CONSCIOUS AND UNCONSCIOUS MENTALITY

Examining their Nature, Similarities and Differences

Edited by Juraj Hvorecký, Tomáš Marvan and Michal Polák

First published 2024

ISBN: 978-1-032-52979-0 (hbk)
ISBN: 978-1-032-52974-5 (pbk)
ISBN: 978-1-003-40952-6 (ebk)

9

THE BRAIN-BASED ARGUMENT FOR UNCONSCIOUS SENSORY QUALITIES

Tomáš Marvan

DOI: 10.4324/9781003409526-12

This Chapter was funded by EU grant No. CZ.02.2.69/0.0/0.0/18_054/0014626

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

9

THE BRAIN-BASED ARGUMENT FOR UNCONSCIOUS SENSORY QUALITIES

Tomáš Marvan

9.1 Introduction

When we consciously see, hear or smell things, we experience sensory qualities of colours, sounds, or smells. An influential tradition in the philosophy of mind, and in cognitive science more broadly, treats these qualities as invariably conscious. The sensory qualities are *consciousness-making* qualities as they are what constitute the difference between unconscious and conscious perception. Given the prevalence of this way of thinking about sensory qualities, the talk of unconscious sensory qualities is countenanced only on condition that a notion of sensory qualities different from that typically adopted in the philosophy of mind is applied.

This route towards unconscious qualities was famously taken by David Rosenthal. Rosenthal (2005) develops a theory of unconscious mental qualities which are different from how we consciously experience them. What we consciously experience is the *subjective appearance* of the qualities which constitutes the 'what-it's-like-ness' of the mental state in question; this subjective appearance is generated by suitable higher-order thoughts. However, the unconscious mental qualities are individuated independently of their subjective appearance, by their perceptual role alone. The perceptual role is fixed by quality spaces for individual sensory modalities. This theory is then consistent with the common intuition that unconscious qualities must be fundamentally different from conscious ones. Unconscious qualities drive behaviour and enable perceptual discrimination. Conscious qualities are characterized by their subjective appearance and 'what-it's-like-ness'. As a consequence, this way of construing the unconscious qualities becomes acceptable even to theorists convinced that qualities *in the common sense* cannot be conscious.

DOI: 10.4324/9781003409526-12

My thinking about sensory qualities was deeply influenced by Rosenthal in more than one respect. In particular, the two points of common emphasis of both accounts are that (1) mental qualities occur both consciously and unconsciously, and (2) the unconscious qualities are made conscious by an independent process or mechanism. However, in contrast to Rosenthal, I claim that the unconscious qualities can be literally the same as the conscious ones. The very same mental state, composed of the very same sensory qualities, can exist consciously as well as unconsciously, depending on the circumstances of perception. In defending this view, I do not wish to deny the importance of the perceptual role of sensory qualities; in fact, one of the arguments for unconscious qualities I offer below draws on the relevant similarity of perceptual roles performed by conscious and unconscious qualities. But I do wish to say that the way consciousness presents sensory qualities is crucial for establishing their true character.

By defending unconscious sensory qualities, I thus want to argue against the tradition which denies that qualities as we experience them when perceiving could exist non-consciously.[1] However, conceptually, my break with this tradition is not as radical as Rosenthal's. I acknowledge that when we speak about sensory qualities, what we typically have in mind are the conscious qualities, the qualities that we subjectively experience. A notion of 'qualities' that is completely severed from the way we normally conceive of qualities is problematic. It is not clear why we should call Rosenthal's relational and structural unconscious entities 'qualities' in the first place, as they have little in common with what we normally call qualities. Taking this conceptual intuition as my starting point, I will be arguing that, although unconscious sensory qualities lack subjective awareness, in their qualitativeness they do not differ from the conscious ones. In other words, although only conscious sensory states possess the experiential 'what-it's-like-ness' and unconscious sensory states lack it, experiential 'what-it's-like-ness' is simply *the sum* of the qualitative character of a sensory state and its being conscious. On this account, the difference between unconscious and conscious sensory qualities is lesser than assumed by most philosophers of mind, and cognitive scientists more broadly.[2]

In the following section I will sketch what I see as the central arguments for my view. The third section then qualifies the view, noting that although there is a part of the mind that produces both conscious and unconscious qualities, the unconscious part of the mind is ultimately broader and more fine-grained than the conscious one.

9.2 Arguments for unconscious sensory qualities

The dominant view of sensory qualities, mentioned in the Introduction, conjoins sensory qualities and consciousness. Therefore, I call this a 'one-factor

view'.[3] Supporters of the one-factor approach do not need to deny the existence of unconscious sensory *processing*. They just insist that this processing is devoid of any qualitativeness (in the ordinary sense). According to this view, as sensory processing unfolds, qualitativeness appears at the very same instant perception becomes conscious.

The view I want to defend here rejects this picture and accepts the idea that unconscious sensory episodes have a genuinely qualitative character. In fact, it proposes that the qualitative character of unconscious sensory episodes is the very same as that possessed by conscious sensory episodes. According to this 'two-factor' view, sensory qualities are fully constituted unconsciously, but need to interact with a separate, non-qualitative process or mechanism that makes them consciously available to the subject. The two-factor view allows a systematic explanation of how perceptual states become conscious, because it divorces the mechanisms of sensory qualities from the mechanisms of sensory consciousness. In the first stage of perception, the mental state is established unconsciously, in all its qualitative aspects. In the second stage of perception, these finalized qualitative states are made consciously available to the subject by a distinctly non-qualitative process, or set of processes; presumably, these processes are shared across different sensory modalities. We could call such additional process(es) 'consciousness-conferring' process(es) (Bachmann 2011; Marvan and Polák 2020).

The two-factor view is not just compatible with much empirical work in experimental psychology and the cognitive neuroscience of consciousness; it is directly supported by this work. In the remainder of this chapter, I will focus on the empirical support for unconscious sensory qualities, the central ingredient of the two-factor view of sensory consciousness. I will start with support from behavioural research. Then I will move to a brain-based argument for unconscious qualities. As this is the stronger of the two, and as other authors have already written on the behavioural argument at some length (Lockwood 1989, chapter 10; Rosenthal 2005; Young 2014; Young et al. 2014; Keller 2016; Coleman forthcoming), I will present the brain-based argument in more detail.

9.2.1 The behavioural argument

The behavioural argument for unconscious sensory qualities is simple. For visual domain, the argument says this:

(1) Absence of consciousness in perceptual states does not completely eliminate all behavioural capacities: The subject can still detect objects, discriminate between them, interact with them, and so on.
(2) The behaviours drawing on unconscious information are the same as those drawing on conscious information.

(3) The behaviours drawing on consciously available information utilize qualitative information (about, for example, the colour, shape, motion, or texture of objects).
(4) The behaviours drawing on unconscious information are the same because they utilize the same qualitative information present unconsciously.
(∴) Unconscious sensory qualities exist, and they have the same nature as the conscious ones.

Let me illustrate the argument with a non-visual example. As Young et al. (2014) and Young (2014) point out, in research on olfactory valence—indicating whether we find a smell pleasant or unpleasant—behaviour in both conscious and unconscious smelling manifests close similarities in 'secondary processing measures'. These measures include behavioural changes in sniff rate and volume: within the first 150 ms after the stimulus onset, sniff rate, and volume is adjusted according to its valence. For instance, when the unconsciously registered smell is evaluated as unpleasant, both sniff rate and volume drop. The same modulation in sniff rate and volume is also present when we smell things consciously. So it seems that the olfactory system treats the olfactory stimulus in the same way, *regardless of whether the subject is conscious of it or not*. According to the behavioural argument, the system treats both conscious and unconscious olfactory qualities in the same way because they share their qualitative character. We *literally* smell things unconsciously, in the qualitative sense of smelling.

The identity of behavioural responses in conscious and unconscious olfactory conditions would be difficult to account for if we presumed that the consciously experienced olfactory qualities were in their nature radically different from the unconscious olfactory processing. Here the standard, one-factor account runs into trouble. It portrays unconscious perception as devoid of qualities typical for conscious experience. But how does one explain the identity in reactions without assuming that the same qualities are present in both conscious and unconscious perceptual conditions? The one-factor account must postulate two very different sets of perceptual mechanisms, one for conscious and one for unconscious perceptual states. Two-factor theory needs only one set of perceptual mechanisms.

Unfortunately, the behavioural argument for unconscious sensory qualities has a serious limitation. Although unconscious qualities can guide our behaviour, it is likely that often they are not as fully and completely connected with the systems and circuits governing behaviour as the conscious qualities. Despite impressive examples of unconscious states driving behaviour, such as the olfactory states regulating the secondary processing measures, we shouldn't assume that unconscious qualities are always able to influence behaviour as well as the conscious ones. In fact, it may be one of the functions, or even the main function, of the conscious mode of perceiving that it fully establishes the link between sensory states and behaviour.[4]

The behavioural argument thus can be only of limited use to the advocates of unconscious qualities (and of the two-factor view). If the existence of unconscious qualities depended solely on the behavioural argument, it would be challenging to defend it convincingly. Fortunately, there is another type of argument for unconscious sensory qualities that is not affected by the limitation of the behavioural argument. It is the argument from shared neural resources of conscious and unconscious sensory states, to which I now turn.

9.2.2 The brain-based argument

The brain-based argument (BBA) for unconscious sensory qualities could be more appropriately called the argument from shared neural resources of conscious and unconscious sensory qualities. The argument's structure is similar to that of the behavioural argument spelled out in the previous subsection. The BBA focuses on what goes on in the brain when we are in perceptual states, and consists of the following steps (again, for visual modality).

(1) There is functional specialization in concrete brain areas for specific types of visual contents.
(2) A brain area specialized for processing a given type of visual contents is directly involved in expressing this type of contents in conscious experience.
(3) The same area is similarly activated when the same visual stimulus is present but the subject is not aware of it.
(4) Conscious visual contents have qualitative character.
(∴) Unconscious visual contents, being produced in the same way in the very same brain areas as the conscious ones, have the same qualitative character as the conscious visual contents.

I will go through premises 1 to 3, sketch the empirical and conceptual support for them, and respond to objections I can anticipate. (Premise 4 seems to me to be obvious, so I won't comment on it.)

On 1. "There is functional specialization in concrete brain areas for specific types of visual contents."

I assume that the first premise of the BBA is not terribly controversial. Although the matter is subject to some debate, thanks to decades of patient work in visual neuroscience we now have well-supported theories about where at least some types of visual contents get processed in the brain. To give a couple of examples that have been prominent in the literature in recent decades, we now have robust evidence of functional specificity of the fusiform face area (FFA) for face perception, of the parahippocampal place area (PPA) for the perception of places, of the extrastriate body area (EBA) for the perception of the human body and its parts, of the visual word form area (VWFA) for

the perception of letters from visual letter shapes, and of the visual motion areas MT and MST for perception of motion. These specialized areas respond more strongly to their preferred stimuli than to other types of stimuli. At the same time, no other areas respond so strongly to these stimuli. This is equally supported by measurements of brain activity in the specialized sites and by their targeted microstimulation (Grill-Spector et al. 2004; Kanwisher 2001; Boly et al. 2017; Downing et al. 2001; Cohen et al. 2000; Bisley et al. 2001; Celebrini and Newsome 1995; Rangarajan et al. 2014).

On 2. "A brain area specialized for processing a given type of visual contents is directly involved in expressing this type of contents in conscious experience."

In contrast to the first premise, the second premise is perhaps the most controversial of the BBA, so let me spend some time defending it from an objection I can anticipate.

Let us take the visual perception of faces as an example. The BBA claims that an area specialized in face processing, viz., the fusiform face area, is a place where the visual face-qualities are not just processed, but also expressed in consciousness. However, it could be claimed that currently available evidence is insufficient to prove this. It could be that the specialized processing of faces in the FFA is just a necessary *preparatory* phase of qualitative visual perception, and that the results of this processing phase need to be taken over by other parts of the brain which, alone, produce truly qualitative contents and express them in the subject's consciousness. This argument would parallel a line of reasoning with its origins at the very beginning of the modern search for the neural correlates of consciousness (NCCs): The Crick and Koch (1995a) argument that the early visual area V1 is not a part of the visual neural correlate of consciousness, despite being engaged in a substantial amount of visual processing. In Crick and Koch's interpretation of the evidence (that was available at the time of their writing), processing in V1 is just one of the necessary preparatory steps of conscious vision. Conscious contents are ultimately expressed outside of V1, in higher visual areas. To capture what V1 is doing for visual perception, Crick and Koch revert to a metaphorical distinction between computations in a computer, and the results of such computations expressed on a computer screen (Crick and Koch 1995b, 294; see also Crick and Koch 1992, 154).

A critic of the BBA could use the same strategy to find fault with the idea that visual face-qualities are constituted and expressed in the fusiform face area. The objection would be that not just V1 but also the higher visual areas in the inferotemporal cortex such as the fusiform face area are necessary only for the preparation of the visual contents. The objection might proceed like this:

> The areas in the inferotemporal cortex just extract all relevant visual information from the stimulus needed for the construction and expression

> of visual qualities, but do so in a way that is wholly non-qualitative. The qualitative features are added later, by other, yet higher areas in the brain—and these yet higher areas need not be visual.

I think that this reasoning has some merit, but ultimately it is unconvincing. The valuable insight present in this view is that we now have empirical evidence suggesting that higher, non-visual areas do somehow participate in conscious visual perception. For example, as Frith (2021, 4) points out, there is converging evidence from studies of semantic priming and of face perception to the effect that activations in the parietal cortex reliably co-vary with conscious visual percepts. What such results strongly indicate is that activations in the specialized visual sites such as FFA or VWFA are necessary but not sufficient for conscious vision.[5]

This evidence is important. It indicates that the entire NCC for a given type of visual content is not restricted to specialized sites in the higher visual pathways such as the FFA. However, the evidence does not cast doubt on the idea that the specialized sites in the visual cortex express visual qualities in consciousness. Rather, the evidence suggests that they cannot do this alone; they need some kind of 'push' from the higher areas in the parietal cortex. At the same time, I would argue that the involvement of these higher areas is not sufficiently content-specific to count as directly contributing to producing and expressing the qualities typical of conscious vision. For example, in one of the most important studies demonstrating the involvement of the parietal cortex in conscious vision, Beck et al. (2001) speak only of "enhanced activity" in parietal areas during conscious perception. Contrast with this the highly selective activations in the FFA and other ventral visual sites. They are specialized and respond in particular, content-specific ways to preferred types of visual stimuli. To my knowledge, there is no comparably detailed and feature-specific processing in the higher areas in the parietal cortex.

My second argument for the insufficient content-specificity of parietal activations is that these areas are implicated in diverse kinds of conscious experiences, not just in visual experiences. There is evidence that parietal cortex activations are also involved in tactile experience and emotional response awareness (for references, see Rees and Frith 2017, 599). This heterogeneous involvement of parietal areas in conscious experience again suggests that they do not furnish the qualitative contents of experience. These higher areas do not possess qualitative resources, so to speak: parietal areas cannot produce and express the visual quality of red colour or the tactile quality of smoothness. Given the evidence, it is likely that they help express these qualities in consciousness. However, they do so in a way that is not content-specific.

In sum, attempts to expel visual qualitative contents from specialized visual sites, and to move them into higher cortical regions, are unconvincing. I will now proceed to discuss the third premise of the BBA.

On 3. "The same area is similarly activated when the same visual stimulus is present but the subject is not aware of it."

Evidence that stimuli of which we are not conscious engage the same sensory areas in much the same way as consciously experienced stimuli is provided by a number of studies and reviews (see, e.g., Kanwisher 2001; Moutoussis and Zeki 2002; Vugt et al. 2018; Pojoga et al. 2020; Frith 2021). Again, faces belong to the most studied category of stimuli. It has been repeatedly demonstrated that faces that escape awareness activate the FFA in much the same way as consciously perceived faces. For instance, in patients with parietal cortex lesions, awareness of a face in the left part of their visual field can be 'extinguished' if an identical face is presented in the right part of the field. That is, the subject can see both the left-side face and the right-side face when they are presented on their own, but sees only the right-side stimulus when both are presented together. However, the extinguished face on the left side still vigorously activates the FFA in the right fusiform cortex of the subject (Rees et al. 2002; Driver and Vuilleumier 2001). Similar effects were found with healthy participants during experiments with masked faces (Kouider et al. 2009).

The BBA assumes that the close similarity in conscious and unconscious processing of the same stimulus concerns not just the same neural location of the processing site, but also the shared content-specificity of visual processing. Some authors are explicit about this. For example, Fahrenfort et al. (2017) studied unconscious perception of illusory shapes such as the well-known Kanizsa triangle with the attentional blink paradigm. The authors concluded that representations of visual stimuli which do not enter consciousness "have a neural signature that is *indistinguishable* from perceptually rich representations that occur for objects that do enter into conscious awareness" (Fahrenfort et al. 2017, 3744; emphasis added).

However, there is, again, a challenge for this premise of the BBA. The third premise claims that there is a close similarity of neural activations subserving conscious and unconscious perception. And this claim may prima facie seem to be inconsistent with a finding that has been reported by a number of studies: unconscious neural responses to the same stimulus in the same perceptual area are typically somewhat weaker than conscious ones (see, e.g., Moutoussis and Zeki 2002; Sterzer et al. 2009; Vugt et al. 2018; Fontan et al. 2021; Stein et al. 2021). For instance, a picture of a face of which the subject is not aware activates her FFA somewhat less strongly than in the conscious condition (Tong et al. 1998). By the same token, masked words elicit activation in the visual word form area, but that activation is somewhat stronger when the word is perceived consciously (Dehaene et al. 2001).

One possible interpretation of this finding is that unconscious sensory qualities do exist but are in some sense weaker, or less pronounced, than their conscious counterparts. This interpretation invites us to contemplate the following scenario: if we could—though it is impossible—peek into the unconscious part of our mind, we would see only fainter versions of the qualities we encounter while sensing consciously. If this interpretation is correct, conscious qualities are not literally the same as the unconscious ones, as the BBA claims. When the visual and other contents are brought into awareness, their qualitative character becomes somewhat more vivid.

I have two things to say to this objection from different activation strengths. First, even on the proposed interpretation of the empirical evidence, conscious and unconscious sensory qualities remain to some degree similar, so this charge might weaken my argument for unconscious qualities, but does not rebut it. Second, and more important, it is possible to explain the differences in signal strength in conscious and unconscious perceptual conditions in a way that is friendlier to the two-factor view and to the idea of unconscious sensory qualities. One could argue that there does not need to be any qualitative difference between conscious and unconscious qualities, because what accounts for weaker signals in unconscious trials is just the absence of consciousness-conferring mechanisms interacting with the qualities—recall that the active presence of such consciousness-conferring mechanisms is postulated by the two-factor theories as the explanation of what makes the sensory qualities conscious. In order for this response to the objection from different activation strengths to work, the consciousness-conferring mechanisms of the two-factor view must be thought of as entirely content-non-specific. Although the theories of consciousness based on them are not usually presented in this way, the mechanisms of neural recurrence (Lamme and Roelfsema 2000) and global broadcasting (Mashour et al. 2020), to name two prominent examples, could be thought of as subserving this function of bringing the sensory contents to the subject's awareness without contributing to their qualitative aspects.

If such content-neutral mechanisms exist, then prior to their engagement, unconscious qualities need not be any weaker or fainter than conscious ones. Interpretation of the neuroscientific evidence appealing to consciousness-conferring mechanisms would then be fully compatible with the BBA. Hence, much depends upon how the crucial experiments are interpreted, and on further experiments which will, hopefully, provide decisive support to one of the competing interpretations.

9.3 The 'finishing line'

The two-factor view based on behavioural and neural arguments asserts the sameness of unconscious and conscious sensory qualities. However, some empirical evidence points to the possibility that unconscious contents differ

in important respects from conscious ones. It has been claimed that some unconscious mental states have a higher spatial and temporal resolution than conscious ones (Lamme 2014; He and MacLeod 2001). A good example of this phenomenon is 'heterochromatic flicker fusion' (Gur and Snodderly 1997). Suppose two isoluminant circles of additive colours, say red and green, are alternately projected at the same place on the screen. When the frequency of their alternations is lower than 10 Hz we consciously see alternating red and green circles. At higher flicker frequencies (more than 10 Hz), instead of two alternating circles with different colours we consciously start to see a single non-flickering circle with a uniform, 'fused' colour (yellow if the two alternating circles are red and green).[6] However, even at these higher frequencies (up to 30 Hz), populations of cells in V1, V2, and other early visual areas continue to respond to the changes from red to green and back. Thus, these areas track the stimuli at temporal frequencies that are too high for conscious visual resolution.

Beyond demonstrating the existence of a quantifiable threshold for colour fusions, experiments on heterochromatic flicker fusion indicate that there is a mismatch between the representational content in early visual areas and the content of consciousness (Chalmers 2000, 21). Faced with such a mismatch, one may feel tempted to speculate that the fused yellow colour *only* appears in consciousness—that only the conscious part of the mind 'cooks up' the yellow circle. In consequence, one may be tempted to conclude that because of this discrepancy between conscious and unconscious sensory qualities, the two-factor view as presented so far must be wrong.

My goal in this last section is to sketch a defence of the two-factor view against this charge. To begin with, although the two-factor view posits the qualitative identity of conscious and unconscious sensory states, it is not committed to hold that there are absolutely no differences between unconscious and conscious contents. It only says that those contents that are conscious were previously formed unconsciously. In line with this claim, the two-factor theorists may argue that the unconscious mind is *richer* or *broader* than the conscious one. And in my view the two-factor theorists should take this line, in light of the data on heterochromatic flicker and similar phenomena. That is to say, when heterochromatic flicker fusion reaches the frequency of colour fusion, the fused yellow circle is formed unconsciously, before entry into the subject's consciousness. The yellow circle would, then, be that part of the sensory content, which is shared between the conscious and unconscious mind at sufficiently high flicker frequencies. But besides this shared yellow quality, the unconscious mind also contains other, more fine-grained alternating qualities of red and green, which are incapable of entering the subject's consciousness. The fact that the unconscious mind is qualitatively richer than the conscious one thus does not preclude the truth of the two-factor theory of consciousness.

This defence of the two-factor theory suggests a 'finishing line' in the brain at which the finalized sensory qualities are poised to be expressed in consciousness. In the heterochromatic flicker fusion example, at temporal frequencies below the fusion threshold, qualities of red and green circles alternate at the unconscious finishing line. At frequencies above the fusion threshold, a non-alternating yellow circle crosses the finishing line. Importantly, though, beyond the fused yellow circle at the finishing line, the earlier visual areas also contain the unconscious qualities of non-fused red and green circles.[7] Understood in this way, the two-factor theory can accommodate a discrepancy between (broader) feature selectivity and (narrower) conscious experience. And because of the broader reach of the unconscious qualitative mind, the two-factor theorist can explain the discrepancies between conscious and unconscious qualities without needing to say that the former misrepresent the latter (as does, e.g., Rosenthal 2012; see Peebles 2022 and Kirkeby-Hinrup 2016 for criticisms of this idea).

Given the bottleneck of conscious visual resolution, the sensory contents at the finishing line are only as finely grained as the conscious part of the mind can handle; some of the qualitative fineness of grain and precision is irretrievably lost along the way to the finishing line. But now a conceptual point needs to be briefly addressed. When the two-factor theorist commits to the existence of unconscious qualities at the finishing line, she can argue by parity with the conscious qualities. One knows how the qualities look and feel in the case of conscious perceptual states, so one can extrapolate to the unconscious versions of the same qualities. Such unconscious qualities are like the conscious ones, but we are just not aware of them. However, with the deeply unconscious mental states that are too fine-grained for conscious vision, the same argument by parity cannot be used. We don't know how such deeply unconscious states look or feel because we can never inspect their conscious versions; we can never experience a red-green colour flicker at 30 Hz alterations. Can we, then, call these deeply unconscious mental states *qualitative*?

I think we can, for two reasons. First, the difference between ordinary unconscious qualities and the deeply unconscious qualities is a difference of degree, not a difference of kind. The difference is just a matter of the finer temporal or spatial resolution of the deeply unconscious qualities. We are not invited to imagine some completely different qualities, incommensurable with the familiar ones. The extrapolation to the deeply unconscious qualities is, therefore, possible. We can, if only imperfectly, imagine what it would be like to experience these non-experienceable qualities, such as the super-fast flickering of red and green circles. Second, one could argue that every mental activity that has to do with the detection and identification of shapes, colours, movements (and so on) is thereby qualitative. By this criterion, even the deeply unconscious mental states qualify as qualitative.

9.4 Concluding remarks

In agreement with a growing body of empirical research, the two-factor view of sensory consciousness acknowledges the existence of unconscious sensory qualities, and their qualitative identity to the conscious sensory qualities. This qualitative sameness does not hold just between different tokens of the same type of mental state; it is numerically the same qualitative mental state that is either conscious or unconscious. This aspect of the two-factor view invites a division of labour in research on sensory consciousness: The research could be split into two branches, each tracking the distinct mechanisms that cooperate in conscious perception. While the first branch would study the processes forming sensory qualities, conscious or not, the second branch would trace the non-qualitative consciousness-conferring mechanisms.

This division of labour is bound to have important consequences for research into the neural correlates of consciousness. If the distinction between qualitative and consciousness-conferring mechanisms proves robust, it will enable us to identify the functions of different neural processes contributing to experienced mental states. In contrast, the currently prevalent NCC methodology simply lumps all correlated neural activity together as 'the NCC'. Another methodological insight stemming from the two-factor view is that the part of the NCC responsible for the qualitative character of conscious mental states also plays the role of a neural correlate of *un*conscious sensory contents (elicited by the same stimuli as the conscious percepts).

Furthermore, thanks to the distinction between qualitative and consciousness-conferring processes, the task of explaining conscious qualitative states may stop appearing intractable, because we no longer face the undifferentiated conglomerate of both consciousness and qualitativeness in the style of the dominant, one-factor approach to sensory consciousness. The hope is that at least the conscious-conferring mechanisms could be recategorized as one of the 'easy problems of consciousness' (Polák and Marvan 2019), as their only function is to uptake contents into the subject's consciousness. And because the consciousness-conferring mechanisms do no qualitative work, they can in principle serve as domain-neutral mechanisms, subserving sensory contents in distinct sensory modalities and multisensory contents.[8]

The combined strengths of behavioural and brain-based arguments for unconscious qualities robustly support the two-factor view of consciousness and challenge the dominant, one-factor view. However, the force of the two arguments is not equal. The BBA carries more weight because it is free of the shortcomings of the behavioural argument. Recall that the behavioural argument depends on a link between unconscious mental states and overt behaviour—a link that in some cases may be weakened, corrupted, or entirely missing. The BBA is unaffected by such adverse conditions: although brain activations are usually measured in the context of a specific task, the BBA can

in principle be invoked even when no behavioural evidence is produced. Even in the absence of overt behaviour, qualitative contents can still be formed in the relevant parts of our brains, and the BBA capitalizes on that. Moreover, the neural data may be more sensitive than the behavioural data (Mudrik and Deouell 2022).

As noted in section 9.2.2, it is possible, though not very probable, that activations in the sensory cortices reflect only the preparatory stages of perceptual processing, with the qualitative aspects being added later by higher cortical regions. The BBA therefore cannot conclusively establish that there is no qualitative difference between conscious and unconscious mental states. It is unclear whether any currently available research method can resolve this conundrum. Meanwhile, the BBA remains our best bet. Together with the behavioural argument, it undergirds the two-factor treatment of sensory consciousness—a treatment that promotes a systematic and general explanation of the processes and mechanisms that make unconscious contents conscious. This explanation would be an extremely important achievement. Therefore, it would be good to see more debates about the *pros* and *cons* of the two-factor view of sensory consciousness, and about how it fares in comparison with the dominant, one-factor theory.

Acknowledgments

I thank David Rosenthal, Sascha B. Fink, Lukas Kob, and the audiences at the Cognitive Science Speaker Series (CUNY Graduate Center), Valencia Philosophy Lab Colloquium, Philosophy Department at Magdeburg University, and the Prague conference *Conscious and Unconscious Mind: Commonalities and Differences* for illuminating discussions. Most of all, I thank the magnificent team of the Phenomenals: Michal, Juraj and Brice.

Work on this chapter was supported by the Czech Science Foundation (GAČR), project no. 20-14445S ('Dual Models of Phenomenal Consciousness') realized at the Institute of Philosophy, Czech Academy of Sciences.

Notes

1 As the term *phenomenal* strongly suggests consciousness, I will rather speak about unconscious qualities instead of unconscious phenomenality. I did use the latter term before (see Marvan and Polák 2017; see also Zięba 2022, and Polák, *this volume*, for closely related positions), but I now prefer, for strategic reasons, to speak just about unconscious (sensory) qualities. In this matter, I am indebted to conversations with David Rosenthal and Sam Coleman.

2 In fact, the only contemporary position on unconscious qualities with a real and close affinity to my view that I am aware of is that of Sam Coleman; see Coleman (2019; 2022; forthcoming).

3 I thank Sascha B. Fink and Lukas Kob for suggesting this helpful label.

4 This conclusion has important consequences for the methodology of studying the differences between conscious and unconscious sensory states. The aim should be to develop paradigms in which the reliance on behavioural criteria is either diminished, or even completely absent. On this, see Soto et al. (2019). The neuroscience-oriented argument presented in the following subsection is, among other things, an attempt to side-step the shortcomings of the behavioural argument.
5 Similar claims have recently been made about the role of prefrontal areas in conscious perception, but I will leave these aside in this chapter. Interested readers may consult Michel (2022).
6 Jiang et al. (2007) set the threshold for colour fusion higher than Gur and Snodderly—at 25 Hz.
7 This does not mean that when the flicker reaches the fusion threshold, the location of the finishing line in the brain must shift. The area may remain the same for both non-fused and fused circles, provided that the same neuronal population is capable of multiplexing both fused and non-fused contents. See Hesse and Tsao (2020) for an account of how the same neuronal populations multiplexed representations of Obama's face and a taco in a binocular rivalry experiment.
8 In this chapter I focused on sensory qualities. A question suggests itself whether the argument for unconscious qualities could be extended to other types of mental states such as emotional, volitional, or cognitive states. For an optimistic take, see Coleman, *this volume*; for more pessimistic assessments, see Bayne, *this volume*, and Hvorecký, *this volume*.

References

Bachmann, Talis. 2011. "Attention as a Process of Selection, Perception as a Process of Representation, and Phenomenal Experience as the Resulting Process of Perception Being Modulated by a Dedicated Consciousness Mechanism." *Frontiers in Psychology* 2: 397. https://doi.org/10.3389/fpsyg.2011.00387

Beck, Diane M., Geraint Rees, Chris D. Frith, and Nilli Lavie. 2001. "Neural Correlates of Change Detection and Change Blindness." *Nature Neuroscience* 4 (6): 645–650. https://doi.org/10.1038/88477

Bisley, James W., Daniel Zaksas, and Tatiana Pasternak. 2001. "Microstimulation of Cortical Area MT Affects Performance on a Visual Working Memory Task." *Journal of Neurophysiology* 85 (1): 187–196. https://doi.org/10.1152/jn.2001.85.1.187

Boly, Melanie, Marcello Massimini, Naotsugu Tsuchiya, Bradley R. Postle, Christof Koch, and Giulio Tononi. 2017. "Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence." *The Journal of Neuroscience* 37 (40): 9603–9613. https://doi.org/10.1523/JNEUROSCI.3218-16.2017

Celebrini, Simona, and William T. Newsome. 1995. "Microstimulation of Extrastriate Area MST Influences Performance on a Direction Discrimination Task." *Journal of Neurophysiology* 73 (2): 437–448. https://doi.org/10.1152/jn.1995.73.2.437

Chalmers, David J. 2000. "What Is a Neural Correlate of Consciousness?" In *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, edited by Thomas Metzinger, 17–39. Cambridge, MA: The MIT Press.

Cohen, Laurent, Stanislas Dehaene, Lionel Naccache, Stéphane Lehéricy, Ghislaine Dehaene-Lambertz, Marie-Anne Hénaff, and François Michel. 2000. "The Visual

Word Form Area: Spatial and Temporal Characterization of an Initial Stage of Reading in Normal Subjects and Posterior Split-Brain Patients." *Brain* 123 (2): 291–307. https://doi.org/10.1093/brain/123.2.291

Coleman, Sam. 2019. "Painfulness, Suffering, and Consciousness". In *Philosophy of Suffering*, edited by David Bain, Michael Brady, and Jennifer Corns, 55–74. London: Routledge.

Coleman, Sam. 2022. "Intentionality, Qualia, and the Stream of Unconsciousness." *Phenomenology and Mind* 22: 42–53. https://doi.org/10.17454/pam-2203

Coleman, Sam. Forthcoming. "An Argument for Unconscious Mental Qualities." *Australasian Journal of Philosophy*.

Crick, Francis, and Christof Koch. 1992. "The Problem of Consciousness." *Scientific American* 267 (3): 152–159. https://doi.org/10.1038/scientificamerican0992-152

Crick, Francis, and Christof Koch. 1995a. "Are We Aware of Neural Activity in Primary Visual Cortex?" *Nature* 375 (6527): 121–123. https://doi.org/10.1038/375121a0

Crick, Francis, and Christof Koch. 1995b. "Cortical Areas in Visual Awareness." *Nature* 377 (6547): 294–295. https://doi.org/10.1038/377294a0

Dehaene, Stanislas, Lionel Naccache, Laurent Cohen, Denis Le Bihan, Jean-François Mangin, Jean-Baptiste Poline, and Denis Rivière. 2001. "Cerebral Mechanisms of Word Masking and Unconscious Repetition Priming." *Nature Neuroscience* 4 (7): 752–758. https://doi.org/10.1038/89551

Downing, Paul E., Yuhong Jiang, Miles Shuman, and Nancy Kanwisher. 2001. "A Cortical Area Selective for Visual Processing of the Human Body." *Science* 293 (5539): 2470–2473. https://doi.org/10.1126/science.1063414

Driver, Jon, and Patrik Vuilleumier. 2001. "Perceptual Awareness and Its Loss in Unilateral Neglect and Extinction." *Cognition* 79 (1–2): 39–88. https://doi.org/10.1016/s0010-0277(00)00124-4

Fahrenfort, Johannes J., Jonathan van Leeuwen, Christian N. L. Olivers, and Hinze Hogendoorn. 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

Fontan, Aurelie, Lenita Lindgren, Tiziana Pedale, Camilla Brorsson, Fredrik Bergström, and Johan Eriksson. 2021. "A Reduced Level of Consciousness Affects Non-Conscious Processes." *NeuroImage* 244: 118571. https://doi.org/10.1016/j.neuroimage.2021.118571

Frith, Chris D. 2021. "The Neural Basis of Consciousness." *Psychological Medicine* 51 (4): 550–562. https://doi.org/10.1017/S0033291719002204

Grill-Spector, Kalanit, Nicholas Knouf, and Nancy Kanwisher. 2004. "The Fusiform Face Area Subserves Face Perception, Not Generic within-Category Identification." *Nature Neuroscience* 7 (5): 555–562. https://doi.org/10.1038/nn1224

Gur, Moshe, and D. Max Snodderly. 1997. "A Dissociation between Brain Activity and Perception: Chromatically Opponent Cortical Neurons Signal Chromatic Flicker That Is Not Perceived." *Vision Research* 37: 377–382. https://doi.org/10.1016/S0042-6989(96)00183-6

He, Sheng, and Donald I. MacLeod. 2001. "Orientation-Selective Adaptation and Tilt After-Effect from Invisible Patterns." *Nature* 411 (6836): 473–476. https://doi.org/10.1038/35078072

Hesse, Janis Karan, and Doris Y. Tsao. 2020. "A New No-Report Paradigm Reveals That Face Cells Encode Both Consciously Perceived and Suppressed Stimuli." *ELife* 9: e58360. https://doi.org/10.7554/eLife.58360

Jiang, Yi, Ke Zhou, and Sheng He. 2007. "Human Visual Cortex Responds to Invisible Chromatic Flicker." *Nature Neuroscience* 10: 657–662. https://doi.org/10.1038/nn1879

Kanwisher, Nancy. 2001. "Neural Events and Perceptual Awareness." *Cognition* 79 (1–2): 89–113. https://doi.org/10.1016/S0010-0277(00)00125-6

Keller, Andreas. 2016. *Philosophy of Olfactory Perception*. Cham: Palgrave Macmillan.

Kirkeby-Hinrup, Asger. 2016. "Change Blindness and Misrepresentation." *Disputatio* 8 (42): 37–56. https://doi.org/10.2478/disp-2016-0002

Kouider, Sid, Evelyn Eger, Raymond Dolan, and Richard N. Henson. 2009. "Activity in Face-Responsive Brain Regions Is Modulated by Invisible, Attended Faces: Evidence from Masked Priming." *Cerebral Cortex* 19 (1): 13–23. https://doi.org/10.1093/cercor/bhn048

Lamme, Victor A. F., and Pieter R. Roelfsema. 2000. "The Distinct Modes of Vision Offered by Feedforward and Recurrent Processing." *Trends in Neurosciences* 23 (11): 571–579. https://doi.org/10.1016/s0166-2236(00)01657-x

Lamme, Victor A. F. 2014. "The Crack of Dawn: Perceptual Functions and Neural Mechanisms That Mark the Transition from Unconscious Processing to Conscious Vision." In *Open MIND*, edited by Thomas Metzinger and Jennifer M. Windt, 1–33. Frankfurt am Main: MIND Group.

Lockwood, Michael. 1989. *Mind, Brain, and the Quantum: The Compound 'I'*. Oxford: Basil Blackwell.

Marvan, Tomáš, and Michal Polák. 2017. "Unitary and Dual Models of Phenomenal Consciousness." *Consciousness and Cognition* 56: 1–12. https://doi.org/10.1016/j.concog.2017.09.006

Marvan, Tomáš, and Michal Polák. 2020. "Generality and Content-Specificity in the Study of the Neural Correlates of Perceptual Consciousness." *Philosophy and the Mind Sciences* 1 (2): 1–17. https://doi.org/10.33735/phimisci.2020.II.61

Mashour, George A., Pieter R. Roelfsema, Jean-Pierre Changeux, and Stanislas Dehaene. 2020. "Conscious Processing and the Global Neuronal Workspace Hypothesis." *Neuron* 105 (5): 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

Michel, Matthias. 2022. "Conscious Perception and the Prefrontal Cortex: A Review." *Journal of Consciousness Studies* 29 (7–8): 115–157. https://doi.org/10.53765/20512201.29.7.115

Moutoussis, Konstantinos, and Semir Zeki. 2002. "The Relationship between Cortical Activation and Perception Investigated with Invisible Stimuli." *Proceedings of the National Academy of Sciences* 99 (14): 9527–9532. https://doi.org/10.1073/pnas.142305699

Mudrik, Liad, and Leon Y. Deouell. 2022. "Neuroscientific Evidence for Processing Without Awareness." *Annual Review of Neuroscience* 45 (1): 403–423. https://doi.org/10.1146/annurev-neuro-110920-033151

Peebles, Graham. 2022. "The Problem of Higher-Order Misrepresentation." *Philosophical Psychology* 35 (6): 842–861. https://doi.org/10.1080/09515089.2021.2016677

Pojoga, Sorin A., Natasha Kharas, and Valentin Dragoi. 2020. "Perceptually Unidentifiable Stimuli Influence Cortical Processing and Behavioral Performance."

Nature Communications 11 (1): 6109. https://doi.org/10.1038/s41467-020-19848-w

Polák, Michal, and Tomáš Marvan. 2019. "How to Mitigate the Hard Problem by Adopting the Dual Theory of Phenomenal Consciousness." *Frontiers in Psychology* 10: 2837. https:// doi.org/10.3389/fpsyg.2019.02837

Rangarajan, Vinitha, Dora Hermes, Brett L. Foster, Kevin S. Weiner, Corentin Jacques, Kalanit Grill-Spector, and Josef Parvizi. 2014. "Electrical Stimulation of the Left and Right Human Fusiform Gyrus Causes Different Effects in Conscious Face Perception." *Journal of Neuroscience* 34 (38): 12828–12836. https://doi.org/10.1523/JNEUROSCI.0527-14.2014

Rees, Geraint, and Chris D. Frith. 2017. "Methodologies for Identifying the Neural Correlates of Consciousness." In *The Blackwell Companion to Consciousness*, edited by Max Velmans and Susan Schneider, 589–606. Malden, MA: John Wiley.

Rees, Geraint, Ewa Wojciulik, Karen Clarke, Masud Husain, Chris Frith, and Jon Driver. 2002. "Neural Correlates of Conscious and Unconscious Vision in Parietal Extinction." *Neurocase* 8 (5): 387–393. https://doi.org/10.1076/neur.8.4.387.16190

Rosenthal, David. 2005. *Consciousness and Mind*. Oxford: Oxford University Press.

Rosenthal, David. 2012. "Higher-Order Awareness, Misrepresentation and Function." Philosophical Transactions of the Royal Society B: Biological Sciences 367 (1594): 1424–1438. https://doi.org/10.1098/rstb.2011.0353

Soto, David, Usman Ayub Sheikh, and Clive R. Rosenthal. 2019. "A Novel Framework for Unconscious Processing." *Trends in Cognitive Sciences* 23 (5): 372–376. https://doi.org/10.1016/j.tics.2019.03.002

Stein, Timo, Daniel Kaiser, Johannes J. Fahrenfort, and Simon van Gaal. 2021. "The Human Visual System Differentially Represents Subjectively and Objectively Invisible Stimuli." *PLOS Biology* 19 (5): e3001241. https://doi.org/10.1371/journal.pbio.3001241

Sterzer, Philipp, Lauri Jalkanen, and Geraint Rees. 2009. "Electromagnetic Responses to Invisible Face Stimuli during Binocular Suppression." *NeuroImage* 46 (3): 803–808. https://doi.org/10.1016/j.neuroimage.2009.02.046

Tong, Frank, Ken Nakayama, J. Thomas Vaughan, and Nancy Kanwisher. 1998. "Binocular Rivalry and Visual Awareness in Human Extrastriate Cortex." *Neuron* 21 (4): 753–759. https://doi.org/10.1016/s0896-6273(00)80592-9

Vugt, Bram van, Bruno Dagnino, Devavrat Vartak, Houman Safaai, Stefano Panzeri, Stanislas Dehaene, and Pieter R. Roelfsema. 2018. "The Threshold for Conscious Report: Signal Loss and Response Bias in Visual and Frontal Cortex." *Science* 360 (6388): 537–542. https://doi.org/10.1126/science.aar7186

Young, Benjamin D. 2014. "Smelling Phenomenal." *Frontiers in Psychology* 5: 713. https://doi.org/10.3389/fpsyg.2014.00713

Young, Benjamin, Andreas Keller, and David Rosenthal. 2014. "Quality-Space Theory in Olfaction." *Frontiers in Psychology* 5: 1. https: //doi.org/10.3389/fpsyg.2014.00001

Zięba, Paweł Jakub. 2022. "Seeing Colours Unconsciously." *Synthese* 200 (3): 260. https://doi.org/10.1007/s11229-022-03687-x