillips 2018b; French and Phillips forthcoming). They also believe that only perceptual experiences have phenomenal characters. The phrase ‘unconscious phenomenal character’ strikes them as an oxymoron, and they insist that the relationalist analysis does not apply to unconscious perception (cf. Phillips 2018b, 472). Consequently, they discard as misguided any critique of relationalism that accuses it of inability to explain unconscious perception. In Sect. 5.4, I contrast this kind of relationalism with an unorthodox formulation of the view.

Equivalence externalism is not well-suited to accommodate UPC. Since any account of UPC compatible with this view is going to entail that unconscious perception lacks phenomenal character, it will also be susceptible to O1 (given the idea that perception in its ordinary sense is identified in relation to its phenomenal nature) and to O3 (given R). For these reasons, the proponent of this view is likely to be sceptical about UPC and UP.

5.3 Inequivalence internalism

The inequivalence internalist approves of the idea of unconscious phenomenal character, and contends that phenomenal character is produced in the mind. On this view, perception acquires phenomenal character before the occurrence of perceptual experience. This approach is arguably more congruent with UPC than the preceding two. While equivalence theories view the evidence for UPC as a problem to be solved, for inequivalence theories that evidence is a source of support.

For example, Marvan and Polák (Marvan and Polák 2017, 6–8) argue that evidence for UP (including evidence for UPC) constitutes a strong reason against equivalence views. If colours constitute the phenomenal character of colour perception, and colours can be seen unconsciously, it follows that the phenomenal character of colour perception is independent of consciousness.

The basic idea behind the inequivalence approach is simple. In case of a conscious perceptual experience, the phenomenal character is conscious and determines what it is like to have that experience. In case of an unconscious perception, the phenomenal character remains unconscious, which is why there is nothing it is like to undergo that unconscious perception²⁰. On the internalist version of the inequivalence view, conscious perceptual experience results from an interaction between two mechanisms: one generates the phenomenal character, the other renders that phenomenal character conscious. Unconscious perception occurs when the second mechanism malfunctions (Polák and Marvan 2019, 2).

Inequivalence internalism is inspired by Rosenthal's insistence that the existence of mental qualities does not depend on their being conscious (Rosenthal 2010). According to Rosenthal's Quality Space Theory, mental qualities are independent of consciousness because we can identify them without reference to conscious experience of them. The idea is that the relations of difference and similarity between mental qualities mirror the relations of difference and similarity between perceptible physical properties of the perceived scene. Since this homomorphism is independent of consciousness, it predicts that seeing colours unconsciously is possible²¹.

Contrary to Rosenthal (Rosenthal 2019), however, Marvan and Polák believe that 'the qualitative nature of a mental state can be directly ascertained only from its conscious appearance to a subject' (Marvan and Polák 2017, 3). Hence they argue that inequivalence internalism mitigates the hard problem of consciousness, i.e. the difficulty of explaining 'why and how do physical processes in the brain give rise to conscious experience' (Chalmers 2018, 6):

'If we allow for the possibility of non-conscious states having phenomenal character, consciousness itself might ultimately transpire to be identifiable with a relatively simple neural mechanism.' (Polák and Marvan 2019, 4).

One may object that this merely shifts the crux of the matter: the hard problem of consciousness gets replaced with the hard problem of phenomenal qualities. But Polák and Marvan (Polák and Marvan 2019, 5) disagree. They contend that dividing the hard problem into two separate problems (i.e. the problem of consciousness and the problem of phenomenal qualities) and showing that the former is in principle solvable does move the research forward.

This response is only partially convincing. Granted, inequivalence internalism is better suited to accommodate UPC than the equivalence accounts, and in this sense makes a move in the right direction. Nevertheless, the charge of shifting the crux of the problem is on point, as Polák and Marvan do not explain unconscious production of phenomenal qualities in visual cortex. While Polák and Marvan are not primitivists about phenomenal qualities such as colours (Polák and Marvan 2018), their conception of phenomenal character could be given a primitivist interpretation. For example, their preferred physicalist explanation of the production of phenomenal

²⁰ Coleman (Coleman 2020, 69–71) defends the same view with regard to phenomenal character of pain and suffering.

²¹ According to Rosenthal (Rosenthal 1999) the term 'colour' is ambiguous, as it can refer either to perceptible physical properties of the environment, or to mental qualities standing in a homomorphism relation to the former. Still, Rosenthal's view is compatible with UPC, since both sides of that homomorphism are independent of consciousness.

qualities in visual cortex could be replaced with an account employing the notion of strong emergence (Chalmers 2006). Nevertheless, the missing explanation is not a detailed story about the anatomy and mechanism of the visual system, but an account of how this specific anatomy and mechanism give rise to that specific phenomenal character. Indeed, the inequivalence internalist agrees with the equivalence internalist that the phenomenal character is produced in the subject when the visual system is given a suitable stimulation. All that they disagree about is where in the brain the production happens, and whether it depends on consciousness or not.

5.4 Inequivalence externalism

According to inequivalence externalism, (i) the phenomenal character of perception is consciousness-independent, and (ii) the properties of which the phenomenal character is comprised (*inter alia* colours) are not produced in the mind; they are mind-independent features of the perceived scene.

A straightforward example of inequivalence externalism is an unorthodox version of relationalism that allows for unconscious phenomenal character while viewing colours as primitive mind-independent properties (it thereby thoroughly accords with the conclusions of Sect. 4). In contrast to orthodox relationalism described in Sect. 5.2, which confines the scope of the relationalist analysis to conscious perception, unorthodox relationalism applies that analysis to all cases of perception (cf. Anaya and Clarke 2017; Zięba 2019).

One may object that the relationalist analysis cannot be consistently applied to unconscious perception. According to relationalist orthodoxy, perceptual experience is a ‘modification of consciousness’ determined by *conscious acquaintance* with a mind-independent object (Brewer 2011, 92). But how could the relation of perceptual acquaintance be sometimes conscious and sometimes not? This seems incoherent, since unconscious perception ‘would seem to involve being acquainted with some element, and yet that element making no contribution to the subject’s conscious perspective on the world’ (Phillips 2018b, 472). If colours constitute conscious visual experience, it cannot be possible to see them unconsciously.

However, this objection would work only if the perceived object were the sole constituent of conscious visual experience. To my knowledge, no relationalist endorses that claim (cf. French 2018). Perceptual relation does not depend solely on (i) the perceived object. It also depends on (ii) the state of the subject and (iii) the circumstances of perception²². And when the conditions (ii-iii) are suboptimal, it is only to be expected that the perceived object’s constitutive contribution to perceptual relation is going to be suboptimal as well. Hence unconscious perception can be considered as either defective or not fully developed instance of perceptual relation. Insofar as it exhibits features characteristic of ordinary conscious seeing despite not being conscious, it deserves the label *unconscious acquaintance*. So there is no inconsistency between UP and the relationalist account of visual phenomenology.

²² This is not to say that all these things constitute perceptual phenomenal character in the same way (see Zięba 2021).

Inequivalence externalism is the main competitor of inequivalence internalism as far as accommodating UPC is concerned. Both accounts gain support from the evidence for UP (and UPC). What differentiates them is what they say about the origin of phenomenal qualities. While the inequivalence internalist holds that phenomenal qualities are produced in visual cortex, the inequivalence externalist locates them out there in the world, waiting, as it were, to be perceived. In order to see which of the two approaches is better suited to accommodate UPC, it is necessary to inspect the sources of their disagreement.

Recall that the internalism vs. externalism distinction in Table 1 maps onto representationalism vs. relationalism distinction in metaphysics of perception. While representationalism is typically combined with colour reductionism²³, relationalism is usually coupled with the objectivist form of colour primitivism. For the reductionist, the hard problem of consciousness is genuine at least in this sense: there really is a need for an explanation of how the phenomenal reduces to the physical. By contrast, the objectivist-primitivist sees the need for such an explanation as an outcome of a misguided research programme, and regards the hard problem as a pseudo-problem. As Campbell puts it,

‘The seventeenth-century science had two great programmes. One was the ambition of explaining all of the physical reality in terms of mathematical laws governing the behaviour of fundamental particles and forces. [...] The second great programme was to analyse the qualitative world in terms of its relation to consciousness. [...] But while the first programme has been a great success, the second programme has simply been a disaster. It really hasn’t worked at all. [...] This whole programme of explaining the qualitative world in terms of its relations to consciousness is simply hopeless and we really should abandon that.’ (Campbell 2012).

In a similar vein, Kalderon (Kalderon 2007, 594–98) points out that it was the puzzle of how to incorporate qualitative properties (*inter alia* colours) into the quantitative description of the world that has led philosophers to think that qualitative properties are mind-dependent. This resulted in well-known and possibly unresolvable difficulties, namely the body-mind problem and the hard problem of consciousness. But if colours are not properties of the mental, locating them ‘in the head’ amounts to committing the ‘introjective error’ (see also Allen 2016, 176–83).

To some extent, UPC reinforces this diagnosis²⁴. While inequivalence internalism fares better than equivalence internalism when it comes to accounting for unconscious perception, equivalence internalism fares better than inequivalence internalism at justifying the mind-dependence of phenomenal qualities. If the latter are consciousness-dependent, as equivalence internalism has it, there is a good reason to think that they are mind-dependent. But if phenomenal qualities are consciousness-independent, as inequivalence internalism requires, an important reason to internal-

²³ Polák’s and Marvan’s proposal exemplifies that (see Polák and Marvan 2018).

²⁴ This is not to say that the diagnosis presupposes UPC. One can consistently approve the diagnosis and endorse equivalence externalism.

ize those qualities is lost. Even though consciousness-independence of phenomenal qualities does not entail that they are mind-independent, it corroborates the hypothesis that attempts to explain the qualitative in terms of its relation to consciousness rest on a mistake.

This might persuade the orthodox relationalist (i.e. the equivalence externalist) that the unorthodox version of their view (i.e. inequivalence externalism) is superior. But the inequivalence internalist will not be persuaded. They will insist that consciousness-independence of phenomenal qualities does not entail that they are mind-independent. To persuade them, more needs to be said to justify the externalization of phenomenal qualities. And since the inequivalence internalist is a representationalist about perception, presumably they will not be persuaded by any other standard motivation of relationalism. However, as I show in the next section, there is a reason to prefer inequivalence externalism over inequivalence internalism that does not figure among relationalism's standard motivations²⁵.

6 A dilemma for the inequivalence internalist

Inequivalence externalism is preferable over inequivalence internalism because it avoids a dilemma faced by inequivalence internalism. The dilemma ensues from the observation that the plausibility of UP (and UPC) is inversely proportional to the plausibility of the phenomenal overflow hypothesis.

Section 6.1 explains overflow and its motivations. Section 6.2 considers reasons against overflow, which in turn reveals a tension between overflow and UP. Section 6.3 sets forth a dilemma based on that tension. Section 6.4 addresses some objections against the reasoning that underlies the dilemma. Section 6.5 is a brief summary.

6.1 Phenomenal overflow

According to phenomenal overflow, one can be phenomenally conscious of something of which one is not access-conscious. Put differently, one can be aware of more things than one is able to report. This is usually motivated by suggesting that the phenomenal character of perceptual experience is so rich that it exceeds the capacity of the perceiver's working memory. It is hypothesized that phenomenal richness of perceptual experience is captured by high-capacity yet short-lasting iconic memory. Only a part of the content of the iconic memory can be transferred into the more durable working memory and become access-conscious. In this sense, phenomenal consciousness 'overflows' access consciousness.

Evidence for overflow reaches back to Sperling's study (Sperling 1960), in which subjects were presented with an array of letters presented in a grid for 5–500 milliseconds. The subjects were able to report 4–5 letters on average. However, when the subjects were cued to a specific row of letters shortly after the grid disappeared from their view, they were able to report almost all of the letters in that row. Moreover, the subjects were expressing a conviction that they saw all of the letters in the

²⁵ My reason is not the only one. There are other reasons of this sort (see e.g. Mandik 2017).

grid. According to the overflow interpretation of this result, the subjects were phenomenally conscious of every letter in the grid, but the limited capacity of working memory prevented them from being able to report them all. They could report only so much that could squeeze in their working memory, i.e. what they were access-conscious of.

Recent evidence for overflow consists of studies which develop Sperling's paradigm in various ways. In a study by Bronfman and colleagues (Bronfman et al. 2014), for example, the subjects were shown a grid of letters as in the Sperling's paradigm, except that the rows varied with respect to colour diversity of the letters. Apart from remembering the letters, the subjects were also asked to estimate the colour diversity in either cued or uncued rows. It was found that the subjects were able to assess the colour diversity in uncued rows without decrease in letter-recall. The authors claim that the additional task did not overload working memory because the information about the colour diversity was stored there as an indeterminate 'ensemble representation'. They also maintain that such compressed representation could not have been formed had the colours of individual letters not been consciously perceived. Hence the conclusion that the phenomenal overflows the accessible.

Another widely discussed study was conducted by Vandembroucke and colleagues (Vandembroucke et al. 2012). The subjects were presented with a display containing eight figures for 500 milliseconds. After a short break, the display was presented again, and the task was to indicate whether any of the figures had changed. The subjects were cued as to where on the display the potential change might occur (it occurred in 50% of the trials). In the Fragile Memory condition (FM), the cue was presented before the presentation of the second display; in the Working Memory condition (WM), the cue occurred for some time during the presentation of the second display. The subjects performed significantly better in the FM condition. The researchers attribute this to 'fragile memory', a type of memory that is less capacious than iconic memory, yet roughly twice as capacious as working memory. The idea is that the fragile memory caused the boost in performance in FM but was overwritten immediately after the second display was presented.

The outcome of this experiment supports overflow if the content of the fragile memory was phenomenally conscious. The authors argue that this was indeed the case by indicating that performance in FM was particularly high when the figures on the displays formed Kanizsa triangles. Assuming that (i) conscious experience is requisite for the occurrence of the Kanizsa modal completion²⁶, and (ii) the boost of performance in FM resulted precisely from that occurrence, it follows that (iii) the fragile memory is phenomenally conscious.

6.2 The tension between phenomenal overflow and UP

At first glance, findings of this sort constitute a strong case for overflow. But the fact of the matter is that they are not univocal. The evidence adduced in favour of overflow admits of being explained without postulating a dissociation between phenomenal consciousness and access consciousness. An alternative explanation regards

²⁶ This is controversial (see Harris et al. 2011; Fahrenfort et al. 2017; Jimenez, Montoro, and Luna 2017).

access consciousness as a graded rather than all-or-none phenomenon. On this view, it is possible that the subjects in the Sperling paradigm can report the letter A while not being able to report the letter B because their access to A is fully developed, whereas their access to B is merely rudimentary (M. Overgaard 2018; see also Caruthers 2017). What about the impression the subjects have about consciously seeing more than they are able to report? Evidence for that impression is questionable (Cova, Gaillard, and Kammerer 2021). Even if it exists, it might be an illusion resulting from the fact that one's working memory is constantly updated by shifts of attention (see e.g. Kouider et al. 2010; Schlicht 2012).

Part of the reason why the status of overflow is hard to settle is that the measurement of phenomenal consciousness is often (perhaps always) mediated by the measurement of access consciousness. This is because the so-called subjective measures of consciousness (confidence ratings, post-decision wagering, direct reports of visibility) effectively measure performance in perceptual discrimination tasks, not visual phenomenology (Irvine 2012; M. Overgaard 2018; Persuh 2018).

The same problem occurs in neuroscience, as the overflow controversy is closely related to the debate about the neural basis of consciousness. If overflow is true, the neural basis of phenomenal consciousness is independent of the neural basis of access consciousness. In the case of visual consciousness, this means that the activity in visual cortex (occipital region) should suffice for the occurrence of visual phenomenology. If overflow is false, the activity in the neural basis of access consciousness (frontal region) is necessary for visual phenomenology to occur. Because the neural basis of access consciousness produces reports, a boost of activity in that region is observed whenever the subjects are asked to report what they experience. This makes it difficult to determine which of the observed brain activity is responsible for experience, and which stands for the reports. Attempts have been made to circumvent this by employing the so-called no-report paradigms, but it is highly controversial whether this manoeuvre solves the problem (Block 2019; M. Overgaard 2018; Phillips and Morales 2020).

According to Phillips, closer inspection of empirical studies purporting to establish overflow reveals a 'yawning gap between an informational story offered in explanation of certain task-performance data, and a corresponding phenomenological story' (Phillips 2018a, 3). It is this explanatory gap that makes room for two conflicting yet just about equally plausible interpretations of the same evidence. In effect, it seems that every single case for overflow can be rebutted in two simple steps:

'(i) accept (for argument's sake) whatever interpretation is offered of the relevant data construed in purely representational or informational terms; (ii) dispute the 'bridging assumptions' used to move from this representational account to claims concerning consciousness.' (Phillips 2018a, 2).

Implementing this recipe to the (Bronfman et al. 2014) study, the opponent of overflow can (i) agree that the observed performance is enabled by rich representations of individual colours, but (ii) deny that these representations are phenomenally conscious. As to (Vandenbroucke et al. 2012) study, the anti-overflow theorist can (i) admit that the boost in performance was caused by Kanizsa modal completion, but (ii) deny that the latter requires phenomenal consciousness.

Phillips' diagnosis illustrates how the plausibility of overflow is inversely proportional to the plausibility of UP. The illustration depicts two possible cases. In the first case, overflow is true. It follows that unreportable (i.e. access-unconscious) phenomenal experience is possible. This fuels O2, i.e. the objection to UP according to which subjects in putative cases of unconscious perception are in fact transiently and/or residually conscious of the stimuli. Therefore, overflow casts serious doubt on UP, even though the two hypotheses are not antithetical. In the second case, overflow is false. It follows that the performance observed in studies such as (Bronfman et al. 2014) and (Vandenbroucke et al. 2012) is due to unconscious perception, which of course validates UP and UPC²⁷.

6.3 The dilemma

This is where the foreshown dilemma for inequivalence internalism comes up. Note that if overflow is true, the neural basis of phenomenal consciousness is distinct from the neural basis of access consciousness. In the case of visual perception, overflow entails that the phenomenal character of visual perception is produced in visual cortex (Block 2005; 2007; 2011a). The crux of the matter is that overflow shares this assumption with inequivalence internalism. Indeed, the evidence for overflow *just is* the evidence for the claim that the phenomenal character of visual perception is produced in visual cortex (cf. Schlicht 2012, 325–28). If that claim is true, overflow is validated, and the inequivalence internalist account of UPC is hard-pressed by O2. This is the first horn of the dilemma. But if phenomenal qualities are not produced in visual cortex, inequivalence internalism is false. This is the second horn, which takes form of the following argument:

- (1) the phenomenal overflow hypothesis is false (assumption);
- (2) there is no phenomenal consciousness (PC) without access consciousness (AC) (from 1);
- (3) the neural basis of PC is the same as the neural basis of AC (from 2)²⁸;
- (4) the neural basis of AC is not located exclusively in visual cortex (an empirical fact);

²⁷ Phillips (Phillips 2016, 432; 2018b, 486–87) mentions this tension between overflow and UP as a source of 'serious internal challenge' in Block's position concerning UP: by accepting overflow, Block renders his case for UP vulnerable to O2. D'Aloisio-Montilla (D'Aloisio-Montilla 2019) has argued recently that Phillips faces an analogous challenge: by taking issue with overflow, Phillips compromises his own employment of O2 against UP. But the two cases are not symmetrical. Block faces the challenge because he explicitly endorses both overflow and UP. Phillips' case is different. As I read him, his view is not so much that overflow and UP are false. Rather, he believes that currently available evidence cannot settle whether UP is true or not, nor whether overflow is true or not. So he protests that it is precipitant to argue that overflow or UP (or both) have been already established on empirical grounds. If anything, the tension between overflow and UP actually reinforces his sceptical position.

²⁸ François Kammerer has disputed the move from 1&2 to 3 by suggesting that the falsity of overflow is compatible with the claim that P-consciousness and A-consciousness have distinct realizers (which would make perfect sense if one thinks, say, that P-conscious properties are realized by some intrinsic property and A-conscious properties by some functional property). Still, if P-consciousness and A-consciousness have distinct realizers, the former can occur without the latter, and granting this much suffices to render overflow possible.

(5) the neural basis of visual PC contains (or is identical to) the neural basis of the phenomenal character of visual perception (from PC's definition, see Sect. 5.1);

(6) the phenomenal character of visual perception cannot be produced exclusively in visual cortex (from 3, 4 and 5).

The dilemma demonstrates that the falsity of overflow is just as problematic for the inequivalence internalist as the truth of it. But the problems of inequivalence internalism do not end here. When denial of overflow is combined with consciousness-independence of phenomenal qualities (i.e. UPC), we run out of candidate brain areas within which the production of phenomenal qualities could happen. It turns out that phenomenal qualities are not produced in the subject at all:

(7) the phenomenal character of perception is consciousness-independent (from UPC);

(8) the activity in the neural basis of AC does not suffice for the phenomenal character to occur (from 7);

(9) phenomenal character of colour perception is not produced in the subject (from 6 and 8).

Before I consider some objections against this reasoning, I want to clarify what the dilemma is, and what follows from it as far as the search for the most plausible reading of UPC is concerned.

On the first horn, the phenomenal character is produced in visual cortex. This validates overflow, i.e. the possibility of phenomenally conscious access unconscious perception, which in turn casts doubt on the possibility of unconscious perception (because it reinforces O2, i.e. the objection that the putative instances of unconscious perception are in fact cases of overflow).

A colour internalist who thinks that unconscious perception is impossible won't have any problem with that, but I am not arguing against colour internalism as such. Instead, my target is the ability of colour internalism to provide a plausible account of UPC. Put differently, the first horn is problematic for the colour internalist who believes that UPC is true. My target is Marvan and Polak's view and any other theory that postulates unconscious production of colours qua phenomenal qualities in visual cortex. To accommodate UPC, the proponents of such theories have to say that the neural basis of the phenomenal character is distinct from the neural basis of consciousness. Hence the prefrontal cortex is not an option for them. And any other brain area or brain activity they might indicate as responsible for phenomenal character will be intercepted by the overflow theorist, who will say that precisely that activity produces phenomenal consciousness.

On the second horn (the one supported by the argument above), the phenomenal character is not produced in visual cortex. In this case, the kind of colour internalism that postulates unconscious production of phenomenal character in visual cortex is false. If the phenomenal character can be unconscious, it cannot be produced in the neural basis of consciousness either (so the prefrontal cortex is not an option). And, as before, indicating any other part of the brain/brain activity as the producer of unconscious phenomenal character will prompt the overflow theorist to argue that this is where phenomenal consciousness is produced, which will undermine UPC.

To sum up, my point is that the tension between overflow and unconscious perception seems to hinder any internalist account of UPC. Admittedly, the dilemma does

not show that colour internalism simply cannot allow for UPC to be true. But it does show that the externalist account of seeing colours unconsciously is in a dialectically better position, for the simple reason that the externalist does not face this dilemma. Inequivalence externalism is more plausible than inequivalence internalism because it is more resilient to objections.

6.4 Responses to objections

One might object that, if Premise 7 is true, Premise 6 fails to follow from Premises 3–5. That is to say, if phenomenal qualities such as colours are consciousness-independent, the neural basis of phenomenal qualities is independent of the neural basis of consciousness; so even if overflow is false and phenomenal consciousness is not realized in visual cortex, it does not follow that phenomenal qualities are not realized there. However, recall that the evidence for overflow *just is* the evidence for the claim that phenomenal qualities are produced in visual cortex. Scepticism about overflow is at least partially motivated by scepticism regarding the claim that phenomenal qualities are produced in visual cortex (Schlicht 2012). Juraj Hvorecký has pointed out to me that this does not yet render overflow and unconscious perception indistinguishable. Yes, but insofar as both phenomena involve an unreportable phenomenal character, there is a legitimate worry that the differences between them might be merely superficial. Currently available evidence does not rule it out.

Another possible objection disputes the move from Premise 7 to Premise 8. If overflow is false, the neural basis of visual consciousness encompasses both occipital and frontal regions in the brain. Is it not clear that a combined activity in both these areas would suffice to produce phenomenal qualities? No, because their consciousness-independence renders that activity unnecessary for their coming into existence. According to the interpretation of UPC recommended in Sect. 4, colours *qua* phenomenal qualities are consciousness-independent. This entails that the neural realizer of phenomenal character differs from the neural realizer of consciousness in some way or another. For if they had the same realizer, they could not dissociate. Hence the claim that the neural basis of consciousness produces colours contradicts the inequivalence view. According to the latter, the neural realizer of consciousness does not realize phenomenal qualities; it only realizes conscious experience of them. Phenomenal qualities are phenomenal in the sense that they determine what it is like to be conscious of them when they are consciously perceived.

François Kammerer has objected that A's independence of B (i.e. the fact that A can exist without B) does not entail that the existence of B is insufficient for the existence of A, as there are possible entities A and B such that A is independent of B and yet the existence of B suffices for the existence of A. Still, if the phenomenal character is consciousness-independent, what reason is there to think that the realizer of A-consciousness realizes the phenomenal character? While it is incumbent on the objector to give such a reason, the tension between UP and overflow suggests that no such reason is there to be given. For example, suppose that some partial activation of the realizer of A-consciousness suffices for phenomenal character but not for A-consciousness, since the latter requires full activation. Why is that partial activation insufficient for P-consciousness without A-consciousness?

Finally, it might be objected that the dilemma ensues from an over-simplistic view about how the brain works. If we acknowledge the role of other brain areas, zoom in on the details of the two already mentioned, and factor in the complexity of connections between them, perhaps we will eventually trace the neural factory of phenomenal qualities, or so the objector may hope. Nevertheless, complicating the picture does not seem to solve the dilemma. Suppose that the inequivalence internalist submits evidence that the production of colours *qua* phenomenal qualities necessarily involves some brain activity B, located outside of both frontal and visual areas. The same evidence will doubtlessly prompt an overflow theorist to recognise B as responsible for access-unconscious phenomenal experience. If the overflow theorist is right, O2 compromises UPC. And whatever reason is given to show that they are wrong is going to controvert the initial suggestion that B is where the production happens. This illustrates that the dilemma reoccurs no matter how detailed a story is told about the brain's anatomy and functioning.

6.5 Taking stock

If UPC is combined with the falsity of overflow, the argument in Sect. 6.3 favours inequivalence externalism over inequivalence internalism. If overflow is true, UPC (and UP in general) is in jeopardy. While it is possible that overflow and UPC are both true (in which case it is less clear which stance is preferable), the tension between overflow and UP makes this possibility very difficult to motivate. Hence, by elimination, inequivalence externalism is the best bet for the UPC-theorist.

This does not yet show that inequivalence externalism is the best conception of visual phenomenology. That depends on the status of UPC, UP in general, overflow, and many other things. For all I have said, each of these hypotheses might be true. But as long as the status of overflow is left open, the UPC-theorist has a good reason to prefer inequivalence externalism over inequivalence internalism. Out of four conceptions of the phenomenal character of visual experience, inequivalence externalism renders UPC the most plausible. Insofar as unorthodox relationalism introduced in Sect. 5.4 is a viable option, the commitment to UPC constitutes a reason (not decisive, but substantial) to embrace the relational theory of perception²⁹ because that theory buys one the most plausible reading of UPC.

²⁹ Tomasz Placek has drawn my attention to the possibility of the following objection. If phenomenal consciousness requires access consciousness (i.e. overflow is false), then all conscious seeing is conceptually structured epistemic seeing (i.e. seeing that something is the case), which means that no conscious seeing is conceptually unstructured non-epistemic seeing (seeing things) (cf. Dretske 1969). It follows that conscious perception has representational content, which is inconsistent with pure forms of relationalism. If so, falsity of overflow does not render inequivalence externalism preferable over inequivalence internalism. For it is problematic for inequivalence externalism too, albeit for different reasons. Note however that the objection assumes that 'access-conscious' means 'accessed', not 'accessible'. In this connection, the relationalist has two responses to choose from. First, they can accept the objector's assumption, and argue that perception has content insofar as it is conscious (i.e. perceptual experience has content), while maintaining that perception as such is fundamentally relational no matter whether it is conscious or not (cf. Hellie 2014). Second, they might reject the assumption by stipulating that access-conscious perception is accessible perception, not accessed perception (i.e. available for conceptualisation, not conceptually structured). The second view leaves open the possibility of non-epistemic conscious seeing.

7 Conclusions

The take-home message of this paper is twofold. First, since none of the three main objections against it is conclusive, UPC is a plausible hypothesis, and its ramifications cannot be ignored. Second, the most plausible reading of UPC construes the phenomenal character of visual perception as (i) constituted by primitive mind-independent qualities and (ii) not essentially tied to consciousness.

Acknowledgements Special thanks are due to Craig French and Mark Calderon, as this paper was inspired by their referee reports on my doctoral thesis. Thanks also to Michel Ghins, Juraj Hvorecký, François Kammerer, Michał Klincewicz, Tomáš Marvan, Takuya Niikawa, Tomasz Placek, Błażej Skrzypulec, and three anonymous referees for their thought-provoking comments on earlier versions of the paper.

Funding This work was supported by the National Science Center (Poland) under grant number 2019/32/C/HS1/00113. Open Access to this article was funded by the "Inicjatywa Doskonałości" scheme at Jagiellonian University's Philosophy Department.

Declarations

Conflicts of interest/Competing interests The author declares that he has no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allen, K. (2016). *A Naïve Realist Theory of Colour*. Oxford University Press
- Anaya, A., & Clarke, S. (2017). 'Naïve Realism and Unconscious Perception: A Reply to Berger and Nanay'. *Analysis*, 77(2), 267–273
- Austin, C. J. (2019). *Essence in the Age of Evolution. A New Theory of Natural Kinds*. Routledge
- Berger, J., & Mylopoulos, M. (2019). 'On Scepticism about Unconscious Perception'. *Journal of Consciousness Studies*, 26(11–12), 8–32
- Bermúdez, J. L. (2000). 'Personal and Sub-personal; A Difference without a Distinction'. *Philosophical Explorations*, 3(1), 63–82
- Bird, A. (2018). 'The Metaphysics of Natural Kinds'. *Synthese*, 195, 1397–1426
- Block, N. (1995). 'On a Confusion about a Function of Consciousness'. *Behavioral and Brain Sciences*, 18(2), 227–287
- Block, N. (2005). 'Two Neural Correlates of Consciousness'. *Trends in Cognitive Sciences* 9 (2): 46–52
- Block, N. (2007). 'Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience'. *Behavioral and Brain Sciences* 30: 481–548
- Block, N. (2011a). 'Perceptual Consciousness Overflows Cognitive Access'. *Trends in Cognitive Sciences* 15 (12): 567–75
- Block, N. (2011b). 'The Higher Order Approach to Consciousness Is Defunct'. *Analysis* 71 (3): 419–31

- Block, N. (2016). 'The Anna Karenina Principle and Skepticism about Unconscious Perception'. *Philosophy and Phenomenological Research* 93 (2): 452–59
- Block, N. (2019). 'What Is Wrong with the No-Report Paradigm and How to Fix It'. *Trends in Cognitive Sciences* 23 (12): 1003–13
- Block, N., & Phillips, I. (2017). 'Debate on Unconscious Perception'. In *Current Controversies in Philosophy of Perception*, edited by B. Nanay, 165–92. New York, London: Routledge
- Boyer, J. L., Harrison, S., & Ro, T. (2005). 'Unconscious Processing of Orientation and Color without Primary Visual Cortex'. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16875–16879
- Breitmeyer, B. G., Ro, T., & Singhal, N. S. (2004). 'Unconscious Color Priming Occurs at Stimulus- Not Percept-Dependent Levels of Processing'. *Psychological Science*, 15(3), 198–202
- Brewer, B. (2011). *Perception and Its Objects*. Oxford, New York: Oxford University Press
- Brogaard, B., & Gatzia, D. E. (2017). 'Is Color Experience Cognitively Penetrable?'. *Topics in Cognitive Science*, 9(1), 193–214
- Bronfman, Z. Z., Brezis, N., Jacobson, H., & Usher, M. (2014). 'We See More Than We Can Report: "Cost Free" Color Phenomenality Outside Focal Attention'. *Psychological Science*, 25(7), 1394–1403
- Burge, T. (2005). 'Disjunctivism and Perceptual Psychology'. *Philosophical Topics*, 33(1), 1–78
- Burge, T. (2010). *The Origins of Objectivity*. Oxford University Press
- Byrne, A., & Hilbert, D. R. (2003). 'Color Realism and Color Science'. *Behavioral and Brain Sciences*, 26(1), 3–21
- Byrne, A., & Hilbert, D. R. (2007). 'Color Primitivism'. *Erkenntnis* 66 (1/2): 73–105
- Campbell, J. (1993). 'A Simple View of Colour'. In *Reality, Representation, and Projection*, edited by J. Haldane and C. Wright, 257–68. New York, Oxford: Oxford University Press
- Campbell, J. (2012). *What Does Visual Experience Have To Do With Vision Science?* <https://youtu.be/mm6vAMpKld4>
- Carruthers, P. (2017). 'Block's Overflow Argument'. *Pacific Philosophical Quarterly*, 98(S1), 65–70
- Chalmers, D. (2006). 'Strong and Weak Emergence'. In *The Re-Emergence of Emergence*, edited by P. Clayton and P. Davis, 244–54. Oxford University Press
- Chalmers, D. (2018). 'The Meta-Problem of Consciousness'. *Journal of Consciousness Studies* 25 (9–10): 6–61
- Clark, A. (2007). 'What Reaching Teaches: Consciousness, Control, and the Inner Zombie'. *British Journal for the Philosophy of Science*, 58, 563–594
- Coleman, S. (2020). 'Painfulness, Suffering, and Consciousness'. In *Philosophy of Suffering. Metaphysics, Value, and Normativity*, edited by D. Bain, M. Brady, and J. Corns, 55–74. Routledge
- Cova, F., Gaillard, M., & Kammerer, F. (2021). 'Is the Phenomenological Overflow Argument Really Supported by Subjective Reports?'. *Mind & Language*, 36(3), 422–450
- D'Aloisio-Montilla, N. (2019). 'Phillips on Unconscious Perception and Overflow'. *Philosophia*, 47(3), 649–662
- Danckert, J., & Rossetti, Y. (2005). 'Blindsight in Action: What Can the Different Sub-Types of Blindsight Tell Us about the Control of Visually Guided Actions?'. *Neuroscience and Biobehavioral Reviews*, 29, 1035–1046
- Devitt, M. (2021). 'Defending Intrinsic Biological Essentialism'. *Philosophy of Science*, 88, 67–82
- Dijksterhuis, A. (2004). 'Think Different: The Merits of Unconscious Thought in Preference Development and Decision Making'. *Journal of Personality and Social Psychology*, 87(5), 586–598
- Dijksterhuis, A. (2006). 'A Theory of Unconscious Thought'. *Perspectives on Psychological Science* 1 (2): 95–109
- Dijksterhuis, A., & van Olden, Z. (2006). 'On the Benefits of Thinking Unconsciously: Unconscious Thought Can Increase Post-Choice Satisfaction'. *Journal of Experimental Social Psychology*, 42, 627–631
- Drayson, Z. (2012). 'The Uses and Abuses of the Personal/Subpersonal Distinction'. *Philosophical Perspectives*, 26(1), 1–18
- Dretske, F. (1969). *Seeing and Knowing*. London: Routledge & Kegan Paul
- Dretske, F. (2004). 'Change Blindness'. *Philosophical Studies* 120 (1/3): 1–18
- Fahrenfort, J. J., van Leeuwen, J., Olivers, C. N. L., & Hogendoorn, H. (2017). 'Perceptual Integration without Conscious Access'. *Proceedings of the National Academy of Sciences of the United States of America*, 114(14), 3744–3749. <https://doi.org/10.1073/pnas.1617268114>

- Fahrenfort, J. J., Snijders, T., Heinen, K., van Gaal, S., Scholte, H., & Lamme, V. A. F. (2012). 'Neuronal Integration in Visual Cortex Elevates Face Category Tuning to Conscious Face Perception'. *Proceedings of the National Academy of Sciences* 109 (52): 21504–9
- Fogelson, S. V., Kohler, P. J., Miller, K. J., & Granger, R. (2014). 'Unconscious Neural Processing Differs with Method Used to Render Stimuli Invisible'. *Frontiers in Psychology* 5: 1–11
- French, C. (2018). 'Naïve Realism and Diaphancity'. *Proceedings of the Aristotelian Society* 118 (2): 149–75
- French, C., and I. Phillips. forthcoming. 'Naïve Realism, the Slightest Philosophy, and the Slightest Science'. In *Contemporary Debates in the Philosophy of Mind*, edited by B. McLaughlin and J. Cohen
- Friend, T. (2019). 'Can Parts Cause Their Wholes?' *Synthese* 196 (12): 5061–82
- Grimes, J. (1996). 'On the Failure to Detect Changes in Scenes across Saccades'. In *Perception*, edited by K. Akins, 5:89–110. Vancouver Studies in Cognitive Science. Oxford University Press
- Haggard, P., & Libet, B. (2001). 'Conscious Intention and Brain Activity'. *Journal of Consciousness Studies*, 8(11), 47–63
- Harris, J. J., Schwarzkopf, D., Song, C., Bahrami, B., & Rees, G. (2011). 'Contextual Illusions Reveal the Limit of Unconscious Visual Processing'. *Psychological Science*, 22(3), 399–405
- Hellie, B. (2014). 'Love in the Time of Cholera'. In *Does Perception Have Content?*, edited by B. Brogaard, 241–61. New York: Oxford University Press
- Irvine, E. (2012). 'Old Problems with New Measures in the Science of Consciousness'. *British Journal for the Philosophy of Science*, 63(3), 627–648
- Jackson, F. C. (1986). 'What Mary Didn't Know'. *The Journal of Philosophy*, 83(5), 291–295
- Jackson, F. C. (2019). 'How to Be an Objectivist about Colour'. *Phenomenology and the Cognitive Sciences* 18 (5): 819–31
- Jimenez, M., Montoro, P. R., & Luna, D. (2017). 'Global Shape Integration and Illusory Form Perception in the Absence of Awareness'. *Consciousness and Cognition*, 53, 31–46
- Kalderon, M. E. (2007). 'Color Pluralism'. *The Philosophical Review*, 116(4), 563–601
- Kalderon, M. E. (2011). 'The Multiply Qualitative'. *Mind* 120 (478): 239–62
- Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). 'How Rich Is Consciousness? The Partial Awareness Hypothesis'. *Trends in Cognitive Sciences*, 14(7), 301–307
- Lamme, V. A. F. (2015). 'The Crack of Dawn. Perceptual Functions and Neural Mechanisms That Mark the Transition from Unconscious Processing to Conscious Vision'. In *Open MIND*, edited by T. Metzinger and M. Windt, 22:1–34. Frankfurt am Main: MIND Group
- MacPherson, F. (2012). 'Cognitive Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism'. *Philosophy and Phenomenological Research*, 84(1), 24–62
- Mandik, P. (2017). 'The Myth of Color Sensations, or How Not to See a Yellow Banana'. *Topics in Cognitive Science*, 9(1), 228–240
- Martin, M. G. F. (2002). 'The Transparency of Experience'. *Mind and Language*, 4(4), 376–425
- Martin, M. G. F., Gendler, T., & Hawthorne, J. (Eds.). (2006). 354–410. Oxford University Press
- Marvan, T., & Polák, M. (2017). 'Unitary and Dual Models of Phenomenal Consciousness'. *Consciousness and Cognition*, 56, 1–12
- Mealor, A. D., & Dienes, Z. (2012). 'Conscious and Unconscious Thought in Artificial Grammar Learning'. *Consciousness and Cognition*, 21, 865–874
- Mizrahi, V. (2006). 'Color Objectivism and Color Pluralism'. *Dialectica*, 60(3), 283–306
- Moutoussis, K., & Zeki, S. (2002). 'The Relationship between Cortical Activation and Perception Investigated with Invisible Stimuli'. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14), 9527–9532
- Nanay, B. (2017). 'Philosophy of Perception: A Roadmap with Lots of Bypass Roads'. In *Current Controversies in Philosophy of Perception*, edited by B. Nanay, 1–19. New York, London: Routledge
- Nelkin, N. (1989). 'Unconscious Sensations'. *Philosophical Psychology*, 2(2), 129–141
- Nimt, C. (2021). 'How Science and Semantics Settle the Issue of Natural Kind Essentialism'. *Erkenntnis*, 86, 149–170
- Nisbett, R. E., & Wilson, T. D. (1977). 'Telling More than We Can Know: Verbal Reports on Mental Processes'. *Psychological Review*, 84(3), 231–259
- Norman, L. J., Akins, K., Heywood, A., & Kentridge, R. W. (2014). 'Color Constancy for an Unseen Surface'. *Current Biology*, 24(23), 2822–2826
- Overgaard, M. (2018). 'Phenomenal Consciousness and Cognitive Access'. *Philosophical Transactions of the Royal Society B*, 373. <https://doi.org/10.1098/rstb.2017.0353>

- Overgaard, M., & Grünbaum, T. (2011). 'Consciousness and Modality: On the Possible Preserved Visual Consciousness in Blindsight Subjects'. *Consciousness and Cognition*, 20, 1855–1859
- Overgaard, S. (2011). 'Disjunctivism and the Urgency of Scepticism'. *Philosophical Explorations*, 14(1), 5–21
- Pautz, A. (2009). 'Colour, Philosophical Perspectives'. In *Oxford Companion to Consciousness*, edited by T. Bayne, A. Cleeremans, and P. Wilken, 144–49. Oxford University Press
- Persuh, M. (2018). 'Measuring Perceptual Consciousness'. *Frontiers in Psychology*, 8, <https://doi.org/10.3389/fpsyg.2017.02320>
- Peters, M., Kentrige, R., Phillips, I., & Block, N. (2017). 'Does Unconscious Perception Really Exist? Continuing the ASSC20 Debate'. *Neuroscience of Consciousness*, 3(1), 1–11
- Peters, M., & Lau, H. (2015). 'Human Observers Have Optimal Introspective Access to Perceptual Processes Even for Visually Masked Stimuli'. *ELIFE*. <https://doi.org/10.7554/eLife.09651.001>
- Phillips, I. (2016). 'Consciousness and Criterion: On Block's Case for Unconscious Seeing'. *Philosophy and Phenomenological Research*, 93(2), 419–451
- Phillips, I. (2018a). 'The Methodological Puzzle of Phenomenal Consciousness'. *Philosophical Transactions of the Royal Society B* 373. <https://doi.org/10.1098/rstb.2017.0347>
- Phillips, I. (2018b). 'Unconscious Perception Reconsidered'. *Analytic Philosophy* 59 (4): 471–514
- Phillips, I., & Morales, J. (2020). 'The Fundamental Problem with No-Cognition Paradigms'. *Trends in Cognitive Sciences*, 24(3), P165–167
- Polák, M., & Marvan, T. (2018). 'Neural Correlates of Consciousness Meet the Theory of Identity'. *Frontiers in Psychology*, 9, <https://doi.org/10.3389/fpsyg.2018.01269>
- Polák, M., & Marvan, T. (2019). 'How to Mitigate the Hard Problem by Adopting the Dual Theory of Phenomenal Consciousness'. *Frontiers in Psychology* 10. <https://doi.org/10.3389/fpsyg.2019.02837>
- Prinz, J. (2012). *The Conscious Brain*. New York: Oxford University Press
- Pritchard, D. (2016). *Epistemic Angst. Radical Scepticism and the Groundlessness of Our Believing*. Princeton University Press
- Railo, H., Salminen-Vaparanta, N., Henriksson, L., Revonsuo, A., & Koivisto, M. (2012). 'Unconscious and Conscious Processing of Color Rely on Activity in Early Visual Cortex: A TMS Study'. *Journal of Cognitive Neuroscience*, 24(4), 819–829
- Ramsøy, T. Z., & Overgaard, M. (2004). 'Introspection and Subliminal Perception'. *Phenomenology and the Cognitive Sciences*, 3, 1–23
- Ro, T., Singhal, N. S., Breitmeyer, B. G., & Garcia, J. O. (2009). 'Unconscious Processing of Color and Form in Metacontrast Masking'. *Attention, Perception, & Psychophysics*, 71(1), 95–103
- Rosenthal, D. M. (1999). 'The Colors and Shapes of Visual Experiences'. In *Consciousness and Intentionality: Models and Modalities of Attribution*, edited by D. Fisette, 95–118. Springer Science + Business Media B.V
- Rosenthal, D. M. (2005a). 'Sensory Qualities, Consciousness, and Perception'. In *Consciousness and Mind*, 175–226. Oxford: Clarendon Press
- Rosenthal, D. M. (2005b). 'Sensory Quality and the Relocation Story'. In *Consciousness and Mind*, 149–74. Oxford: Clarendon Press
- Rosenthal, D. M. (2005c). 'The Independence of Consciousness and Sensory Quality'. In *Consciousness and Mind*, 135–48. Oxford: Clarendon Press
- Rosenthal, D. M. (2010). 'How to Think about Mental Qualities'. *Philosophical Issues* 20 (1): 368–93
- Rosenthal, D. M. (2011). 'Exaggerated Reports: Reply to Block'. *Analysis* 71 (3): 431–37
- Rosenthal, D. M., Coates, P., & Coleman, S. (Eds.). (2015). 33–65. Oxford University Press
- Rosenthal, D. M., Coleman, S. (Ed.). (2019). 32–61. Cambridge University Press
- Schlicht, T. (2012). 'Phenomenal Consciousness, Attention and Accessibility'. *Phenomenology and the Cognitive Sciences*, 11(3), 309–334
- Schurger, A., Pereira, F., Treisman, A., & Cohen, J. (2010). 'Reproducibility Distinguishes Conscious from Nonconscious Neural Representations'. *Science*, 327(97), 97–99
- Schurger, A., Sarigiannidis, I., Naccache, L., Sitt, J., & Dehaene, S. (2015). 'Cortical Activity Is More Stable When Sensory Stimuli Are Consciously Perceived'. *Proceedings of the National Academy of Sciences* 112 (16): E2083–92
- Schwitzgebel, E. (2008). 'The Unreliability of Naive Introspection'. *The Philosophical Review*, 117(2), 245–273
- Schwitzgebel, E., Smithies, D., & Stoljar, D. (Eds.). (2012). 29–47. Oxford University Press
- Shepherd, J., & Mylopoulos, M. (2021). 'Unconscious Perception and Central Coordinating Agency'. *Philosophical Studies*, 178(12), 3869–3893

- Skrzypulec, B. (2021). 'Contents of Unconscious Color Perception'. *Review of Philosophy and Psychology*
- Sperling, G. (1960). 'The Information Available in Brief Visual Presentations'. *Psychological Monographs: General and Applied*, 74(11), 1–29
- Stoneham, T. (2008). 'A Neglected Account of Perception'. *Dialectica*, 62(3), 307–322
- Suhler, C. L., & Churchland, P. S. (2009). 'Control: Conscious and Otherwise'. *Trends in Cognitive Sciences*, 13(8), 341–347
- Taylor, H. (2020). 'Fuzziness in the Mind: Can Perception Be Unconscious?'. *Philosophy and Phenomenological Research*, 101(2), 383–398. <https://doi.org/10.1111/phpr.12592>
- Vandenbroucke, A. R. E., Sligte, I. G., Fahrenfort, J. J., Ambroziak, K. B., & Lamme, V. A. F. (2012). 'Non-Attended Representations Are Perceptual Rather than Unconscious in Nature'. *PLoS ONE*, 7(11), e50042
- Weksler, A., Jacobson, H., & Bronfman, Z. Z. (2021). 'The Transparency of Experience and the Neuroscience of Attention'. *Synthese*, 198(5), 4709–4730
- Wong, H. Y. (2014). 'Personal and Sub-Personal: Overcoming Explanatory Apartheid'. In *Communicative Action. Selected Papers of the 2013 IEAS Conference on Language and Action*, edited by T.-W. Hung. Springer
- Wu, W. (2013). 'The Case for Zombie Agency'. *Mind*, 122(485), 217–230
- Yang, E., Brascamp, J., Kang, M. S., & Blake, R. (2014). 'On the Use of Continuous Flash Suppression for the Study of Visual Processing Outside of Awareness'. *Frontiers in Psychology*, 5(724), 1–17
- Zięba, P. J. (2019). 'Naïve Realism about Unconscious Perception'. *Synthese*, 196(5), 2045–2073. <https://doi.org/10.1007/s11229-017-1570-1>
- Zięba, P. J. (2021). 'Selectionism and Diaphaneity'. *Axiomathes*. <https://doi.org/10.1007/s10516-021-09608-w>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.