11 The Emperor's New Phenomenology? The Empirical Case for Conscious Experiences without First-Order Representations

Hakwan Lau and Richard Brown

A major divide in philosophical theories of phenomenal consciousness is that between higher-order and first-order approaches. First-order states are those states that represent objects or properties in the world, and higher-order states are those that represent other (first-order) mental states.[1] So the divide between first-order theories and higher-order theories corresponds roughly to acceptance of what David Rosenthal calls the transitivity principle, which states that a perceptual state's being *phenomenally* conscious consists in awareness of oneself as being in that state (in some suitable way) or, conversely, that a perceptual state of which one is in no way aware of being in cannot be phenomenally conscious.[2] For instance, on a specific version of the higher-order view known as the higher-order thought theory (Rosenthal 2005), to make a first-order perceptual state conscious, a thought-like mental representation to the effect of being in that particular perceptual state is needed. This is because awareness of being in some state or other requires that one represent oneself as being in that state, and a plausible way of doing that is to deploy an intentional thought-like representation to the effect that one is in some first-order state or other. There are different ways of implementing the transitivity principle, but most, if not all, versions of the higher-order approach adopt this principle. In contrast, first-order theories reject this principle and hold that merely having the right kind of first-order state is necessary and sufficient for phenomenal consciousness.

Because the higher-order approach involves two levels of representations, critics have long challenged the view by considering hypothetical cases in which the content of the first-order and higher-order representations mismatch (Neander 1998; Balog 2000; Levine 2001). If individuals are aware of themselves as seeing blue when they really

We would like to thank Rocco Gennaro, Pete Mandik, Adam Pautz, Ian Phillips, Adrianne Prettyman, Jesse Prinz, David Rosenthal, Daniel Stoljar, and Josh Weisberg for helpful comments on an earlier draft. An earlier version of this chapter was presented at the New York Consciousness Project meeting September 21, 2011, and we are grateful for the enormously helpful discussion at that meeting. Hakwan Lau is supported by the Templeton Foundation (grant number 21569).

have a first-order representation of redness, what is it like for them? If they consciously experience redness in such cases, then in a sense we have rejected the higher-order approach; the conscious experience follows the first-order content, and this seems to render the higher-order representation irrelevant. But if they consciously experience blueness, then the first-order states seem to play no role in consciousness. In response to this, Rosenthal and others have argued that it is the higher-order representation that determines the character of the conscious experience (Rosenthal 2005, 186, 187, 203; Weisberg 2011). Though the first-order representations play a relatively indirect role with respect to consciousness, it has been independently argued they nonetheless be considered as qualitative mental states (Rosenthal 2005, 38–39).[3] Thus, because it is the higher-order representations that really determine the character of conscious experience—that is, determine what it is like for an individual to have the experience—even when the first-order representation is missing, that person's conscious experience should not be altered.[4] As a result of this, many higher-order theorists thought that nothing more is needed to be said about this problem. Misrepresentation may seem intuitively odd, but it is nothing that the higher-order theory cannot handle.

Block has put a novel twist on this traditional worry (Block 2011a). He argues that if we claim that we are just as conscious in the empty higher-order case as in the Full-Conscious-Experience case, this seems to violate the transitivity principle. If conscious perception happens in the empty higher-order case, it would no longer be because a first-order perceptual state becomes conscious as we become aware of being in it. There is no such first-order state to begin with, and a nonexistent state cannot become conscious. So what is conscious seems to be the higher-order state, not the first-order state. Thus, the transitivity principle, which states that a first-order state becomes conscious when we are aware of being in it, seems to be wrong.

Though proponents of the higher-order view have resisted this argument (Rosenthal 2011; Weisberg 2011), Block takes it that considerations of empty higher-order states motivate rejecting the higher-order view and in turn support his own view, which is a variant of a first-order view.[5] Our goal in this chapter is to bring a new perspective to this debate by employing a strategy for which Block himself is well known: we consider empirical cases that bear on the philosophical issues. We argue that there are plausible cases in which subjects seem to enjoy conscious experience and yet the relevant first-order states are either missing or insufficient to explain the reported phenomenology. These may be considered to be empirically plausible empty higher-order state cases. That empirical results suggest that such cases actually exist means these cases are not just meant to be a hypothetical conceptual problem for a particular theory. Any successful theory will have to be able to interpret these cases. The upshot is that it turns out that this strategy backfires, as they may actually be more troublesome for a first-order view rather than a higher-order view. In brief, the first-order view has difficulty accounting for these cases. In addition, we argue that the higher-order view can

account for them under certain interpretations of the transitivity principle without giving up the core ingredients of the higher-order approach.

Plausible Empirical Cases of Empty Higher-Order Representations with Phenomenally Conscious Experience

In this section we discuss three empirical cases that we take to be plausible candidates for empty higher-order representations. These three cases are Rare Charles Bonnet Syndrome, Inattentional Inflation, and Peripheral Vision (together we call these the Empirical Cases). Our aim in this section is to simply present these cases without arguing for any interpretation of them. In the subsequent section we examine various interpretations of the Empirical Cases.

To relate the philosophical issues to empirical data we first need to consider plausible neuroanatomical interpretations of the substrates of first-order and higher-order representations. Fortunately, the general picture is not controversial. Authors, including Block himself, typically find it plausible that first-order representations are reflected by activity in early sensory regions and that higher-order representations are reflected by activity in the prefrontal cortex—and possibly the parietal areas too (Lau and Rosenthal 2011). This lines up nicely with what is traditionally meant by these terms. First-order states are those that represent the world outside, so to speak, and those representations are likely found in the early sensory areas. Likewise, higher-order states represent other mental states, and those representations are found in the higher-functioning areas of the brain. We hereby adopt this interpretation for the purposes of this chapter, such that when we say "first-order representations," we simply mean early sensory activity that represents the contents of perception.[6] In this sense, when the relevant early sensory activity is missing, the first-order representation is missing, even though the perceptual content may be somehow represented (e.g., at some higher level of processing).[7]

Though there is general agreement about activity in early sensory regions being involved, there is some debate as to what the exact neural correlates of first-order representations really are. One view suggests that awareness critically depends on activity in the feedback projection of extrastriate activity back to the primary visual cortex (V1) (Lamme 2006). For instance, for conscious perception of motion, cortical activity typically first arises at V1, and then it travels through feed-forward connections to the motion-sensitive area MT/V5. Conscious experience seems to critically depend on projection of activity from MT/V5 back to V1. Though popular, this view has its critics (e.g., Silvanto and Rees 2011). However, Block endorses, or at least favors, this view (Block 2005, 2007), and therefore we consider feedback to V1 as the primary candidate for the correlate of first-order representations. In suitable places below we consider the alternatives and consequences if Block gives up this specific view.

Assuming that first-order representations are supported by feedback to V1, cases in which subjects report conscious visual experience following damage to V1 can be considered potential cases of conscious awareness with empty higher-order representations. This occurs in a rare form of Charles Bonnet Syndrome.

In typical cases of Charles Bonnet Syndrome (Ffytche 2005), which in general can result from various forms of brain damage, patients report that they experience vivid hallucinations of objects such as faces, familiar persons or objects, and complex geometrical designs.[8] However, the patients are otherwise cognitively intact. Unlike in schizophrenic or other psychotic experiences, sufferers of Charles Bonnet Syndrome typically show no sign of irrational fear regarding the hallucinations and often can lucidly describe the content of the hallucination and accept that it is a visual deficit. Therefore, there are good reasons to believe that they actually go through these visual experiences, rather than their merely thinking that they do. Here we are concerned with a rare form of Charles Bonnet Syndrome (we call these Rare Charles Bonnet Cases), which results from damage to the occipital areas including the primary visual cortex (Duggal and Pierri 2002; Ashwin and Tsaloumas 2007; Contardi et al. 2007). Since these patients lack an intact primary visual cortex, on which first-order representations presumably critically depend, it seems plausible that some higher mechanisms are driving such hallucinations.[9]

Such cases are dramatic but rare. In the laboratory, we can demonstrate a related but less dramatic phenomenon. Of course, we cannot deliberately lesion the visual cortex of human subjects to completely abolish the first-order representations, but we can find cases in which the first-order representations seem too weak to generate the reported conscious experience. The strength of a first-order representation can be estimated in terms of signal-to-noise ratio, which can be indirectly assessed through behavior (Green and Swets 1966) or by brain imaging measurements. We can try to find cases in which the relevant early sensory brain areas are activated but not robust enough to account for the level of phenomenology reported by subjects. For instance, we can have two cases in which the levels of activation in the visual cortex are the same, and yet in one case the level of phenomenology is reported to be richer or more vivid. This would constitute a case in which the putative higher-order representations are not exactly empty, because there is actually a relevant first-order representation. However, this is similar in spirit to the empty higher-order cases because the first-order representation is not strong enough to account for the conscious experience, and thus there is still some degree of emptiness that needs to be explained. All we mean by this is that there is more to the experience phenomenologically than can be accounted for by the first-order representations (as reflected by activity in early sensory areas).

For example, Rahnev, Maniscalco, and colleagues (2011) presented grating patterns of strong luminance contrast to the unattended locations of the visual field and grating patterns of weak luminance contrast to the attended locations, such that the

forced-choice task performance was matched between the attended and unattended. Specifically, the authors used signal detection theoretic analysis to estimate the signal-to-noise ratio of visual processing—a measure known as d' (pronounced d-prime). Under matched d' conditions, subjects produced more hits and false alarms for detection of gratings in the unattended locations; that is, subjects reported they see the grating target more frequently in the unattended location, even though the strength or capacity of signal processing in that location was no higher than in the attended location. In another condition, the task was to determine the orientation of the grating (left tilted vs. right tilted) rather than to detect a target. In this case, subjects gave higher visibility ratings for stimuli presented in the unattended locations, even though they were no better at discriminating them (d' was again matched between unattended and attended).

The signal-to-noise ratio measure, d', is estimated from behavioral data. To directly assess the level of neural activity in the brain, functional magnetic resonance imaging (fMRI) is available. So in a follow-up study, Rahnev, Bahdo, and colleagues (2012) used fMRI to track the spontaneous fluctuation of brain activity in brain areas that are known to be critical for spatial attention (in the intraparietal region and frontal eye fields). When such activity was low, which presumably reflected a low state of attention, subjects gave higher confidence ratings in a discrimination task—again, even though their ability to discriminate between the stimuli was not higher. Also, the average intensity of activity in the early visual areas was not higher in either the attended or unattended conditions.

We can call this phenomenon Inattentional Inflation of Subjective Perception (Inattentional Inflation for short), which on the face of it contradicts the general notion that attention boosts visibility and the subjective appearance of objects (Carrasco, Ling, and Read 2004). The critical difference here is that in the experiments discussed, the focus is on the changes in subjective ratings (confidence or visibility) or detection bias (propensity to give a yes response in detection, resulting in higher hit rate and false-alarm rate) when effects of signal processing capacity (as assessed by d', task performance, and other factors) were already controlled for. Previous studies have found that attention boosts visibility at the same time that it positively changes the signal processing capacity itself, as well as boosted brain activity in the visual cortex (Carrasco 2011). On the other hand, the experiments described here (Rahnev, Maniscalco, et al. 2011; Rahnev, Bahdo, et al. 2012) showed that, independently from this positive impact on signal processing capacity and visual activity, there was also a negative impact on subjective reports of conscious experience.

The results of Rahnev, Maniscalco, and colleagues (2011), and Rahnev, Bahdo, and colleagues (2012) are based on precise laboratory measurements, but they are also somewhat technical and perhaps unintuitive. So finally we consider a case from everyday experience, which is the case of Peripheral Vision.

Introspection suggests, to us at least, that Peripheral Vision is colorful and vivid. Perhaps it is subjectively less so than foveal vision, but typically the sense is that a fair amount of color can be seen in the periphery and that, when we take a single quick glance at an unknown scene, the perception of detail is relatively uniform—that is, the detail in the periphery is not entirely missing. However, it is not clear if the physiology of the retina and the wiring of its input to the primary visual cortex can offer such detail of processing (Azzopardi and Cowey 1993; Newton and Eskew 2003). It seems likely, as Rahnev, Maniscaldo, and colleagues (2011) found, that some similar subjective inflation is at work for Peripheral Vision.

This last example of Peripheral Vision is perhaps the least decisive, because it relies on introspection and some may disagree about the exact phenomenology. However, the point is to complement the results of Rahnev, Maniscaldo, and colleagues (2011).

To sum up, there are three kinds of Empirical Cases—Rare Charles Bonnet Cases (i.e., Charles Bonnet Cases that result specifically from damage to the primary visual cortex), Inattentional Inflation (i.e., the results of Rahnev, Bahdo, et al. 2012), and Peripheral Vision (introspective evidence from everyday life). The three cases serve slightly different purposes. The Rare Charles Bonnet Cases highlight the possibility of vivid conscious experience in the absence of primary visual cortex. If we take the primary visual cortex as the neural structure necessary for first-order representations, this is a straightforward case of conscious experience without first-order representations. In Inattentional Inflation, the putative first-order representations are not missing because of lack of attention, but they are not strong enough to account for the inflated level of reported subjective perception, in that both behavioral estimates of the signal-to-noise ratio of processing and brain imaging data show that there was no difference in overall quality or capacity in the first-order perceptual signal, which concerns not only the primary visual cortex but also other relevant visual areas. Finally, Peripheral Vision gives introspective evidence that conscious experience may not faithfully reflect the level of detail supported by first-order visual processing. Though this does not depend on precise laboratory measures, it gives an intuitive argument that is not constrained by specific experimental details.

We argue that these are plausible empty higher-order cases, in the sense that since the first-order representations in question seem to be either missing or too weak to account for the conscious experience, it seems plausible that stipulating higher-order mechanisms may be necessary to provide a coherent account. However, exactly how this argument goes, and what implications it has for the first-order and higher-order theories, depends on our interpretation of the phenomenology in these Empirical Cases. Authors like Block may not want to take the reported conscious experience in these cases at face value. In the following sections, we argue that the higher-order approach can account for these cases and that the first-order view has difficulty accounting for them.

The No-Conscious-Experience Interpretation

One possible interpretation of our putative cases of empty higher-order representations is to deny that there is actual conscious experience when the first-order representations are missing. That is, in our Rare Charles Bonnet Cases, perhaps the patients were only *thinking* that they have those phenomenal experiences without actually having them. Likewise, for the case of Inattentional Inflation, one can try to deny that there is actually a higher degree or intensity of conscious experience when the strength of the first-order representations remained similar between attended and unattended conditions. Subjects reported a higher degree or intensity to be the case only because of some cognitive or reporting bias, but such reports did not faithfully reflect the actual character of the relevant conscious experience. Similarly for Peripheral Vision, perhaps we do not actually experience vivid colorful details. We only *think* that we do.

Denying the empirical plausibility of these cases will keep Block's philosophical position intact; that is, if there is no conscious experience in the Empirical Cases, they are in general compatible with a first-order view. But we find such denial outlandish. Though other authors have in other cases denied reported conscious experience as real, such views have not been popular. For instance, attempts have been made to deny phenomenology in dreams (Dennett 1976), by arguing that the dreamer is actually not enjoying any conscious experiences while dreaming. It is only during wakeful recall that the dreamer re-creates the conscious experience and remembers the dreams *as* conscious. Though some authors hold this view, like many others we find it implausible. Importantly, in dreams individuals are rarely asked whether they are conscious during the moment of dreaming. But in the empirical cases we review here, they are not asked to *remember* if a certain experience was conscious. In many of these cases a person can be asked the question at the moment of the conscious experience. If the person says he or she is vividly experiencing something *right now*, who are we to deny such claims?[10]

Importantly, as explained above, in our Rare Charles Bonnet Cases the patients were cognitively largely intact. We have no more reason to doubt their introspective reports of hallucinations than we do to doubt that ordinary subjects are truthful in claiming themselves to be conscious. It is perhaps hard to ascertain whether such hallucinations are phenomenologically identical to normal perception, but to deny that there is any conscious experience associated with the reported hallucinations seems extremely difficult to defend.

Denying that the subjective reports in the Inattentional Inflation experiments actually reflect conscious experience is also somewhat problematical. Recall that in those experiments, when task performance was matched between attended and unattended locations, subjects responded yes more frequently in a detection task (i.e., higher hit rates and false-alarm rates), and they also gave higher subjective ratings of visibility in a

discrimination task. Crucially, it is the *combination* of both results of detection bias and subjective ratings of visibility that makes it appealing that there is some genuine difference in phenomenology between the attended and unattended. If a combined increase in the frequency of saying, "Yes, I see the target," and higher visibility ratings is not good enough evidence that phenomenology changed, what else can count as good evidence? To deny that is to deny the common standard of interpretation of experiments. Also, these results are unlikely to be just a cognitive effect (i.e., a subject tries to use different responding strategies for attended vs. unattended locations), because in Rahnev, Bahdo, and colleagues (2012) there were also conditions under which subjects were encouraged not to be biased by being given trial-by-trial feedback as to what were the correct answers and so on, and it was found that the differences between attended and unattended locations were resistant to these changes in experimental context, as if the bias was automatic, over which subjects have little control.

Also, there is the case of Peripheral Vision to consider. That did not depend on any particular procedure of a psychological experiment, but just that introspectively, it seems (to us at least) as though Peripheral Vision gives a higher degree and intensity of conscious experience than can be afforded by the underlying physiological mechanisms at the early visual processing level.

To sum up, we think denying that our putative empty higher-order cases involve real conscious phenomenology is unattractive. One of these cases may be resisted, but it is difficult to see how a unified interpretation can resist all three at the same time. It is extremely implausible that there is no phenomenology at all for the hallucinations in the Rare Charles Bonnet Cases. One can argue that those cases involved lesions including only V1 and not the entire visual cortex, and therefore maybe some weak first-order representations still exist. But that requires giving up the notion that feedback to V1 is critical for awareness, something that Block may not want to do (more discussion below). Inattentional Inflation and Peripheral Vision involve converging evidence from different kinds of reports that reflect conscious experience. To claim that in all three cases all reported phenomenology merely reflects some cognitive or reporting biases seems outlandish.

The Full-Conscious-Experience Interpretation

Since denying that there is any conscious experience in all three of the Empirical Cases seems an unappealing option, one may choose to just accept that there is normal phenomenology in these cases. To do this would be to hold that in Rare Charles Bonnet Syndrome subjects actually consciously experience the things they say they do, in Inattentional Inflation subjects are actually consciously experiencing the unattended stimuli more strongly and richly than they do for the attended targets, and in Peripheral Vision we really do experience color and vivid details as introspection suggests.

If we accept that there is full conscious experience in the Empirical Cases, that may be problematic for the first-order view, because the first-order view holds that having first-order representations is necessary and sufficient for conscious experience. Here, however, in Rare Charles Bonnet Cases the primary visual cortex is missing, and so if we accept that such patients enjoy normal conscious experience, it seems to violate the first-order view.

At this point the first-order theorist may suggest that there are enough first-order representations to account for the phenomenology in the Empirical Cases. After all, in none of the Rare Charles Bonnet Cases described was the entire visual cortex damaged. Yet the damage seemed to involve V1, and this poses a challenge to the view that first-order representations critically depend on feedback to V1. Block is not necessarily committed to this empirical claim and can alternatively identify first-order representations with extrastriate activity (i.e., visual areas other than V1; see Prinz 2005).[11] The notion that awareness critically depends on feedback to V1 has been independently criticized on empirical grounds (Macknik and Martinez-Conde 2008; Silvanto and Rees 2011), and we think locating first-order representations in extrastriate cortex is superior to the feedback-to-V1 view.

However, one reason that Block may prefer something like the feedback-to-V1 view relates to his philosophical position. For a standard first-order representational theorist, the content of awareness is driven by the *content* of the first-order representations. Although we do not have a complete understanding of the content reflected by activity in different visual areas, we surmise that it is plausible that extrastriate areas can support the suitable contents for the hallucinations in the Rare Charles Bonnet Cases, which are mainly objects, geometric shapes, faces, and the like. However, on Block's view, it is the *biological substrate* of the first-order representation that is critical for conscious phenomenology. Presumably, the feedback-to-V1 view is attractive to him because the recurrent processing reflected by the feed-forward and feedback waves of neural activity seems to give a flavor of a specialized biological phenomenon. If Block is to abandon this view, he would need to specify what is special about extrastriate activity that allows it to support conscious phenomenology. Is it not just normal neural coding, which sometimes can reflect unconscious processing too?

And also, there are the cases of Peripheral Vision and Inattentional Inflation for the first-order theorist to worry about, if we accept the suggested conscious experience at face value. In Peripheral Vision, it is not clear how the relevant first-order representations can exist, because even at the retinal level the relevant input is not rich enough. One can perhaps argue that the color sensation and vividness of details in the first-order representation are created from top-down mechanisms, but one needs to substantiate such empirical claims. In our own introspective experience, even if we open our eyes for a brief period to a new scene, we get the phenomenological feeling that the periphery is not exactly monochrome and devoid of details. It seems, at least to us, as if there

is a kind of phenomenological inflation there, and it is unlikely that we are filling in instantaneously in the periphery with memory (because the scene is novel to us).

In the case of Inattentional Inflation, we also argue against similar top-down mechanisms. Not only is it odd to suppose extra top-down mechanisms are at work when one is *not* paying enough attention (which is a top-down mechanism in itself); the fMRI results showed no difference in overall intensity of activity in the early visual areas between the conditions in which spontaneous fluctuation of attention differed. And yet subjects gave higher subjective ratings for their visual discrimination when attention was at a low state, even though they were no better at the discrimination. Importantly, Rahnev, Bahdo, and colleagues (2011) offer a computational account of this finding that is essentially compatible with a higher-order approach. The model assumes that attention reduces the variability of the (first-order) perceptual signal. Because of the higher variability of the perceptual signal in the unattended case, the quality or strength of the perceptual process is low under the lack of attention. However, as in standard models of perception (Green and Swets 1966), subjective perception happens when the signal crosses a threshold or criterion. Importantly, the model assumes that this criterion is set after consideration of the statistical properties of both the internal signals for the attended and unattended stimuli. This makes the criterion setting essentially a higher-order mechanism in that representing the properties of the first-order states sets the criterion. Because human subjects must use the same criterion for both the attended and unattended if they are presented simultaneously (a known psychophysical fact based on previous work; Gorea and Sagi 2000), the higher variability of the internal signal under the lack of attention turns out to lead to more frequent crossing of the criterion—that is, more frequent occurrence of subjective perception. This model provides a good fit to the experimental data and accounts for why, under the lack of attention, subjects are more likely to report that, yes, they see the target and to give higher confidence and visibility ratings in discrimination.

In other words, within the context of first-order versus higher-order mechanisms, attention does change the variability of the perceptual signal itself, which we can consider an influence on the first-order representation. However, even when we present a stronger stimulus to the unattended location such that the signal-to-noise ratio of the first-order representation would be matched between the attended and the unattended locations, there would still be a difference in subjective perception. This is because, according to the model, subjective perception happens when the first-order signal crosses a criterion. When the criterion is fixed, a more variable signal, albeit noisy, can cross the criterion frequently because of the higher fluctuation. We argue that this criterion is determined on the basis of higher-order mechanisms, because setting the criterion requires taking into account the statistical properties of internal signals, such as the baseline activity level when no target is presented. To represent things like baseline activity level and its variance is to have higher-order representations, because these are

properties of first-order representations, rather than properties of objects in the world. Importantly, even if there are other ways to determine the criterion, such mechanisms likely reside in the prefrontal cortex (Lau and Rosenthal 2011), which means that they count as higher order for the purpose of the arguments in this chapter (as stipulated at the beginning of this section).

So if there is indeed normal conscious experience in the three Empirical Cases, the first-order view may be in trouble. But why should we think that there is really normal conscious experience in these cases? When we considered the no-conscious-experience interpretation above, we gave reasons for why denying the reported experience is an unattractive option. But is there any reason to think that accepting such reports at face value is attractive? One argument for why this may be positively attractive is that this allows for a parsimonious view of how different lines of evidence converge. This is similar in spirit to Block's very own mesh argument (Block 2007, 2008).

The mesh argument has roughly the following form. We should adopt the theory of conscious experience that allows for the most parsimonious explanation of the relationship between data at the neuroscientific level and data at the psychological level. When evaluating theories we should take into account as wide a swath of evidence as possible and look for the theory that gives a unified simple explanation of the various empirical discoveries. Block has tried to use this argument to show that we should accept the claim that there are two separable systems of consciousness: one supporting phenomenally conscious experiences and one supporting conscious access to those phenomenally conscious experiences. If we accept this, then we can parsimoniously interpret two results. The first is what he calls phenomenological overflow; that is, in many experiments such as change blindness and inattentional blindness, subjects can report the details of only a few objects, but subjectively they seem to see the entire visual scene in front of them. The other result is the neurobiological finding that the posterior visual system in the occipital and temporal lobes seems to have a higher informational capacity and resolution than the prefrontal system. If we map phenomenally conscious experience onto the posterior visual system and conscious access to the prefrontal system, then we can see why there is phenomenological overflow: because the former has higher informational capacity and resolution than the latter. Therefore, there is phenomenal conscious experience that we cannot report or access.

If this conclusion is true, perhaps so much the worse for ambitious versions of the higher-order theory, which denies that there can be phenomenal conscious experience without a higher-order representation, the latter of which presumably resides in the prefrontal cortex. If phenomenally conscious experience can overflow the prefrontal system, it seems that the ambitious version of higher-order theory is wrong. However, in the light of our Empirical Cases, we suggest that the ambitious version of higher-order theory can actually account for the possibility of phenomenological overflow just as well (Brown 2012a). In fact, it provides a better mesh for all evidence—because

it allows us to accept the reported conscious experiences in the Empirical Cases at face value.

First, we note that the ambitious higher-order approach fits just as well to the neurobiological finding that there is a high-capacity early visual system and a low-capacity prefrontal system. Both the higher-order view and the first-order view (i.e., one favored by Block) assume that there are two stages of visual processing. What is at issue is whether the correlates of conscious awareness are supported by the late-stage low-capacity system or the early-stage high-capacity system. On the first-order model the correlates of conscious experience go with the high-capacity system. On the higher-order model the correlates of conscious experience go with the low-capacity system. Both claims are compatible with the neurobiological finding.

What differs between the two views is the way they account for phenomenological overflow. For the first-order view, because access is supported by the late-stage system, which has lower capacity than the early-stage system that supports conscious experience, one naturally expects phenomenological overflow. On the other hand, the higher-order view may seem to have some difficulty in accounting for phenomenological overflow. If conscious phenomenology is associated with the late-stage low-capacity system, how can it seem richer than what is reflected by access or task performance? On the higher-order view, the capacity for task performance is reflected by the capacity of first-order representations. If we associate the first-order representation with the early-stage visual system, which has high informational capacity and resolution, shouldn't we expect the opposite of phenomenological overflow, that conscious phenomenology is less detailed than what is reflected by performance?

A natural solution for the higher-order theorist is to deny that the richness of conscious experiences is determined by the informational capacity of the relevant neural system. Equating the richness of conscious experiences with the informational capacity of the relevant neural system is appealing only if we assume that the conscious experience is veridical. In phenomenological overflow, when the subject apparently experiences vivid details of the entire visual scene but can report the identity of only a few objects, one interpretation is that the experience of richness is not veridical. This is not to deny that subjects experience such vividness but to say that, in reality, they do not represent the visual scene in such a vivid and rich way. In the famous Sperling postcue experiments (1960), subjects had the impression of clearly seeing twelve letters, despite being able to report accurately the identity of only about four of them at a time. An interpretation of the results is that the higher-order mental state represents oneself as vividly seeing the identities of all twelve letters without specifying what they are (Brown 2012a, 2014). Because the higher-order system does not actually represent the identities of the twelve letters, the representation that one is vividly seeing the identities of all letters is nonveridical. But because such higher-order representations do not carry

the actual information regarding the identities of the letters, this does not require an actual information capacity for all twelve items.

However, in the Sperling experiments, subjects could report *any* four of the letters if they were cued immediately after the letters disappeared. This suggests that subjects must have had some form representations of all twelve letters too. We do not deny these results, but such representations are traditionally considered a form of iconic memory, and it is unclear whether they are *conscious* representations. Block (2011c) has argued that there is a general lack of evidence for such unconscious memory (see Brown 2014 for a response). We agree that when considering unconscious memory as in the experiments in which the stimuli were masked and could not be perceived at all in the first place, it is not clear if subjects could form memories with such high capacity. But we should not conflate whether the memory representations are conscious with whether the stimuli are presented consciously in the first place. What we suggest is that in Sperling experiments or their variants, when the stimuli were consciously presented, the iconic memory representations for the stimuli may nonetheless not be conscious throughout the delay period. This issue may need to be empirically resolved by future studies. But at least so far there is no empirical evidence directly against the possibility that subjects did not have conscious detailed representations of the letters throughout the delay. The subjective impression that such memory is phenomenally conscious may well be supported by nonveridical higher-order representations.

But is this move of invoking nonveridical higher-order representations ad hoc? Certainly not, for this is exactly what we expect on the basis of the Empirical Cases that motivate this chapter. For example, in the case of Inattentional Inflation, subjects claimed to see more clearly under lack of attention, despite being no better at performing the visual tasks. This is compatible with the interpretation that the higher-order system can overestimate the richness of perception. Also, as with Inattentional Inflation, phenomenological overflow tends to happen when focused attention is lacking (e.g., in the Sperling experiments or in inattentional and change blindness experiments).

Given that the higher-order view can account for phenomenological overflow just as well as the first-order view, we can say that the higher-order view is superior, because it allows us to accept the Empirical Cases at value face. That is, unlike the first-order view, it does not require us to make ad hoc claims that deny the phenomenology in the Empirical Cases. Thus, in the spirit of the mesh argument itself, one should prefer the (ambitious) higher-order view.

To sum up: we have good reason to believe that there is real phenomenology in our empirical cases, and we argue that this favors the higher-order view.

Nevertheless, Block seems to think otherwise. If there are indeed full conscious experiences in the Empirical Cases, Block's challenge to the higher-order view is that it seems to violate the transitivity principle, which states that a first-order state is conscious if

and only if an individual is conscious of being in such a state. In the Empirical Cases, especially the Rare Charles Bonnet Cases, the putative first-order state does not exist. Certainly, a nonexistent state cannot be conscious. How can the higher-order view meet this challenge?

Rosenthal's reply (2011) is to point out that even when the relevant first-order state is *actually* missing, the higher-order state nonetheless represents oneself as being in a certain first-order state, which just happens to be nonexistent. Therefore, according to the transitivity principle, the nonexistent first-order state is phenomenally conscious; the relevant conscious experience is determined by the content of the higher-order state, and that state represents one as being in a (nonexistent) first-order state, so what it is like for the subject will be like being in the first-order state even when he or she is in fact not in that first-order state. This may sound odd but is perhaps not so if we consider that nonexistent objects in general have all sorts of properties. A nonexistent communication can have representational content: for example, I thought you had written a complaint letter (even though you did not) about a new university policy (because we had talked about complaining about it, but unbeknownst to me you never actually wrote that letter). So likewise, *perhaps* it is not problematical at all to speak of conscious experiences arising because of our being conscious of such nonexistent first-order states. It becomes less problematical when we realize that all that is meant by saying the nonexistent state is conscious is that the state is described in the content of the higher-order state.

We emphasize "perhaps" because in our experience we find that some readers find it insurmountably odd and downright unacceptable to say that a (nonexistent) first-order state is phenomenally conscious. On the other hand, some readers find it perfectly sensible.

Instead of becoming bogged down arguing over this, we offer an alternative solution to accommodate the Empirical Cases within a higher-order framework, under the full-conscious-experience interpretation. Specifically, Brown (2012b, 2015) has argued that phenomenal consciousness consists in *implementing* the transitivity principle, which is to say that it consists in the occurrence of the higher-order state itself. On this alternative, phenomenal consciousness is just a higher-order representation. Thus, for a first-order state to be (state) conscious, one does need to represent oneself as being in that first-order state, and we can happily say that the (as it happens) nonexistent first-order state is state conscious in this sense (or that it is the content of the higher-order state). If we think of phenomenal consciousness as the property of there being something that it is like for the subject of the experience, then it is the higher-order state that has that property. It is that state that is like something for the subject to be in; without it there is nothing that it is like for the subject. This is equivalent to saying that the first-order state is never *phenomenally conscious*. Phenomenal consciousness is just having the appropriate higher-order state.[12]

This is not to give up the transitivity principle. The transitivity principle, as we construe it, says that phenomenal consciousness consists in being aware of oneself as being in some first-order state. It is sometimes thought that the transitivity principle relies on there being a relation between the first-order state and the higher-order state. We explore a relational version of the transitivity principle in the next section, but it is not the case that the transitivity principle *must* be interpreted as involving a relation between a first-order state and a higher-order state. On this alternative nonrelational view of the transitivity principle, phenomenal consciousness relies on the instantiation of a specific kind of awareness. On this construal, phenomenal consciousness is just being aware of oneself as being in a first-order state. This is what the ambitious higher-order theory under the full-conscious-experience interpretation should be interpreted as saying. This solves Block's puzzle. There is phenomenology in the empty case, and it consists in having the appropriate higher-order state.

We grant that it may seem counterintuitive that the redness of my conscious experience is nothing more than the occurrence of an appropriate thought-like representation to the effect that I am seeing red. It certainly doesn't seem to be that way when one has the conscious experience! But introspection cannot be reliable here, since it is the very nature of the higher-order representation to make it seem to us as though we are in fact in the first-order state. Thus, we would naturally expect it to be the case that it doesn't seem to us as though we are having a thought. It will seem to us as though we are seeing red.

In our view either of the responses is adequate as a response to the Empty Higher-Order problem, and we do not need to endorse one over the other. Our aim, rather, is to show that despite Block's contention that the higher-order approach has a problem accounting for Empty Higher-Order Representations, the reverse is true. The higher-order approach has ready responses; indeed there may be others that we have not yet encountered. In fact, as we have tried to show, it is the first-order view that really has trouble accounting for Empty Higher-Order cases.

The Partial-Conscious-Experience Interpretation

The previous two interpretations are both extremes. One possible intermediate position holds there is conscious experience in the Empirical Cases, but such conscious experience is not full-fledged as it would have been had the first-order representations been intact. That is, there is some reduced form of conscious experience in the Empirical Cases that is qualitatively different from what one expects in normal cases.

The higher-order view can certainly entertain this interpretation as much as it does for the Full-Conscious-Experience Interpretation. All it takes is to stipulate that in the Empirical Cases, because of the weak or missing first-order representation, the higher-order representation contains less detailed perceptual information.

On the other hand, the first-order view seems to face the same problem as it does in the Full-Conscious-Experience Interpretation. If there is even just a spark of phenomenology that is not explainable in terms of the missing or too-weak first-order representations, it violates the view.

One can perhaps imagine a response from a first-order theorist based on the strategy of divide and conquer. For example, one can deny the reported conscious experience in the Inattentional Inflation case and then argue that in the case of Peripheral Vision first-order representations are created via top-down mechanisms or found in extrastriate areas, despite the lack of sufficient retinal input.[13] And finally for the Rare Charles Bonnet Cases, one can argue that although the primary visual cortex (V1) is missing, activity in the remaining visual cortex is nonetheless sufficient for a partial conscious experience. Though we suspect this may well be the reply Block would favor, we note that this strategy would involve a fair amount of patchwork—one needs to deny the reported conscious experience in Inattentional Inflation—that is, essentially denying that a change in detection bias (subjects saying yes more frequently to target detection) *and* an increase in visibility rating in a discrimination task *together* do not constitute a reliable reflection of a change in conscious experience. While denying this single case is probably not an unarguable position, this seems ad hoc, and goes against the spirit of finding a simple coherent interpretation for *all available* evidence (Block 2007). And for Peripheral Vision, one needs to substantiate or at least commit to predicting the existence of the putative top-down mechanisms for creating those detailed first-order representations. And finally, one also needs to abandon the feedback-to-V1 view for first-order representations for conscious experience to deal with the Rare Charles Bonnet Cases.

What we find relatively more intriguing is the possibility that, under this Partial-Conscious-Experience Interpretation, an intermediate view captures some of the flavor of both the first-order and higher-order views. The basic idea behind this option is that conscious experience perhaps *jointly depends* on both higher-order and first-order representations. On this view, we cannot tell what it is like for a subject just by looking at the first-order representations, nor can we tell what it is like for the subject just by looking at the higher-order representations. It is the combination of the first- and higher-order representations that jointly determines the qualitative character of conscious experiences.

What does it mean to say that conscious experience jointly depends on both higher-order and first-order states? Note, first, that we are not claiming that the first-order state plays some indirect causal role in determining the final conscious experience, because higher-order theories usually allow first-order states to be causally relevant in normal circumstances.[14] Here we are discussing a different view, that the first-order state partially *constitutes* the conscious experience.

A simple analogy might help. In classical Newtonian physics, an object's acceleration depends jointly on the force applied to it and its mass. The rate of acceleration

of an object cannot be calculated from the force alone. So too, consciously seeing red may jointly depend on both higher-order and first-order states. This kind of Joint-Determination relationship may be quite general in nature and not necessarily ad hoc. But an analogy is not specific enough. Below we explore what exactly this view concerning the basis of conscious experience could be.

Others have argued for such a view as well, including Uriah Kriegel (2003, 2006, 2009). However, here we focus on a version of this Joint-Determination view that is an extension of Lau (2008). On this view, the higher-order representation *refers* to the relevant first-order representation for the specific content, and together the two states determine the exact nature of what it is like for the subject. For example, a higher-order state may represent something like "I am vividly perceiving the content of first-order representation F," where the first-order representation F may have red* as its content. The intensity of the conscious experience is determined by the higher-order representation (i.e., vividly perceiving, as opposed to having some faint and uncertain impression), but ultimately, the detailed content (e.g., of color) is determined by the first-order states, by virtue of it being referred to by the higher-order state. So in the normal case of consciously seeing red, a higher-order state exists to the effect that one is seeing some determinate shade of red as specified by some first-order state and the relevant first-order state represents the specific shade of red. When the relevant first-order state is missing, the phenomenology would be different: subjects would still experience seeing color but without any specific color consciously experienced. That is, the subjects are confident that they have perceived the color of the relevant object, except that in trying to name the exact color, they may fail, because the first-order representations are missing. In this case the subject will experience what we can call fake phenomenology—having a conscious experience of a determinate color without it being of any determinate color. This may seem odd at first, but we suspect everyday peripheral vision is very much often like this. Not only do we think we see color in the periphery, but there is a phenomenological experience of perceiving color. However, upon more careful introspection or under rigorous laboratory testing, it seems that we do not actually experience any determinate color in the periphery. Thus, each state contributes something to the overall phenomenology of normal conscious experience.

The appeal of the Joint-Determination view is that it retains some of the main motivations for both the first-order and higher-order views. In a sense, this is essentially a higher-order approach, following the transitivity principle. That is, the first-order representation F (as in the example above) gets to contribute to the qualitative character of conscious experience only because one represents oneself as being in F. F on its own does not give rise to conscious experience; it is only when F is targeted by an appropriate higher-order representation that there is any phenomenology at all. A mere change in higher-order representation, keeping F constant, can lead to a change in conscious experience. However, a critical difference here is that on this Joint-Determination view

the perceptual content is not being duplicated in the higher-order system. Following the standard higher-order view, in a sense, the Joint-Determination view also holds that the qualitative character of conscious experience is determined by the content of the higher-order representation. However, it is determined by not only the narrow content but also the broad content.[15] That is, whether the higher-order content is veridical would matter; if F does not exist, rendering the higher-order state nonveridical, the qualitative experience would be different. Thus, the details of the qualitative experience ultimately come from the content of the first-order states to which the relevant higher-order state refers. This way, one also preserves the empirical intuition that the qualitative details of conscious experience may be too fine grained to be represented by the higher-order system.

The Joint-Determination view would allow us to say that in the Empirical Cases there is some conscious experience (because of the presumed existence of the relevant higher-order states), but such experience is not full blown; in other words, it is qualitatively different from normal cases (because of the absence or the low representational quality of the first-order states). For instance, in the case of Inattentional Inflation, the first-order state may be constant between the attended and unattended case, but the higher-order state under inattention may represent one as having a more reliable and intense perceptual experience than under attention.

Note that unlike the standard higher-order view, the Joint-Determination view cannot allow full-blown conscious experience to occur if there are no first-order representations at all. If the higher-order state represents oneself as vividly seeing an object with shape as specified by F1, with color as specified by F2, and so on, and if it turns out that the relevant first-order states (F1, F2, etc.) are actually completely missing, one should just experience a sense of seeing something without being able to say what that something is. To be more precise, it is not just that subjects are unable to *say* it; the conscious experience itself also lacks the specific content of what it is that they are experiencing. While this may seem odd, once again we note that perhaps this happens not infrequently in everyday peripheral vision. However, this is certainly different from what is reported in Rare Charles Bonnet Cases, in which the patients claim to see vivid objects and are able to name them precisely. Therefore, the interpretation based on this Joint-Determination view is that in Rare Charles Bonnet Cases there is some impoverished first-order representation, despite the primary visual cortex being damaged. This is possible because one can hold the Joint-Determination view without holding the view that first-order representations critically depend on (feedback to) the primary visual cortex. One can hold that the first-order representation in Rare Charles Bonnet Cases is perhaps impoverished but not nonexistent, thus the representations of the colors and shapes of the hallucinated objects may be less precise and distinct. However, the precision and distinctness of the percept may well be subjectively inflated because

of the idiosyncratic nature of the higher-order representations in these cases. In other words, committing to this intermediate Joint-Determination view would also involve rejecting the feedback-to-V1 view for first-order representations.

Regarding Block's Empty Higher-Order Representation argument, when the first-order representations are completely missing (which is not necessarily the case in the Empirical Cases, as noted earlier), there is still some conscious experience, albeit impoverished Thus, the higher-order representations play an important role. However, such conscious experience would lack specific content. Usually, when one consciously perceives an object (or any specific content), it is natural to speak of the first-order mental representation of the object as being conscious. However, here in true empty higher-order cases, there is no specific content to speak of, and therefore, no first-order mental states are conscious. In this sense, prima facie, the transitivity principle—that a first-order state is conscious only when we are conscious of having that state—is not violated. There is no such first-order state in this case, and the theory does not claim that such state is conscious.

The Joint-Determination view is clearly different from the first-order view because on the former, conscious experience arises only when the relevant higher-order states exist. However, one may wonder whether it is it truly distinct from the higher-order view. As noted earlier, in a sense, the qualitative character of the conscious experience is determined entirely by the content of the higher-order state—if we consider the *broad, externalist* content, including what it refers to and whether it is veridical. In this case this broad content would be the relevant first-order representation. One may also wonder, when the first-order state is completely missing, why there is conscious experience (albeit nonspecific). Does one not need to assume then that phenomenal consciousness is a property of higher-order states? Is this then not the same reply as the one by Brown (2015) in defending the higher-order view (under the Full-Conscious-Experience Interpretation)? The answer is, true enough, in a limited sense, by taking this Joint-Determination view, one is probably conceding that the conscious sense of perceiving something is essentially driven by the higher-order but not the first-order state. But an important difference here is that one need not claim that first-order states are never phenomenally conscious. When a first-order perceptual state is conscious, it is phenomenally conscious. It is just that even when no first-order states are phenomenally conscious, one can still have nonspecific conscious experiences due to the higher-order states.

One disadvantage of the Joint-Determination view is that it gives up the explanatory power that is a main motivation for accepting the more traditional version of the higher-order approach. Consider a normal case of consciously experiencing a pain. On the higher-order thought theory, one is in a higher-order thought-like state that represents oneself as being in pain. This is why it is painful for you to be in this state,

because the higher-order state deploys the intentional concept of pain, and this is what accounts for it seeming, from your point of view, that you are in pain. On the Joint-Determination view, we seem to lose that explanation. It is not clear how a first-order state's being referred to can account for the conscious experience. What does this higher-order referring do? With only a mere reference to the first-order state, unlike the higher-order thought that contains the relevant intentional concept, it seems relatively difficult to construct the kind of naturalistic explanation of consciousness that some higher-order theorists aim for.

Another criticism could be that if, ultimately, it is the higher-order state (which represents oneself as being in a particular first-order state) that *leads* to the conscious experience, why does it have to indirectly refer to the content of the first-order state? Why cannot the higher-order state represent the color content itself? In other words, what is the motivation for holding this Joint-Determination view? If it follows from the transitivity principle, isn't the ambitious higher-order view described in the last section more straightforward?

However, we maintain that, these potential drawbacks aside, we should not reject the Joint-Determination view offhand, because there may be advantages that outweigh the disadvantages. For instance, because the higher-order state, on the Joint-Determination view, does not duplicate the perceptual content from the first-order state, it avoids a possible outright mismatch in content, such as a higher-order state representing oneself as seeing red, but the first-order state representing greenness. As noted at the chapter's beginning, it has been argued that the possibility of mismatch poses a challenge to the higher-order view. But the Joint-Determination view bypasses these problems. Higher-order theorists have offered other replies to this challenge of mismatch; however, we do not think that bypassing the mismatch problem alone makes the Joint-Determination view superior.

We suspect that the strongest motivation for the Joint-Determination view may be empirical. Recall the neuroanatomical interpretation adopted throughout this chapter: higher-order representations depend on activity in the prefrontal and parietal cortices, and first-order representations depend on activity in the early visual areas. It may be most plausible that the qualitative character of a conscious experience cannot be determined fully by the activity in the prefrontal cortex alone, as representations in the prefrontal cortex may not have the fineness of grain to capture the richness of the perceptual content in conscious experience. This is an open empirical question, and we believe that no a priori theorizing can settle the matter at this point.[16] Having explanatory power is one kind of theoretical virtue, but so too is fitting the data. At present we are unable to settle this issue and are happy to leave disputes between Joint-Determination and standard higher-order views at the mercy of these future empirical results. In other words, we do not argue strongly for the Joint-Determination view, and

we note the potential disadvantages. We only describe it here as a possibility that future work may explore further.

Resisting the Trilemma?

The foregoing discussion is presented as considering three different interpretations of the Empirical Cases. However, it can also be seen as an argument against the first-order view and in support of the higher-order approach. The main crux of the argument roughly boils down to the following:

(1) There is either conscious experience or not (No-Conscious-Experience Interpretation) in the Empirical Cases.

(2) It is extremely implausible to deny that there is conscious experience in the Empirical Cases.

(3) If there is conscious experience in the Empirical Cases, it is either like normal conscious experiences (Full-Conscious-Experience Interpretation) or it is not full-fledged (i.e., impoverished and nonspecific) (Partial-Conscious-Experience Interpretation).

(4) If the conscious experience in the Empirical Cases is like normal conscious experience (Full-Conscious-Experience Interpretation), the higher-order theory is more plausible than the first-order theory.

(5) If the conscious experience in the Empirical Cases is not like normal conscious experience but is, rather, impoverished and nonspecific (Partial-Conscious-Experience Interpretation), then either the higher-order theory is true or conscious experience is jointly determined by first-order and higher-order states.

We take it that points 1–3 are not so controversial. Point 4 depends on some loose ends; namely, should the higher-order theorist accept that it is fine to speak of nonexistent first-order states being conscious (Rosenthal 2011) or accept that first-order states are never phenomenally conscious (Brown 2015)? There may yet be other solutions. But what we maintain is that, under the Full-Conscious-Experience Interpretation, the first-order theory does seem to be in much more trouble than the higher-order theory.

We suspect point 5 is probably the most controversial. As pointed out in the last section, it is not impossible that a first-order theorist may find ways to defend a modified first-order view by taking the Partial-Conscious-Experience Interpretation. We do not find such defense plausible and note that, at the very least, the first-order theorist would have to give up the feedback-to-V1 view that Block endorses. And likewise, if one accepts the Joint-Determination view, one also has to give up the feedback-to-V1 view for first-order representation.

Conclusion

Just as in Hans Christian Andersen's fable in which the emperor's thinking that he has gotten new clothes does not give him real clothes, merely thinking of being in a conscious perceptual state should not give one real conscious experience. However, taking this analogy literally would be a disservice to the higher-order thought theory, because one should not confuse normal conscious thinking (as in the case of the emperor) with the specific higher-order representations stipulated by higher-order approaches. According to the theory, such higher-order states, unlike normal conscious thoughts, do not have to be conscious themselves, unless they become the content of yet another higher-order representation. Opponents may argue that unconscious thought leading to conscious perception is odd. However, being odd is not the same as being incorrect. As Block himself has suggested (2009), something counterintuitive may well be needed to account for the infamous explanatory gap (Levine 2001). We recognize the counterintuitiveness of some interpretations of the standard higher-order approach, but if our best scientific theories point in that direction, we should be prepared to accept the conclusion. To those who cannot take this oddness straight up, we offer the Joint-Determination view as an alternative to help preserve some of the intuition of the first-order view.

We offer the following four points as a summary of what we suggest in the chapter's argument.

(I) Recurrent feedback loops to V1 have been hypothesized to be the neural correlate of conscious visual experience, but the Empirical Cases we appeal to suggest that this is most likely not the case. This is true under the Full-Conscious-Experience Interpretation as well as the Partial-Conscious-Experience Interpretation. Thus, to save the feedback-to-V1 view the only option is to take the No-Conscious-Experience Interpretation, which we argue to be so implausible as to be outlandish.

(II) The first-order view faces serious challenges regarding these empirical cases. This is because it is very unappealing to take the No-Conscious-Experience Interpretation, which would have been compatible with the first-order view. It is unclear how the first-order view can cope with the Full-Conscious-Experience Interpretation. On the Partial-Conscious-Experience Interpretation, we argue that there may be some hope for defending some core elements of the first-order view, but the first-order theorist will also need to make some important adjustments and sacrifices.

(III) In light of the preceding two points, we argue that a version of the standard higher-order approach should be considered less problematical than the first-order view. We do not argue that the higher-order approach faces no difficulty whatsoever. The challenge seems to be that, under the Full-Conscious-Experience Interpretation (or the Partial-Conscious-Experience Interpretation), one needs to decide how to preserve the transitivity principle. Such solutions are not straightforwardly intuitive to

everybody, but we argue that their problems pale in comparison to the problems facing the first-order view.

(IV) A new alternative we see is to adopt the Joint-Determination view about conscious experience, which holds that phenomenology is jointly determined by both first-order and higher-order states. This view may have its problems, but also potential merits, and may well be the happy medium where first-order theorists can preserve some core intuitions of their theory amid the present challenges.

Among these points, we feel most strongly about point (I): the recurrent-feedback-loop view has other problems and, as we noted earlier, has been independently criticized (Macknik and Martinez-Conde 2008; Silvanto and Rees 2011). Block (2007, 499, note 10) has also noted previously that his view is not committed to this empirical claim. Though point (II) seems to us hard to deny, we are somewhat less sure about (III) and (IV), in the sense that we feel somewhat ambivalent as to whether the Joint-Determination view is truly a good alternative—and to arbitrate between them would be to go beyond the scope of this chapter. Ultimately, it may depend on future empirical evidence. However, in any case, although the Joint-Determination view has some flavor of preserving some important intuitions behind the first-order view, we note that to take this option is likely to involve giving in quite a bit to the higher-order approach. This is not just the modest version of the theory that aims to explain what makes a conscious mental state one that one is aware of. This is about the very nature of phenomenal consciousness. The position holds that the higher-order state is an essential component of phenomenal conscious experience and that it partially determines the overall nature of what it is like to have a conscious experience. Taking this option is to concede a lot to the higher-order theorist.

Thus, rather than being "defunct" (Block 2011a), the higher-order approach to consciousness is alive and well indeed.

Notes

1. We use the phrase "mental representation" interchangeably with "mental state" or sometimes just "representation" or "state." Also, unless otherwise specified, we always use the term "conscious experience" to refer to *phenomenally* conscious experience. Other similar and related terms are "conscious phenomenology," "conscious perception," "what-it's-likeness," and "phenomenally conscious state."

2. When we talk about the higher-order approach to consciousness or higher-order view, we always mean what Block calls the ambitious version of the theory as opposed to the modest version of the theory (Block 2011a). That is, in this chapter we always take it to be a theory of phenomenal consciousness as opposed to merely a theory of state consciousness. Because of this some theorists may reject the way that we have formulated the transitivity principle (see, e.g., Rosenthal 2011, 435).

3. In brief, the argument for this is that the first-order states do have a large functional role to play and that they do account for most of our performance. It is just with respect to conscious experience that they have an indirect role.

4. Rosenthal technically holds that there are two aspects to a typical conscious experience. On the one hand is the higher-order state, which accounts for what it is like for the individual to have the experience. And on the other hand the individual has the first-order qualitative state that accounts for the functioning and perceptual role. So in the empty case we do have something odd going on, but there is no difference in conscious phenomenology (though there will be a difference in performance).

5. On our interpretation, we can call Ned Block a first-order vehicle theorist, who holds that having the right kind of biological substrate for the first-order representation is partly necessary for consciousness.

6. We also assume an identity between neural activity in early sensory areas and representations of, say, red. Thus, we can switch between talking about neural activity, spiking, and other activity and representations of, for example, red, blue, or pain.

7. Unfortunately, the term "early sensory regions" is often not technically delineated, even in the neuroscience literature. Here, concerning the visual modality, we use it to refer to primary visual cortex (i.e., striate cortex, also known as V1), extrastriate areas (including V2, V3, V4, MT), and other areas in the occipital and temporal lobes that are known to contain a high number of neurons explicitly coding for visual objects (e.g., fusiform face area). All other areas are considered higher-order, but we typically consider such higher-order areas to be in the lateral prefrontal cortex and around the intraparietal areas.

8. In the philosophy literature, sometimes the word "hallucination" is used when a first-order representation refers to a nonexistent object. Here we are not restricted by this usage. By "hallucination" we simply mean nonveridical conscious experiences, regardless of whether it is driven by a first-order or higher-order representation.

9. Also relevant are cases in which, through intracranial magnetic brain stimulation, conscious motion percept can be induced in a patient whose spatially relevant part of the primary visual cortex is damaged (Silvanto et al. 2007). We do not consider these cases in more detail here because they have been discussed at length elsewhere (Silvanto and Rees 2011). But we note that they corroborate with the conclusion from the Rare Charles Bonnet Cases here that conscious visual experience can occur in the absence of the primary visual cortex.

10. This is of course not to say that our experience of the world is always right and always veridical. We may hallucinate. But the point is that, even in hallucinations, there is genuine conscious experience. We cannot argue against such conscious experience, because the object of perception may not actually exist. Thus, from a third person's point of view, it is hard to argue for or against the existence of a genuine conscious experience without some kind of appeal to the first-person reports.

11. Indeed, Block himself acknowledges this in "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience" (2007; see esp. 499, note 10).

12. This may not be all that different from the way that Weisberg (2011) should be taken in his response to Block. He says, while talking about a thought experiment in which future neuroscientists have stimulated just the higher-order regions and elicited a conscious experience of pain in the subject, that

> "...in this case the lower-order state does not exist. What is to be said? Perhaps the thing to say is that the higher-order state itself is the conscious state. And there's a certain meaningful sense of 'conscious state' in which that is correct—the higher-order state is responsible for there being something it's like for the subject" (Weisberg 2011, 442).

13. One could also take the view that the representations in Peripheral Vision are indeterminate. (See, e.g., Block 2010, 52–53.)

14. On Rosenthal's account, the first-order state is needed to acquire the concept used in the higher-order state, but once the concept is acquired, it is the higher-order state that determines what the experience is like. So in a way the overall phenomenology does depend on the first-order state in that it is necessary to have the concept, and so its distinctive characteristics will be captured in the concept deployed to represent it. The overall phenomenology is thus jointly determined by the higher-order state and the characteristics of the first-order state, though not the actual first-order state that may have been targeted. However, this is not the idea of joint determination that we have in mind. Rosenthal's account seems to be no more than causal connection, like the kind that holds between states of the retina of visual representations.

15. We adopt the terms "narrow content" and "broad content" here merely for illustrating this point. We are not committed to theories regarding these.

16. We note that there is another wrinkle here. Rosenthal argues (2005) that the content of the higher-order states is essentially comparative. Thus, instead of needing to represent every specific color that we encounter in the world, we need only represent their characteristic similarities and differences. If this is so, then much less processing power would be required by the prefrontal cortex. We cannot decisively resolve this issue currently, since we do not really understand how the brain encodes this information.

References

Ashwin, P., and M. Tsaloumas. 2007. Complex visual hallucinations (Charles Bonnet syndrome) in the hemianopic visual field following occipital infarction. *Journal of the Neurological Sciences* 263: 184–186.

Azzopardi, P., and A. Cowey. 1993. Preferential representation of the fovea in the primary visual vortex. *Nature* 361: 719–721.

Balog, K. 2000. Phenomenality and higher order thought. *Consciousness and Cognition* 9: 215–219.

Block, N. 2005. Two neural correlates of consciousness. *Trends in Cognitive Sciences* 9: 46–52.

Block, N. 2007. Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences* 30: 481–548.

Block, N. 2008. Consciousness and cognitive access. *Proceedings of the Aristotelian Society* 108: 289–317.

Block, N. 2009. Comparing the major theories of consciousness. In *The Cognitive Neurosciences*, 4th ed., ed. M. Gazzaniga, 1111–1122. Cambridge, MA: MIT Press.

Block, N. 2010. Attention and mental paint. *Philosophical Issues* 20: 23–63.

Block, N. 2011a. The higher-order approach to consciousness is defunct. *Analysis* 71: 419–431.

Block, N. 2011b. Response to Rosenthal and Weisberg. *Analysis* 71: 443–448.

Block, N. 2011c. Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences* 15: 567–575.

Brown, R. 2012a. The myth of phenomenological overflow. *Consciousness and Cognition* 21: 599–604.

Brown, R. 2012b. The brain and its states. In *Being in Time: Dynamical Models of Phenomenal Experience*, ed. S. Edelman, T. Fekete, and N. Zach, 211–238. Advances in Consciousness Research. Amsterdam: John Benjamins Press.

Brown, R. 2014. Consciousness doesn't overflow cognition. *Frontiers in Psychology* 5: 1399.

Brown, R. 2015. The HOROR theory of phenomenal consciousness. *Philosophical Studies* 172: 1783–1794.

Carrasco, M. 2011. Visual attention: The past 25 years. *Vision Research* 51: 1484–1525.

Carrasco, M., S. Ling, and S. Read. 2004. Attention alters appearance. *Nature Neuroscience* 7: 308–313.

Contardi, S., G. Rubboli, M. Giulioni, R. Michelucci, F. Pizza, E. Gardella, F. Pinardi, I. Bartolomei, and C. Tassinari. 2007. Charles Bonnet syndrome in hemianopia, following antero-mesial temporal lobectomy for drug-resistant epilepsy. *Epileptic Disorder* 9: 271–275.

Dennett, D. 1976. Are dreams experiences? *Philosophical Review* 73: 151–171.

Duggal, H., and J. Pierri. 2002. Charles Bonnet syndrome: Neurobiological insights. *Indian Journal Psychiatry* 44: 289–292.

Ffytche, D. 2005. Visual hallucinations and the Charles Bonnet syndrome. *Current Psychiatry Reports* 7: 168–179.

Gorea, A., and D. Sagi. 2000. Failure to handle more than one internal representation in visual detection tasks. *Proceedings of the National Academy of Sciences* 97: 12380–12384.

Green, D., and J. Swets. 1966. *Signal Detection Theory and Psychophysics*. New York: John Wiley and Sons.

Kriegel, U. 2003. Consciousness, higher-order content, and the individuation of vehicles. *Synthèse* 134, no. 3: 477–504.

Kriegel, U. 2006. The same-order monitoring theory of consciousness. In *Self-Representational Approaches to Consciousness*, ed. U. Kriegel and K. Williford, 143–170. Cambridge, MA: MIT Press.

Kriegel, U. 2009. *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford University Press.

Lamme V. 2006. Towards a true neural stance on consciousness. *Trends in Cognitive Sciences* 10: 494–501.

Lau, H. 2008. A higher-order Bayesian decision theory of perceptual consciousness. In *Progress in Brain Research*, ed. R. Banerjee and B. Chakrabarti, 168: 35–48.

Lau, H., and D. Rosenthal. 2011. Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences* 15 (8): 365–373.

Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.

Macknik, S., and S. Martinez-Conde. 2008. The role of feedback in visual masking and visual processing. *Advances in Cognitive Psychology* 3: 125–152.

Neander, K. 1998. The division of phenomenal labor: A problem for representational theories of consciousness. *Philosophical Perspectives* 12: 411–434.

Newton, J., and R. Eskew Jr. 2003. Chromatic detection and discrimination in the periphery: A postreceptoral loss of color sensitivity. *Visual Neuroscience* 20: 511–521.

Prinz, J. 2005. A neurofunctional theory of consciousness. In *Cognition and the Brain: The Philosophy and Neuroscience Movement*, ed. A. Brook and K. Akins, 243–259. Cambridge: Cambridge University Press.

Rahnev, D., L. Bahdo, F. de Lange, and H. Lau. 2012. Pre-stimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception. *Journal of Neurophysiology* 108: 1529–1536.

Rahnev, D., B. Maniscalco, T. Graves, E. Huang, F. de Lange, and H. Lau. 2011. Attention induces conservative subjective biases in visual perception. *Nature Neuroscience* 14: 1513–1515.

Rosenthal, D. 2005. *Consciousness and Mind*. New York: Oxford University Press.

Rosenthal, D. 2011. Exaggerated reports: Reply to Block. *Analysis* 71: 431–437.

Silvanto, J., A. Cowey, N. Lavie, and V. Walsh. 2007. Making the blindsighted see. *Neuropsychologia* 45: 3346–3350.

Silvanto, J., and G. Rees. 2011. What does neural plasticity tell us about role of primary visual cortex (V1) in visual awareness? *Frontiers of Psychology* 2: 6.

Sperling, G. 1960. The information available in brief visual presentations. *Psychological Monographs* 74: 1–29.

Weisberg, J. 2011. Abusing the notion of what-it's-like-ness. *Analysis* 71: 438–443.