Synthese
https://doi.org/10.1007/s11229-020-02613-3

S.I.: MATERIALISM & METAPHYSICS

Consciousness and information integration

Berit Brogaard[1,2] · Bartek Chomanski[1] · Dimitria Electra Gatzia[3]

Received: 5 October 2018 / Accepted: 27 February 2020

Abstract

Integration information theories posit that the integration of information is necessary and/or sufficient for consciousness. In this paper, we focus on three of the most prominent information integration theories: Information Integration Theory (IIT), Global Workspace Theory (GWT), and Attended Intermediate-Level Theory (AIR). We begin by explicating each theory and key concepts they utilize (e.g., information, integration, etc.). We then argue that the current evidence indicates that the integration of information (as specified by each of the theories) is neither necessary nor sufficient for consciousness. Unlike GWT and AIR, IIT maintains that conscious experience is both necessary and sufficient for consciousness. We present empirical evidence indicating that simple features are experienced in the absence of feature integration and argue that it challenges IIT's necessity claim. In addition, we challenge IIT's sufficiency claim by presenting evidence from hemineglect cases and amodal completion indicating that contents may be integrated and yet fail to give rise to subjective experience. Moreover, we present empirical evidence from subjects with frontal lesions who are unable to carry out simple instructions (despite appearing to understand their meaning) and argue that they are irreconcilable with GWT. Lastly, we argue that empirical evidence indicating that patients with visual agnosia fail to identify objects they report being conscious of present a challenge to AIR's necessity claim.

Keywords Amodal completion · Attended intermediate-level representation theory · Attention · Consciousness · Feature integration · Global workspace theory · Illusory contours · Information integration theory · Spatial neglect · Visual agnosia

---

The names appear in alphabetical order: each author contributed equally.

✉ Dimitria Electra Gatzia
degatzia@uakron.edu

Extended author information available on the last page of the article

1 Introduction

A growing number of philosophers and cognitive scientists have recently argued that informational integration should play a central role in our attempts to understand conscious experience. Various theories have been proposed to explain how information is integrated in such a way as to produce conscious experiences, such as seeing a blue square. In this paper, we will present evidence from a range of empirical studies that casts doubt on the strength of the explanatory role information integration plays in the consciousness debate. After discussing the relevant notions of consciousness and information integration, we outline some of the most prominent theories of consciousness that give center stage to the notion of information integration. We then present evidence for consciousness in the absence of integration, and vice versa.

2 Representation, consciousness, and information integration

We assume that mental states are representational: that is, as a minimum, they carry information about real-world properties and relations. When a mental state represents a square shape as blue, for example, it carries information about blueness and squareness. These properties are components of the mental state's representational content.[1]

Human beings and other creatures are conscious when they are aware of themselves or their environment. This sort of awareness, also known as 'creature consciousness,' can change gradually when, say, we are falling asleep or are getting intoxicated. It can also change more rapidly when, for instance, a boxer gets knocked out in the ring. Creature consciousness is commonly taken to be distinct from state consciousness (Dretske 1993; Carruthers 1996, 2000; Lycan 1996; Block 1997; Rosenthal 1997; Manson 2000).

State consciousness is a property of mental states that is closely tied to what the content of the mental state makes its bearer aware of. An individual who is (seemingly) aware of a property *P* is in a conscious mental state that has a content that contains *P* (or a component that refers to *P*). Although not universally accepted, we shall here understand creature consciousness in terms of state consciousness as follows:

Individual *S* is creature conscious just in case some mental state of *S*'s is conscious.

In this paper, we will mostly be concerned with consciousness as a property of mental states. Consciousness is viewed as a great scientific challenge, in part, because it is a phenomenon that involves subjective experiences that have a certain phenomenology or qualitative 'feel' (Chalmers 1998; Bayne and Chalmers 2003). For example, experiencing a blue square has a different phenomenology or qualitative 'feel' than experiencing a red circle. Having such experiences 'feels like something' from a first-person point of view. Any theory of consciousness should be able to explain why the underlying anatomical, functional, or informational structures (or connectivity) give

[1] A theory of consciousness should be able to explain the neural substrates of mental states and their behavioral manifestations.

rise to subjective experiences that have a certain phenomenology. Indeed, the aim of theories we discuss in this paper is to provide this sort of explanation.

With respect to subjective experience, we take integration to be the bringing together, or combination, of sensory information from different sources. For example, given a normally functioning human perceptual system, information about properties such as shape, color, motion, and spatial location may be combined to yield an integrated percept of, say, a blue square moving from left to right. When the subject is in a mental state whose representational content conveys integrated information about a range of properties, that state can then be said to have integrated content. Thus understood, integration is a property of the contents of mental states.

With respect to theories of consciousness, the aim is to explain why and how certain physical mechanisms give rise to conscious experiences. Some theoretical approaches to consciousness that rely on information integration begin by identifying the essential properties of subjective experience and then ask what sorts of physical mechanisms would give rise to them (see e.g., Tononi 2004). Others start from the brain and then ask how it could give rise to subjective experiences (see e.g., Baars 1988, 2005). As we shall see, the notion of 'information integration' is used differently by proponents of Information Integration Theories. For example, according to Giulio Tononi's axiom of information (2012, p. 297), each experience is specific; and according to the postulate of information (posited to explain the above axiom) information is a "cause-effect repertoire" (defined as "the probability distribution of potential past and future states of a system as constraint by a mechanism in its current state") which is integrated, i.e., it is not reducible to its parts. For Baars and Franklin (2003), it is integrated sensory information that enters the global workspace, which, when spotlighted, leads to conscious experiences; and for Prinz (2012), it is the integration of intermediate-level representations that serve as the contents of conscious experience. Although each of these theories utilize different notions of information integration, what they all have in common is that they tie the notion of *information integration* to subjective experience. This is not surprising given that the aim of such theories of consciousness is to explain subjective experiences in terms of the sort of information integration that can give rise to them.

The general thesis that information integration and consciousness are deeply intertwined can be given a number of more specific interpretations as follows:

(A) *The Identity Thesis:* for a mental state to be conscious just is for it to have integrated informational content.
(B) *The Equivalence Thesis:* having an integrated informational content is necessary and sufficient for consciousness.
(C) *The Necessity Thesis:* having an integrated informational content is necessary for consciousness.
(D) *The Sufficiency Thesis:* having an integrated informational content is sufficient for consciousness.

Once again, it is important to keep in mind that each of the theories we discuss specifies "integrated informational content" in its own distinctive way; here, it is used as a mere placeholder. A and B imply C and D. We will argue that the current empirical data indicate that consciousness can occur in the absence of informational integration, and that a mental state can have an integrated content without being conscious. Hence, we will argue that theses (A) through (D) are false.

(C) and (D) can be cashed out as follows:

(C) *Necessity of Integration*: Necessarily, if mental state M is conscious, then M's informational content is integrated ($C \Rightarrow I$).
(D) *Sufficiency of Integration*: Necessarily, if mental state M's informational content is integrated, then M is conscious ($I \Rightarrow C$).

We shall here limit ourselves to the following three theories: Information Integration Theory (IIT), Global Workspace Theory (GWT), and Attended Intermediate-level Representation Theory (AIR). All three theories hold that integration of the informational content of experience is necessary for phenomenal consciousness. In other words, they are all committed to (C). IIT is furthermore committed to the sufficiency of integration, as expressed by (D).

As we will argue below, there is evidence to suggest that subjects with frontal lesions are unable to carry out simple instructions, despite appearing to understand the meaning conveyed by such instructions. This, we argue, is irreconcilable with GWT. While IIT arguably has the resources to accommodate these results, it appears to fall prey to empirical evidence demonstrating the possibility of the experience of simple features without feature integration. Moreover, by claiming that integration of informational content is not only necessary but also sufficient for consciousness, IIT tempts fate twice. As we will see, in hemineglect cases, content may be integrated and yet fail to give rise to subjective experience. Analogous mechanisms also turn out to underlie the phenomena of amodal completion. So, while it is quite plausible that IIT constrains consciousness in some ways, we will argue that it doesn't successfully pin down its nature. Lastly, we will argue that data from visual agnosia present a problem for AIR. The failure of agnostic patients to identify objects they report being conscious of is sometimes based on insufficient content integration, as defined by AIR. This suggests that integration, as construed by AIR, isn't necessary for consciousness.

3 Integration theories

In this section, we will briefly outline the specifics of the three target theories: information integration theory, global workspace theory, and attended intermediate-level theory.

3.1 Information integration theory (IIT)

Information Integration Theory (IIT), originally advanced by Giulio Tononi, accepts the identity thesis (A): for a mental state to be conscious just is for it to have an integrated content (Tononi 2004, 2008, 2010). Unlike other theories of consciousness, IIT starts with the phenomenology of experience. Reflection on the phenomenology of experience yields (five) axioms (e.g., information, integration, exclusion, etc.).[2] The theory then specifies the conditions a physical mechanism (e.g., neurons and their connections) or a system (e.g., a brain) must satisfy in order to generate phenomenal experiences (referred to as postulates).[3] Whether this approach is better suited to the study of consciousness is controversial. Bayne (2018), for example, argues that such an axiomatic approach is not well-suited to the study of consciousness, in part, because axioms, while closely associated with mathematics and logic, are absent in accounts of explanation within neuroscience or psychology.

According to the axiom of information, each experience differs phenomenally from other possible experiences (Oizumi et al. 2014). For example, an experience of a blue square differs phenomenally from an experience of a red square. According to the axiom of integration, each experience is irreducible to its parts (Oizumi et al. 2014). For example, when one sees the word 'honeymoon', one does not perceive it as a mere co-occurrence of two distinct experiences: an experience of the word 'honey' and an experience of the word 'moon'. Rather the word 'honeymoon' constitutes a single experience irreducible to the simple combination of an experience of 'honey' and an experience of 'moon' (Tononi and Koch 2015). Similarly, when looking at a blue square, one is not simply aware of the uninstantiated blue color and an achromatic square shape. Rather, one is aware of blue and square as instantiated together in the same place.[4] The experience of the blue square cannot be reduced to experiences of its components, i.e., an experience of the blueness and an experience of the squareness.[5] These examples indicate that the axiom of integration concerns representational unity (see, e.g., Bayne and Chalmers 2003). IIT treats these axioms as fundamental properties of experience.[6]

In light of the aforementioned axioms or fundamental properties of experience, several postulates are posited to explain what kinds of physical mechanisms could account for the phenomenology of experience. According to the two postulates of information and integration, information is integrated when a mechanism or system generates information not simply through the combination of its discrete components but rather through a kind of integration that yields 'much more' information than the mere combination of its parts. ITT posits that in order for a mechanism to be able to generate experiences with specific/distinct phenomenology, it must specify

[2] Bayne (2018) argues that the phenomenological axioms posited by ITT fail to capture the essential features of every experience.

[3] ITT's postulates identify an experience with the set of all mechanisms (i.e., the "conceptual structure") and the maximally irreducible probability distribution of potential past and future states of a system as informed by a mechanism in its current state (i.e., its "cause–effect repertoire. See Tononi and Koch 2015).

[4] But why, one might ask, is there more information contained in the representation of a blue square at some location L1, than there is in the representation of a color at L1 and a shape at L1? Why wouldn't the first representation be reducible to its parts? The reason for this seems to be that the former representation carries the information that the color and the shape are properties of a single object (and thus their fates are non-accidentally correlated), whereas the latter representation lacks this information.

[5] IIT's contention that the experience of a blue square has its integrated content essentially is closely tied to the unity of consciousness: conscious unity can be understood in terms of the irreducibility of its components (see Bayne and Chalmers 2003).

[6] For arguments against this and other claims pertaining to the axioms see Bayne (2018).

information (i.e., a “cause-effect repertoire”) that is not reducible to its parts (Oizumi et al. 2014, p. 4).

On this view, consciousness is a fundamental property (as fundamental as mass or charge) of physical systems that have (internal) causal powers that determine past and future states, which

> can be accounted for by the intrinsic cause-effect power of certain mechanisms in a state — how they give form to the space of possibilities in their past and their future. An analogy is mass, which can be defined by how it curves space-time around it — except that in the case of experience the entities having the property are not elementary particles but complexes of elements, and experience comes not in two but in a trillion varieties. In this general sense, at least, IIT is not at odds with panpsychism.

Although IIT may not be at odds with panpsychism, it should be distinguished from it.[7] Whereas panpsychism holds that all things are conscious, IIT attributes consciousness only to systems capable of distinguishing among alternatives (Tononi 2008). Moreover, unlike panpsychism, IIT takes consciousness to be graded, which is to say, not all conscious systems are conscious to the same degree.[8]

Since consciousness is identified with integrated information that is generated by a mechanism or a system over and above the information generated by its parts, it can be quantified. A mathematical quantity, φ, which is, at least in principle, measurable, is thus used by proponents of IIT to quantify consciousness (Oizumi et al. 2014; Tononi 2004, 2008). The idea is that a mechanism or a system that has higher φ will produce more conscious states than a mechanism or a system that has a lower φ. For example, a particular quality of an experience such as the blueness of a square is generated by the integration of information at the substrate level. The fact that the experience of a blue square has a distinctive phenomenology than, say, an experience of a green square is expressed by its φ value. This value measures consciousness by measuring whether a mechanism or a system causes specific experiences when it integrates information.

3.2 Global workspace theory (GWT)

Global Workspace Theory (GWT), originally proposed by Bernard Baars, accepts thesis (C) but rejects (D). It thus maintains that information integration is necessary but not sufficient for consciousness (Baars 1988, 1997, 2002). On this view, integration requires more than a unification of sensory information originating from different sources. In order for sensory information to become integrated in the way required by

7 For discussions of panpsychism see, e.g. Chalmers (2015), Strawson (2006), and papers in Brüntrup and Jaskolla (2016).

8 One might wonder how IIT would account for experiences in different sensory modalities that pick out the same feature of a given object. Consider, for example, the case of seeing and holding a ball. The visual and tactile experiences of the ball’s spherical shape have different phenomenal characters. How would IIT account for this difference in phenomenology, given that both experiences pick out a spherical shape? One way for IIT to accommodate this sort of phenomenal difference is to take the shared feature to be integrated differently in the different sensory modalities or submodalities. For example, the sphericality of the ball may be taken to be integrated in a visual way in one case but in a tactile way in the other (for a similar solution to the general problem, sometimes referred to as ‘Molyneux’s question’, see Chalmers 2004).

GWT, it needs to be broadcast to an executive cognitive system referred to as the 'global workspace' (Baars 1988, 1997, 2002; Franklin and Graesser 1997; Dehaene et al. 2006; Dehaene and Changeux 2011).[9] Global workspace is a working memory system for temporarily storing and managing information required to carry out complex cognitive tasks such as decision-making, reporting and complex problem solving (Baars 1988, 1997, 2002; Franklin and Graesser 1997; Dehaene et al. 2006; Baars and Franklin 2009; Dehaene and Changeux 2011). 'Broadcast' here is a term of art. In order for information to be broadcast to global workspace, it does not suffice that it is available for use in cognitive tasks, if by 'available' we simply mean availability for potential encoding in working memory. All of the information from specialized motor, perceptual and cognitive subsystems, or 'modules' that is needed for a given cognitive task must be stored, or encoded, *together* in working memory. This sort of encoding ensures that information from different sources is combined in the way that is required for carrying out a given cognitive task. As Stanislas Dehaene and Lionel Naccache put it:

> Through the workspace, modular systems that do not directly exchange information in an automatic mode can nevertheless gain access to each other's content. The global workspace thus provides a common 'communication protocol' through which a particularly large potential for the combination of multiple input, output, and internal systems becomes available (Dehaene and Naccache 2001, p. 13).

To see this, suppose you are a research participant saddled with the task of pressing a button only when you both hear a sharp tone and see a blob with soft edges. In order to complete the task correctly, you will need to encode the information contained in the instruction as well as the information provided to you visually and auditorily. On the basis of those informational inputs, you will need to make a cognitive decision about whether or not to press the button. If you are unable to make such a decision on the basis of the information you have received from the different sources, then the information from these different sources is not encoded in a way that makes it useful for completing the relevant task. The information in this case is not sufficiently integrated for it to become conscious.

While encoding in working memory is necessary for consciousness to arise, it is not sufficient. For the encoded information to reach consciousness, it must also be amplified by selective attention (Baars leaves open the possibility that additional conditions are necessary for consciousness) (Dehaene et al. 1998; Dehaene and Naccache 2001; Posner 1994; Posner and Dehaene 1994). Only information that is broadcast to working memory in combined form and is amplified by selective attention has the potential to reach the level of consciousness.

To shed light on the claim that both encoding and amplification are needed in order for sensory information to become consciousness, defenders of GWT employ a theater metaphor (Baars 1997). Consciousness arises when a 'spotlight of selective attention'

---

9 The global workspace cannot be associated with a fixed set of brain areas because various cortical areas can contain workspace neurons with suitable long-distance and widespread connectivity needed to give rise to conscious experiences. However, the fact that workspace neurons seem to be denser in certain areas such as the prefrontal cortices (PFC) and the anterior cingulate (AC) suggests that these areas play a dominant role in the function of the global workspace.

lights up a specific portion of the stage. The actors, their actions, and the scene props (information) that are illuminated by the narrow, intense beam of light constitute the visible activities that contribute to the overall narrative (conscious content). The theater personnel, such as the director (executive processes), the playwrights, scene designers, and choreographers are in charge of what becomes visible on the stage, but they are not themselves visible.

GWT's requirement that attention is needed for consciousness is somewhat dubitable if one holds that phenomenal consciousness and access consciousness actually come apart (Block 1995).[10] However, this requirement is far less contentious given the assumption underlying GWT, namely that consciousness simply *is* access consciousness.[11]

Another worry about GWT, raised by Jesse Prinz (2012, pp. 29–31), is based on GWT's claim that working memory can contain information that isn't amplified by attention and therefore isn't conscious. The problem with this claim, Prinz argues, is that it requires that attention operate on higher-order information already stored in working memory. Yet, he argues, there is no evidence to suggest that attention amplifies information stored in working memory. Rather, the evidence strongly indicates that when we attend to things in the external world, attention alters the content of the corresponding perceptual states.

However, there is a fairly simple reply to this worry. Recent evidence seems to support a distinction between working memory proper and fragile working memory, at least in the case of visual experience (Neisser 1967; Vogel et al. 2001; Sligte et al. 2008). Fragile visual working memory is an intermediate form of visual working memory that is thought to last longer than iconic memory but not as long as visual working memory. When information is processed by the visual system, it is first stored in iconic memory. From there it enters fragile visual working memory and then visual working memory. Information stored in fragile visual working memory is only potentially available for use in cognitive tasks. Given the distinction made by Block (1995) between phenomenal consciousness and access consciousness, information in fragile visual working memory may be phenomenally conscious but it is not access conscious. Barring this distinction (and going along with GWT),[12] information in fragile visual working memory is not conscious and therefore needs to be amplified by attention in order for it to enter visual working memory and thereby become conscious.

[10] Block (1995, p. 234) argues that there is a natural use of 'consciousness' and 'awareness' corresponding to 'access consciousness' and 'phenomenal consciousness', respectively. According to Block, phenomenal consciousness can be understood as awareness, whereas access consciousness is better understood as consciousness proper. When a content is both P- and A-conscious, Block suggests that we speak of 'conscious awareness'.

[11] Raffone and Pantani (2010) propose a variant on GWT that purports to accommodate Block's (1995) distinction between phenomenal and access consciousness.

[12] Many of those who agree that there is a meaningful distinction to be drawn between phenomenal and access consciousness argue that the two do not actually come apart. David Chalmers (1997), for example, argues that while phenomenal consciousness and access consciousness, as defined by Block (1995), coincide (i.e., are both present or both absent in the actual world), it is conceivable and, therefore, possible that they come apart. His view supports those like Baars and Prinz who take P-consciousness and A-consciousness to coincide.

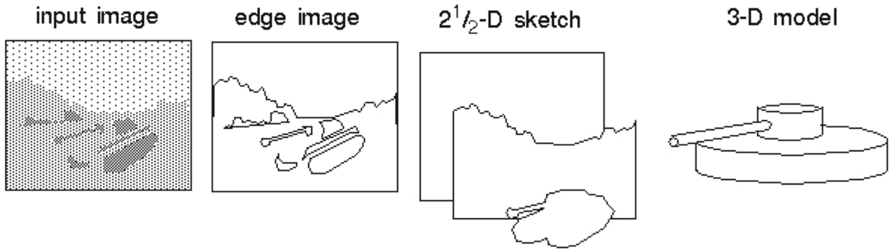

Fig. 1 On Marrs' view, which is similar to Prinz's, the pattern of light hitting the retina first gives rise to a low-level representation of blobs and edges, or what Marr calls 'the primal sketch' (the edge image) (also sometimes referred to as 'retinotopic representation'). From this representation, a 2½-D sketch is generated. This serves as the input to the viewpoint-independent 3D model of the object (here intentionally distorted in such a way to be shown from the viewer's perspective)

Thus, given the evidentially supported distinction between fragile working memory and working memory proper, Prinz's worry does not present a challenge for GWT.

3.3 Attended intermediate-level representation theory (AIR)

The attended intermediate-level representation theory of consciousness (AIR), advocated by Prinz (2000, 2012), holds that conscious experiences are attended intermediate-level representations. Intermediate-level representations differ from low-level and high-level representations in several ways. Low-level representations represent local features of a stimulus, such as edges or contours, in a disunified, two-dimensional way. High-level representations represent abstract viewpoint-independent features of three-dimensional objects that are generated by abstracting away from the vantage point of view and surface details. Intermediate-level representations, by contrast, represent objects and the features they instantiate from the perceiver's point of view. Because they reflect the retinal imprint (when veridical) yet capture information about Gestalt grouping, e.g., depth and orientation, which are also referred to as '2½-D sketches' in David Marr's (1982) terminology (see Fig. 1).

According to AIR, only representations that occur at the intermediate level have the sort of integration needed for them to become available to working memory. Attention, which Prinz takes to have a neural correlate in a pattern of neural oscillation known as 'gamma waves', is a modulation of the intermediate-level representations that makes them *available for encoding* in working memory. Attention is thus a 'gateway' to working memory (Prinz 2012, p. 334). Encoding in working memory, Prinz argues, just is the subjectively experienced maintenance of information over a brief time interval. As he puts it:

> The simplest explanation for this [intimate relationship between attention and working memory] is an identity claim: attention can be identified with the processes that allow information to be encoded in working memory. When a stimulus is attended, it becomes available to working memory, and if it is unattended, it is unavailable (Prinz 2012, p. 93)

The reason only intermediate-level representations can serve as the contents of conscious experience, according to Prinz, is that the low-level representations of 'oriented edges, bars, ends, and blobs' that are first generated by the visual system are too disunified to carry any meaning, whereas the high-level representations of viewpoint-independent objects fail to capture what we are aware of when we have a conscious experience (Prinz 2012, 50 ff; see also Prinz 2000).[13] For example, when you are looking at a quarter viewed at an angle, you can thereby come to know that the quarter is disk-shaped but it does not show up as disk-shaped in your conscious experience. Rather, it shows up as elliptical. Prinz argues that perceptual constancies, such as the disk-shape of the quarter, are not presented in *sensory* experience. If we experience the quarter from atop, it looks circular-shaped (or round), but not disk-shaped. However, we can infer that the quarter is disk-shaped on the basis of past experiences of quarters from many different perspectives.

Since attention is the processes that allow information to be encoded in working memory, on Prinz's view, consciousness essentially involves attention (Prinz 2000, 2012). Attention makes intermediate-level representations available for potential encoding in working memory, which suffices for them to become conscious (Prinz 2012, p. 106).[14] But, Prinz argues, empirical evidence suggests that no actual encoding is necessary (Prinz 2012, pp. 99–102). This can be seen from cases of change blindness. Change blindness, which has been demonstrated to occur in real-world interactions, involves a failure to detect changes in a scene (Ballard et al. 1994). In one famous experiment, subjects failed to notice that the person (A) they were giving directions to on the street had been replaced with a different person (B) after a temporary distraction created by carrying a mirror or a door between the subject and A (Simons and Levin 1998). This indicates that while the subject had separate experiences of A and B, the information pertaining to the change from A to B failed to be encoded in working memory. As a result, the subject failed to notice that she was no longer talking to A. Prinz argues that while the change from A to B is not actually encoded in working memory, the fact that the subject experienced both people makes this information available for recall (provided that the research participant was asked immediately after viewing). This indicates that what's needed here is availability to working memory, not actual encoding in working memory.

Because GWT takes consciousness to require actual encoding in working memory, whereas AIR merely requires availability for encoding, Prinz (2012) notes that AIR could be understood as a dispositional version of GWT. As he puts it:

[13] In the case of vision, the intermediate level is anatomically located in a family of areas involved in processing color (hue), motion and three-dimensional shape (extra-striate brain regions). By comparison, the lower level is anatomically located in primary visual cortex (V1) and some subcortical structures such as the visual nuclei of the thalamus and the superior colliculus; the high level recruits structures in inferior temporal areas (such as TE, TEO, and sections of the superior temporal sulcus), the lateral occipital complex, and some structures in parietal cortex (such as the ventral and posterior inter-parietal areas).

[14] Note that, on this view, high-level representations are used to mediate encoding once an attended item has been selected for use in cognitive tasks or for retention in long-term memory, but Prinz argues, high-level representations are not themselves modulated by attention and therefore do not themselves reach conscious awareness (Prinz 2012).

> The AIR theory might be described as a dispositional global workspace theory, although there is one empirical issue that would need to be resolved before that claim could stick (Prinz 2012, p. 335).

The issue that would 'need to be resolved', Prinz says, is that of whether consciousness would break down if individuals with brain lesions could use stored information directly for cognitive tasks without global broadcasting (Prinz 2012, p. 335).

Setting aside this issue, AIR and GWT agree that integration is facilitated by attention. The main difference between AIR and GWT is that the latter takes integration be the result of actual encoding in working memory, whereas the former takes integration to happen at the intermediate level of processing and consciousness to happen when the information becomes available to working memory. Since attentional amplification is required in order for intermediate-level representations to reach conscious awareness, on Prinz's view, integration is a precondition for consciousness to arise.

4 Integration is not necessary for consciousness

Having provided an overview of each of the target theories, we now turn to the question of whether information integration is necessary for consciousness. As stated above, the necessity thesis (C) can be articulated as follows:

> (C) *Necessity of Integration*: Necessarily, if mental state M is conscious, then M's representational content is integrated ($C \Rightarrow I$).

According to C, having an integrated content is necessary for consciousness. So, if there are cases of conscious mental states that fail to have integrated content, then (C) is false. In the next section we will look at empirical findings that suggest that feature awareness can occur without feature integration. This possibility presents a counterexample to thesis (C). We begin by discussing some empirical results that appear to create trouble for IIT. We provide evidence against GWT and AIR in subsequent sections.

4.1 Evidence against IIT from failures of feature integration

Different parts of the perceptual system are sensitive to different features of the environment such as shape, motion, luminance, and color. The information about these features from different sources needs to be combined to yield the familiar representations of our environment as filled with stable, coherent objects. The question of how the perceptual system manages to produce unified representations of this kind has become known as the 'binding problem' (see, e.g., Hardcastle 1994; Treisman 1996; Hummel 2001; Bayne and Chalmers 2003; Bayne 2010).

Investigations into how to solve the binding problem have yielded a sizable body of research, including discussions of binding failures; that is, cases where the normal processes of combining distinct types of information about different features fail. Such failures of feature integration take different forms. In one type of failure, distinct features are mismatched. For example, when subjects were presented with, say, a green

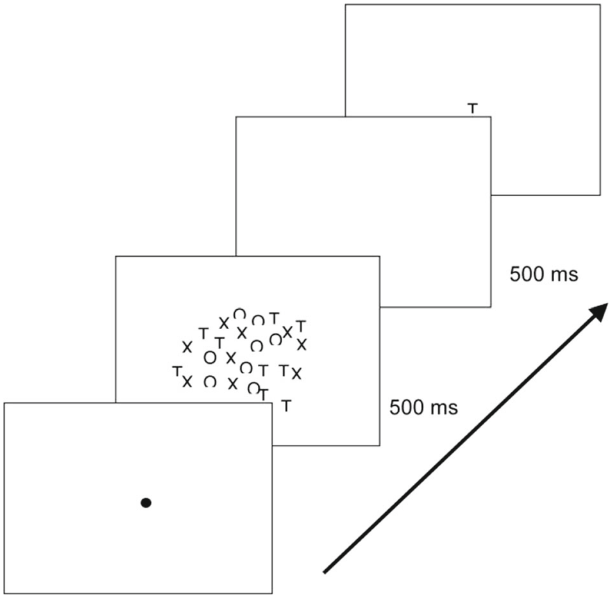

Fig. 2 The set of relevant targets to be reported, i.e., red items, Os, or red Os, was specified immediately after the display was presented. Adopted from Treisman (2006)

letter O and a red letter T, some reported seeing ‘illusory conjunctions’ of red Os and green Ts (see e.g., Robertson et al. 1997; Cohen and Ivry 1989; Robertson 2003). Such failures do not challenge IIT since these subjects are still aware of the illusory conjunction of distinct features, even though these features are mismatched. This is not the only way for feature integration to fail, however.

As illustrated by the studies discussed below, subjects can sometimes detect several features of the same object independently of each other. For example, they may detect color but not extension, or vice versa. Conscious awareness of isolated features that are unbound appears to present a counterexample to IIT. This is because IIT is committed to the claim that conscious experience cannot exhibit this kind of disintegration (as indicated by its axioms). As Tononi (2004) writes,

> The integration of information in conscious experience is evident phenomenologically: when you consciously ‘see’ a certain image, that image is experienced as an integrated whole and cannot be subdivided into component images that are experienced independently. For example, no matter how hard you try, you cannot experience colors independently of shapes, or the left half of the visual field of view independently of the right half (Tononi 2004, eprint pub.).

However, the experience of colors independently of shapes and of colors independently of orientation is precisely what the next two studies show *can* happen.

For starters, Anne Treisman’s experiments on feature binding appear to show that we are aware of features in a scene prior to being aware of how they are bound together. In other words, subjects were found to be aware of features in the scene

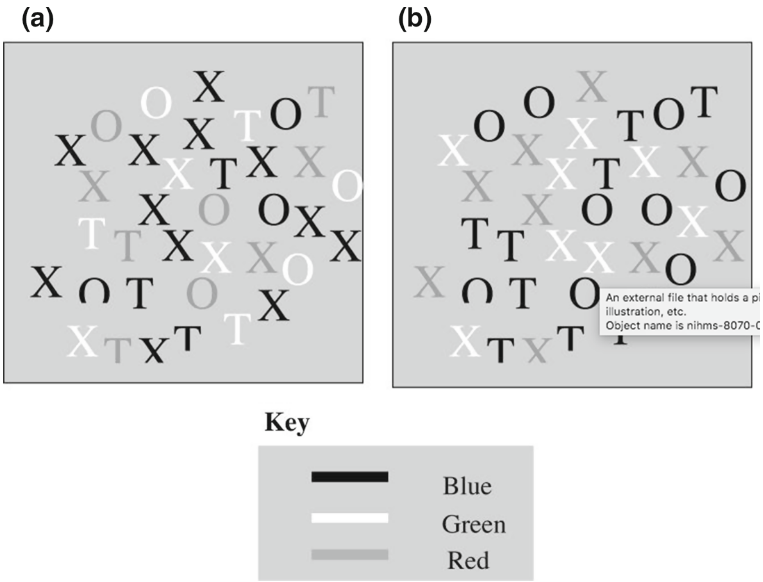

Fig. 3 Examples of two displays with equal numbers of each target type but different numbers of various conjunctions. The display in **a** has 36% blue Xs but the display in **b** has none. Adopted from Treisman (2006)

before they were being integrated (bound) as features of different objects (Treisman 2006). Subjects were presented with brief displays of colored letters, i.e., O, X, and T in red, green and blue, and were asked to estimate what proportion of the stimulus was, for instance, red, or Os, or red Os. The experimenter post-cued which feature or conjunction was relevant. On each trial, the stimulus was presented for 500 ms followed by a 500 ms of a blank screen before a single color patch or a single white letter or a single colored letter was used as a probe (Fig. 2). Subjects were asked to report what proportion of the display a particular feature or conjunction represented (Fig. 3). The results indicated that subjects '*were good at judging the proportions of the separate features that were present, but very poor at judging the proportions of conjunctions*' (Treisman 2006; emphasis added).

These results suggest that the participants were conscious of features such as colors and shapes separately, as suggested by their accurate estimations of proportions of each feature in a scene. However, they did not seem to be conscious of combined colors and shapes, as evidenced by their poor estimation of the proportion of conjunctions of features. These results present a challenge for the IIT. Recall that, according to the IIT, one of the essential properties of consciousness is that it is unified (Tononi 2012; Tononi and Koch 2015). As a consequence, you should not be able to experience, say, colors independently of shapes, 'no matter how hard you try' (Tononi 2004, eprint pub.). Yet, these results indicate that such unbound experiences can occur. This, in turn, shows that the axioms posited by IIT fail to adequately capture the actual phenomenology of experience; and along with them the corresponding postulates that are posited to explain the underlying mechanisms of the phenomenology of experience.

It may be objected that IIT does not entail that colors (generally speaking) cannot be seen independently of shapes as Treisman's experiments seem to indicate. Rather, the theory maintains that the features of a single stimulus (say, a blue square) can only be experienced as an integrated whole. However, as we have argued, Treisman's experiments show the opposite: namely, subjects did not experience the features of a single image as an integrated whole. For while subjects were shown two distinct displays, these displays contained exactly the same proportions of the different features (half blue letters, half Xs, a quarter each red letters, green letters, Ts, and Os). The only difference between the two displays was that the features were *differently bound*: one display had 33% blue Xs and other had none. Yet, the judgments subjects made about feature proportions and proportions for bound objects differed significantly in their accuracy. This suggests that the subjects' experiences served much better as sources of evidence for estimating the proportions of unbound features than they did as sources of evidence for estimating the proportions of bound features. It seems natural to locate the reason for this difference in the experiences' phenomenology. There was something it was like for the subjects to experience unbound features, and there was nothing it was like for the subjects to experience bound features. This explains the more accurate report about the unbound features and the less accurate report about the bound features.

The proponent of IIT may nevertheless insist that the differences in the subjects' responses do not present a problem for IIT because either the nature of the task influenced what subjects experienced or it influenced how subjects directed their attention. However, neither of these explanations are applicable to Treisman's experiments since the subjects were given the task instructions *after* they were shown the target displays. It seems unlikely, then, that the nature of the task altered what they experienced or how they attended to the experience itself. Moreover, Treisman notes that these differences cannot be attributed to changes in focused attention (which involves conscious access to stimuli with bound features in their current locations) because the subjects deployed global attention (which involves global and statistical properties of the general layout of a scene).[15]

In another study, conducted by Neri and Levi (2006), subjects were shown an array of rectangular segments. Each segment could be either horizontally or vertically oriented, and each could be either dark or bright in color. The subjects were then asked to report whenever they saw an array where (1) one segment had different color, regardless of orientation (e.g., a dark segment among bright segments); (2) one segment had a different orientation, regardless of color (e.g., a horizontal segment among vertical segments); and (3) one segment had a different orientation and color compared to the rest (e.g., a horizontal dark segment among vertical bright segments) (Fig. 4). The arrays were presented for 150 ms to avoid saccades. The detection threshold is taken to be the size of the colored segments at which over 82% of the subjects' answers were correct.

When the arrays were presented to the fovea (i.e., in the center of the visual field), detection thresholds for orientation and for orientation-and-color were the same. The subjects were as good at determining which of the segments was the odd-one-out by orientation alone as they were at determining the same thing by the combination of

[15] We are grateful to an anonymous reviewer for suggesting these objections.

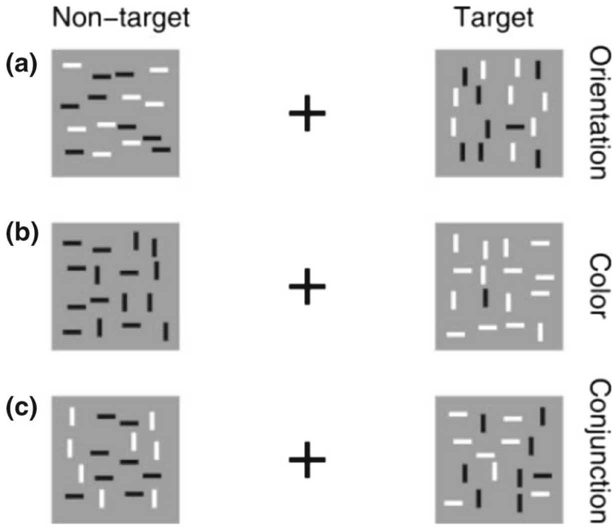

Fig. 4 This figure represents example stimuli. **a** The orientation discrimination task: Non-targets (left) contain segments of the same orientation while targets (right) contain 1 segment of different orientation. **b** Uses the same idea for color discrimination. **c** The conjunction task: here non-targets and targets were equated in overall color and orientation content. Adopted from Neri and Levi (2006)

color and orientation (color alone is not mentioned because detection thresholds for color are lower than detection thresholds for orientation).

But when the stimuli were presented in peripheral vision, detection thresholds for orientation-and-color were much higher than detection thresholds for orientation alone (meaning that the combination was much more difficult for the subjects to detect than individual features). This suggests that there are cases in which each of the two features (color and orientation) is visible separately but they are not visible in combination. At the very least, there are cases where the subjects can discriminate the color of the stimulus and the orientation of the stimulus separately but cannot discriminate the combination of color and orientation. In this case, feature integration appears to be an obstacle to consciousness, which goes against the claims made by advocates of IIT. Again, these findings indicate that the axioms posited by IIT fail to adequately capture the actual phenomenology of experience; and along with them the corresponding postulates that are posited to explain the underlying mechanisms of the phenomenology of experience.

Our interpretation of the results assumes that verbal reports from study participants indicate that they were phenomenally conscious of what they reported. While the issue of the connection between reportability and consciousness is fraught with controversy, we do think it is relatively uncontroversial to assume that report is good evidence for consciousness. Moreover, the length of time for which stimuli were presented in both studies described above (150 ms in the Neri and Levi study and 500 ms in the Treisman study), gives us further confidence that the stimuli were indeed consciously perceived. This is because there is empirical evidence that stimuli presented for much shorter periods are likely to be consciously processed by subjects (Pins and Ffytche

2003). While, admittedly, it could be the case that experience outstrips report in both studies we cite, the onus is on the proponents of IIT to explain why we should not take subjective reports at face value. Tononi and Koch (2015) argue that behavior can be misleading and cite cases in which a person may sleepwalk or be unresponsive. Although we agree that such case reports can be misleading, there is nothing so alarming in these experimental conditions that would lead us to outright dismiss their veracity. For we are dealing with responsive adult humans who are fully capable of understanding the experimenter's instructions as opposed to, say, unresponsive adults, infants, or animals.[16]

To summarize: Treisman's (2006) study indicates that unity (in Tononi's sense) need not apply to all conscious experiences. Recall that the accurate estimations of proportions of each feature in the scene indicated that the subjects were conscious of features such as colors and shapes *separately*. At the same time, the poor estimation of the proportion of conjunctions of features indicated that subjects did not appear to be conscious of color and shape simultaneously. Similarly, in the Neri and Levi (2006) study, detecting the combination of orientation and color in peripheral vision was more difficult than detecting either orientation on its own or color on its own. In other words, detection thresholds for orientation-and-color were much higher than detection thresholds for orientation alone. This suggests, in line with the authors' own conclusion, that each of the two features (color and orientation) were experienced separately but not in combination. It follows that, contrary to IIT's axioms, subjects did not have a unified experience of color-and-orientation (or color-and-form) but rather an experience of color and an experience of orientation (or form).

Recall that IIT's approach to the question of consciousness is the opposite of the approach that is typically taken in neuroscience. IIT first posits a set of axioms that describe the phenomenology of experience as primary. It then posits a set of postulates that purport to explain how these phenomenological axioms can be implemented by physical mechanisms. But if the axioms fail to adequately capture the phenomenology of experience, then the postulates will also fail to capture how they are implemented by the brain.

4.2 Evidence against GWT from failures of global broadcast

As long as we take the subjects' performance (i.e., forced-choice guesses or reportability) to genuinely reflect what the subjects are conscious of, it is difficult to see how the above studies present a challenge to GWT. Indeed, some proponents of GWT maintain that conscious experience must be evaluated by subjective measures, i.e.,

[16] As an anonymous reviewer noted, proponents of IIT could insist that these subjective reports should not be taken at face value because the Φ is maximal at posterior parts of the brain. Since the frontal activity associated with reportability is not part of the neural correlates of consciousness, reportability is suspect as a guide to the phenomenology of experience. However, this objection seems misguided. Our claim is not that reportability is part of conscious experience but rather that conscious experience is, in the good cases, accurately reportable. If IIT rejects this claim, then it is unclear what sort of evidence it can provide (apart from reportability) for the claim that its axioms govern the phenomenology of experience. If proponents of IIT claim instead that the Treisman cases are not "good" for whatever reason, but that report does generally serve as a guide to phenomenology (especially when it comes to their axioms), then the latter claim, given the disassociation between frontal activity and consciousness, starts seeming exceedingly mysterious.

reportability (Dehaene et al. 2006).[17] So if all that the subjects experience in those cases are disembodied colors and colorless shapes, then it follows from GWT that this information (i.e., disembodied colors and colorless shapes) will be broadcast to different parts of the cognitive system.

GWT is not immune to criticism, however. A study conducted by Baddeley and Wilson (1988) lends support to the view that there could be conscious experiences whose content is not properly integrated, because it is unavailable for use by the subject's cognitive system. Baddeley and Wilson studied patient R.J. who suffered from frontal lobe lesions which resulted in serious deficits in cognitive functioning tasks. In one of the tests conducted by the experimenters, R.J. was tasked to complete a very simple 12-piece jigsaw puzzle. The researchers describe his performance as follows:

> [R.J.] was consistently held up at particular points, often taking a piece that the picture indicated could not possibly fit in a given location, and systematically changing its orientation in the hope that it would fit. The problem appeared to be one of strategy, indicating both *an apparent failure to be able to take account of information from two sources such as shape and pattern simultaneously*, and a tendency to perseverate in the attempt to fit a piece in the wrong place. (Baddeley and Wilson 1988, pp. 221–223, emphasis added)

In contrast to the first two studies, there is no independent indication in this study that the aberrant observations are the result of disturbances or abnormalities to the proper function of the visual system. Since R.J.'s visual experience of the chosen piece and the piece required by the picture was not aberrant, he must have seen the pattern and shape of the chosen piece as well as the pattern and shape of the piece required by the picture. Yet he was unable to use the visual information to solve the puzzle. The most plausible reason for this seems to be that the visually provided information failed to combine when it was broadcast to executive areas. Recall that GWT holds that integration of the sort that gives rise to consciousness consists in the actual encoding of the information in global workspace. But this requires that the information is available for use in cognitive tasks such as reporting, decision making, or problem solving.

This case presents a challenge for GWT. R.J. appears to have a conscious experience of the piece in his hand as well as the location he is instructed to fit it into. Yet the content of his experience is not jointly encoded in global workspace, which would be needed for it to be available for use in cognitive tasks such as fitting the piece in the right place. It follows that this is a case of consciousness in the absence of proper integration.

In another study, conducted by Konow and Pribram (1970), patient I.S. who suffered frontal lobe damage was asked to draw simple shapes on the basis of the experimenters' verbal instruction. I.S. was unable to follow the instructions correctly: on one occasion, when asked to draw a square, she kept on drawing circular shapes; on another, she drew an A-shape in response to almost any instruction. I.S. understood the content

[17] While objective measures typically involve asking subjects to make forced-choice guesses about what they have seen, subjective measures typically involve reportability (Szczepanowski and Pessoa 2007; Kunimoto et al. 2001). When subjects are asked to report on whether they saw a stimulus, negative responses are taken as evidence that the stimulus was not experienced consciously.

of the instruction was: she would, for instance, make remarks to the effect that she immediately recognized that she had made an error (despite the persistent inability to correct it). Moreover, she was also able to notice when someone else made an error of the same sort (e.g., drawing a circle in a response to being instructed to draw a square). Her verbal behavior, therefore, demonstrated that the inability to follow the instructions was not the result of a failure to understand the instructions. Rather, the observations indicate that her drawing deficits resulted from a failure to integrate the content of the instruction with the representation of the shape she ends up drawing. This runs counter to what GWT predicts. Although I.S. fails to use the instructions correctly, she is nonetheless conscious of their content. She is also conscious of the circular shapes she is drawing. It follows that, contrary to GWT, integration is not necessary for consciousness; if it were, we would expect I.S. not to be conscious of the content of the instructions, the content of the drawings, or both.

Proponents of GWT might argue that because the content of the instructions and the shape I.S. sets out to draw are not unified, there isn't a single experience with a content that lacks integration but rather two conscious experiences each of which is integrated. However, this reply is not available to defenders of GWT since they deny that unification plays a role in individuating experiences. Co-consciousness of the contents suffices for there to be a single experience. Yet the contents of that conjoint experience are not integrated in such a way as to be available for use in the requested cognitive task. Hence, contrary to what GWT implies, the content is not integrated in a way that suffices for it to reach conscious awareness.

Defenders of GWT may attempt to refute the objection in a different way. They might argue that while I.S. clearly represents the shapes she eventually draws, it is in no way obvious that she has a conscious experience of those shapes. After all, action is guided by motor representations in the vision for action stream, and there is evidence that these motor representations are not conscious (Milner and Goodale 1995). Suppose you reach to and grasp a coffee cup on your desk. In order to grasp the cup effortlessly you must rely on a conscious experience of the cup and its location but the motor representations that ultimately guide your action must represent various additional features of the situation, including the hand aperture needed for you to successfully grasp the cup and the force you must apply to lift the cup. Nevertheless, you don't sit down and calculate the required hand aperture and muscle force before reaching for the cup. It may thus be argued that these other aspects are not available to consciousness. Since the vision-guided motor representations are not available to consciousness, it may be further argued that she does not have a conscious experience of the shapes she sets out to draw. It follows that there isn't a single conscious experience without integrated content.

While this sort of move may hold some promise for the defender of GWT, it doesn't ultimately succeed. As Milner and Goodale (1995) have argued, action on the fly is governed by unconscious motor representations in the sensorimotor cortex in the vision-for-action stream (see also Goodale et al. 1991; Goodale and Milner 1992; Hu and Goodale 2000; Westwood and Goodale 2003). Everyday actions like reaching for your coffee cup, eating with a knife and fork, using a computer mouse, riding your bike, and parking your car on a busy street are not typically preceded by careful planning (Brogaard 2011b; Brogaard and Gatzia 2017). Yet when an action is planned ahead of time, as in the case of following a recipe, playing darts, or making a move in a game of chess, a conscious representation of the planned action helps us execute the action as planned (Brogaard 2011b). Not all drawing activities are based on careful planning, of course. Doodles are, by definition, an excellent example of an action carried out on the fly. But many other kinds of drawing seem to be the result of planning. Arguably drawing isn't altogether different in this respect from completing a JigSaw puzzle. Putting down a piece on the table isn't an action done on the fly but rather an action that you intend to align with the outcome of your prior deliberation. I.S. was indeed given ample time to plan her drawing but failed owing to the lack of integration of the information from the different sources. These considerations suggest that I.S. does indeed have a conscious experience composed of an experience the represents the shape she is asked to draw and an experience that guides her when drawing the shapes she actually draws. As the components of the content of the experience are not integrated, this case presents a problem for GWT.

Yet another avenue for the proponents of GWT to respond would be to suggest that while the information about what shape to draw is available for use in the drawing task, the failure to carry out the instruction correctly is caused by some fault in the system consuming this information *after* it has been made available for use. If this were the case, I.S. would no longer constitute a counterexample to GWT, as the proponents of the theory could still claim that the information about what shape to draw is conscious but I.S.'s inability to draw it doesn't show the information is not available for other systems. The failure to carry out the instruction could be attributed to some fault in the processing downstream from integration.

However, as Pribram and Konow report, I.S. is able to draw the requisite shape if instructed to copy a shape she is shown. This suggests that there is nothing wrong with the system consuming integrated information for use in the drawing task. Rather, it remains plausible that the essence of I.S.'s problem lies in the content of the verbal instruction failing to be made available for use in the first place.

4.3 Evidence against AIR from failure of shape and depth perception

Recall that AIR holds that consciousness occurs when and only when an integrated intermediate-level representation is made available to working memory by being attentionally amplified. While attentional attenuation is necessary and sufficient for consciousness, integration of information into a 2½-D sketch is only necessary. This kind of integration is necessary, according to AIR, because low-level representations lack meaning, whereas high-level representations are too abstract and subject-independent to be able to serve as the content of *subjective* experience.

AIR is not subject to the problems that threaten to undermine IIT and GWT. Since AIR doesn't require only availability for encoding, not actual encoding in working memory for consciousness to arise, it avoids the above objections to GWT. Like IIT, AIR requires integration of low-level features for consciousness to arise. However, Prinz maintains that the notions of integration and binding should be kept apart (Prinz 2012, p. 40). Feature binding can fail when stimuli are presented very quickly to subjects or are presented in the periphery of their visual field (Treisman 2006). In

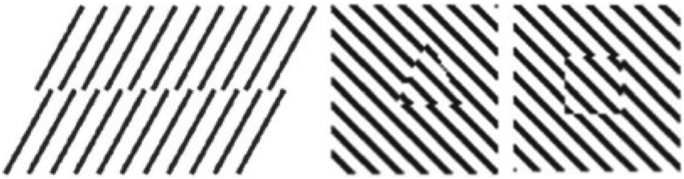

Fig. 5 Examples of illusory contours owing to Gestalt grouping by similarity in texture stimuli. Despite being consciously aware of the stimuli, D.F. was unable to perceive the illusory contours resulting from the Gestalt groupings (Milner et al. 1991)

such cases, subjects may only become aware of the color or the shape of the object or location. Feature binding can also fail as a result of lesions to the posterior part of the visual cortex and/or regions of the temporal lobe, as is seen in patients with visual form agnosia (Prinz 2012, p. 40). But, according to Prinz, while feature binding fails in these cases, the content of the object representation is sufficiently integrated for it to be become available to working memory and thereby acquire consciousness. As these experiences have integrated content, failure of feature binding does not present a problem for AIR.

There is, nevertheless, reason to question Prinz's claim that the content of patients with visual form agnosia is integrated. Milner et al. (1991) studied visual agnosia patient D.F., who had permanent lesions to the lateral visual cortex after recovering from carbon monoxide poisoning. In spite of the lesions, she retained the ability to consciously perceive brightness and hue but had a significant deficit in shape perception, which resulted in a failure of object recognition. She also had a reduced ability to perceive brightness and motion. Her deficit in shape processing was caused by an inability to perceive boundary shape and orientation, whether conveyed by color, intensity, depth, motion, proximity, good continuity, or similarity. For example, she was unable to perceive Gestalt grouping by similarity in texture stimuli, as illustrated by Fig. 5.

Milner et al. (1991) argue that D.F.'s inability to perceive boundary shape and orientation, regardless of how these boundaries were conveyed, shows that her deficit doesn't arise because of an inability to detect local edge differences on the basis of luminance contrast. Rather, it reflects a failure to integrate local differences into a global whole.

So despite her conscious experience of color and local edge features, D.F. was unable to integrate these features into a 2½-D sketch or intermediate-level representation in Prinz's terminology. The fact that D.F. was able to describe her experiences of colors and local edges shows that the primal sketch (or low-level representation) was made available to working memory through attentional modulation, thereby acquiring consciousness. This finding goes against Prinz's claim that only intermediate-level representations, which occur between low-level representations and abstract 3D models, can become conscious when modulated by attention.

In another study, Turnbull et al. (2004) examined a patient, D.M., who suffered from visual agnosia following a closed head injury from a car accident three years prior to testing. The brain lesions had resulted in an impairment of recognizing drawn

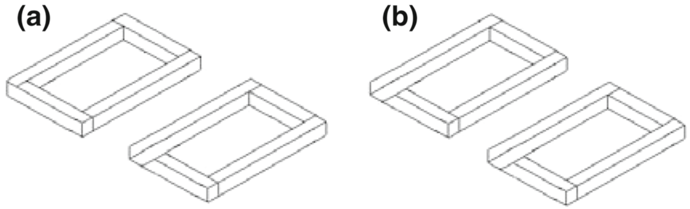

Fig. 6 **a** One ‘possible’ and one ‘impossible’ figure was shown to DM. **b** Two figures that both could be ‘possible’ or ‘impossible’ were then presented to DM. Adopted from Turnbull et al. (2004)

objects that was caused by a failure to derive the appropriate three-dimensional structure from the pictorial depth-cues in two-dimensional images. In spite of the lesions, D.M.’s ability to obtain two-dimensional information from the images such as 2D shape, brightness, acuity, color, and motion enabled him to complete tasks that merely required relying on the images’ two-dimensional structure. Yet owing to the failure of depth perception, he was unable to complete tasks that required extracting depth cues from the images.

In one 3D discrimination task, D.M. was presented with one ‘possible’ and one ‘impossible’ line drawing as shown in Fig. 6a and told that he would see two drawings ‘one of which is a good drawing, while one is a bad drawing that doesn’t make sense if you look closely’. Using a forced-choice paradigm (an objective measure), he was then asked to select the ‘bad’ drawing by pointing to it in a series of ten trials during which the position of the correct item was varied.

In a 2D control task, D.M. was presented with pairs of drawings (Fig. 6) that in one half of the trials were both possible or both impossible (Fig. 6b) and in the other half consisted of one possible and one impossible figure (Fig. 6a). In the control task, D.M. was asked to determine whether the drawings were the ‘same’ or ‘different’, regardless of whether they were possible or impossible.

D.M. fared no better than chance in the 3D possible-impossible discrimination trials but answered correctly in the 2D control trials. As the 2D task can be completed without extracting three-dimensional structure from the drawn objects, these results show that D.M. had cognitive access to two-dimensional information about the drawn objects but was unable to extract information about their three-dimensional structure.

D.M.’s deficits in depth perception made him unable to derive three-dimensional structure from line drawings presented to him. While the 2D features D.M. obtained from the drawings were not integrated into a 2½-D sketch, the study shows that they were nonetheless available to working memory. The study thus demonstrates that low-level information can acquire consciousness, contrary to what AIR predicts.

Further evidence for DM’s failure of depth perception came from the fact that unlike normal perceivers, he wasn’t susceptible to the Ponzo and Müller–Lyer illusions (Fig. 7). In the Ponzo illusion, the line at the top appears to be longer than the line at the bottom even though both lines have the same length. Similarly, in the Müller-Lyer illusion, when the fishhooks are pointed outwards, the line appears longer than when the fishhooks are pointed inwards. In both cases, the lines that are represented as being

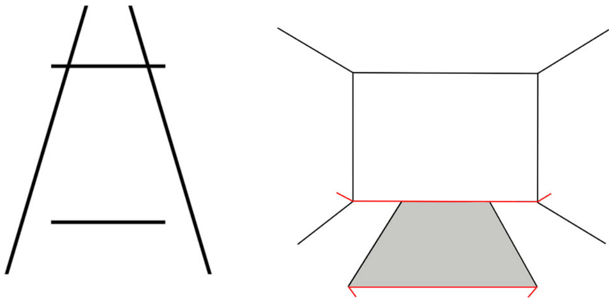

Fig. 7 The Ponzo illusion (left): the top horizontal line appears longer than the bottom line, even though they have the same length. The Müller-Lyer illusion (right): the line with the outward pointing fishhooks appears to be longer than the line with the inward pointing fishhooks, even though the lines have the same length

further away from us appear longer than the lines that are represented as being closer to us.

Whereas normal perceivers fail to accurately judge the two-dimensional extent of the lines in the two illusions, D.M. effortlessly judged the lines to be of equal length. D.M. made accurate judgments in the two illusions because he didn't experience the lines as located in depth at some distance from himself. When the lines are seen as lying in the picture plane, they are naturally interpreted as having the same length. Unlike normal perceivers, D.M. is thus making a judgment about the extent of the lines without extracting three-dimensional structure from the illusory figures. This shows that D.M. had conscious access to two-dimensional information about the illusory figures, a finding that goes against Prinz's claim that only intermediate-level representations, which provide information about depth, have the kind of integration that is required for conscious access.

Interestingly, Turnbull et al. (2004) found that while D.M.'s ability to recognize drawn objects was severely compromised, he had retained some ability to recognize real objects on the basis of richer surface-based cues absent from line drawings, such as information about texture and shading. They take these findings to support the view that object recognition can be accomplished on the basis of two-dimensional information alone. This suggests that many cognitive tasks can be accomplished without depth perception as long as surface-based cues are made available to working memory, contrary to what AIR predicts.

5 Integration is not sufficient for consciousness

Having challenged the claim made by our three target theories that integration is necessary for consciousness, we now turn our attention to the question of whether information integration is sufficient for consciousness:

(D) *Sufficiency of Integration:* Necessarily, if mental state S's representational content is integrated, then S is conscious ($I \Rightarrow C$).

According to (D), if mental state *S*'s representational content is integrated, then *S* is conscious. So, if there are cases of mental states with integrated content that fail to be conscious, then (D) is false.

The current evidence suggests that unconscious processing is possible both at perceptual and semantic levels (for reviews, see e.g. Köhler and Moscovitch 1997; Shacter et al. 1998; Brogaard 2011a; Brogaard and Gatzia 2017). Take, for example, studies of hemispatial neglect, a neurological disorder that occurs as a result of unilateral brain damage (Driver and Vuilleumier 2001). It is characterized by an absence of visual awareness of stimuli located on the opposite side of the lesion: if the lesion is on the right, the patient has no visual awareness of the left side of the visual scene. Hemineglect also results in a failure to perform actions that would normally be directed toward the neglected hemifield, such as orienting behaviors and exploratory search. Many neglect patients behave as though half of their surroundings is missing (Driver and Vuilleumier 2001). Despite the lack of conscious awareness of stimuli in the neglected hemifield, priming measures indicate a considerable amount of processing of the neglected stimuli at perceptual and semantic levels (McGlinchey-Berroth et al. 1993). Hemineglect cases, as we will see, provides evidence against the sufficiency of information integration for consciousness.

5.1 Evidence against IIT from hemineglect

Before we discuss the evidence any further, another note about our assumptions is in order. When we talk about hemispatial neglect and the empirical data collected from the patients, we assume that the report of no conscious awareness on the part of the subject is very good (though, once again, not incontrovertible) evidence of no conscious experience. When patients say they do not experience one half of their visual field, we, once again, take them at face value.

In a well-known study of hemispatial neglect, patients were presented with two pictures of a house (Marshall and Halligan 1988). One showed an ordinary house, whereas the other depicted a house with one side being on fire. The burning half of the house was presented to the neglected hemifield. While the patients reported seeing no difference between the two pictures, they consistently picked the picture that depicted the ordinary house, not the half burned house, as the house they would prefer to live in. This suggests that while the burning side of the house was not consciously processed by the patients, they nevertheless were able to detect the information presented to the neglected hemifield and direct their behavior accordingly.

The initial interpretation of the results was that the hemineglect patients can unconsciously process the neglected stimuli at the level of meaning. That is, it was thought that they have the capacity to unconsciously recognize what the meaning of the flames shown in the picture is. However, failure to replicate the result in subsequent studies encouraged a rival interpretation, according to which the patients' responses were based on relatively low-level features of the stimuli such as symmetry (Bisiach and Rusconi 1990). While this interpretation is consistent with the view that information in the neglected hemifield can be unconsciously detected and integrated into a conscious percept, the rival interpretation of the study weakens its evidential status.

A more recent study suggests that semantic features can indeed be unconsciously detected by patients with hemispatial neglect (Kanne 2002). The results of this study revealed that when a word was shown to the neglected field, the subject's performance on the subsequent search task was improved when the neglected word shared semantic affinity with the items searched for, but not when the word showed orthographic or phonetic similarity. For example, when the word presented to the neglected hemifield was 'dog', the neglect subjects did better in a task requiring them to find animal-related words than when the word shown to the neglected hemifield was semantically unrelated to the task (even when the primed word looked or sounded like the target word). This suggests that the subjects were able to unconsciously process and take into account semantic features of the stimuli.

Even emotional features of stimuli are processed unconsciously in neglect patients. Vuilleumier et al. (2002) found that when a fearful face was presented to a subject's neglected hemifield, the face still evoked responses in the parts of the brain implicated in emotional processing, in contrast to neutral faces. Furthermore, the neural activity in these brain areas was not affected by whether the patient reported being aware of the face or not. The brain responded in the same way irrespective of whether the stimulus was consciously seen. These results suggest that emotional features of the stimulus were processed by the patient even in the absence of visual awareness of said stimulus.

The hemineglect studies thus demonstrate that perceptual, semantic, and affective features can be combined into an integrative whole without the integrated information entering the hemineglect patients' consciousness. The data from hemineglect thus present a challenge to IIT.

Proponents of IIT might insist that unconsciously detected information is not integrated to a degree that suffices for consciousness to emerge. However, this sort of reply rests on the implausible assumption that only a rather high degree of integration suffices for conscious experience. This, however, runs counter to the claim that consciousness is gradable. If consciousness comes in degrees, as IIT argues, then we should expect information to be unconscious only when it lacks integration or has a very low degree of integration.

Another way for IIT to defuse the objection is to insist that hemineglect patients do not lack consciousness in the neglected hemifield but simply have a diminished degree of consciousness in the neglected hemifield. This suggestion is reminiscent of the debate about whether blindsight patients who report lacking awareness in their blind field really do lack awareness in their blind field.[18] Most blindsight researchers who draw a distinction between type 1 and type 2 blindsight argue that subjects with type 2 blindsight can detect certain features visually in spite of having extensive V1 damage that results in abnormally low levels of visual consciousness in contralesional regions of the visual field. Subjects with type 1 blindsight, by contrast, have the capacity to detect certain features visually in spite of having extensive V1 damage that results in cortical blindness in contralesional regions of the visual field.

There is considerable evidence for the view that subjects with type 2 blindsight have spared visual capacities in spite of the abnormally low levels of visual conscious-

[18] See the special journal issue published in Consciousness and Cognition, 2015 Mar. volume 32. Edited by Robert Foley and Bob Kentridge.

Fig. 8 A Kanizsa-style illusion consisting of four spatially separate, occluded black circles (pacman-like shapes) gives rise to an illusory experience of a bright white rectangle

ness. Patient report ‘seeing something’, ‘seeing something move’, ‘seeing something occluded behind a dark shadow’, and so on (Macpherson 2015; Brogaard 2015). However, unlike patients with type 2 blindsight, hemineglect patients do not make reports indicative of spared consciousness in the neglected hemifield. Absent any reports of spared consciousness in hemineglect, the envisaged objection lacks credibility.

5.2 Evidence against IIT from amodal completion

There are other ways in which the phenomenon of hemineglect turns out to threaten the claim that information integration is sufficient for consciousness. While hemineglect patients fail to be conscious of low-level features of stimuli presented to them in the neglected hemifield, these features can nonetheless contribute to the overall phenomenology of their experience. In Vuilleumier and Landis’s well-known study (1998), for example, hemineglect patients were asked to bisect (i.e., mark the midpoint of) Kanizsa-style figures similar to the one below.

To neurotypical perceivers, the display in Fig. 8 visually appears as a white rectangle hovering over four black circles. This appearance is generated via a perceptual process known as ‘amodal completion’ (Michotte et al. 1991). Although the only objects in Fig. 8 capable of making an imprint on the retina are the four pacman-like shapes, the process of amodal completion completes or ‘fills in’ the unseen parts of the stimulus to yield a unified, holistic image. The visually detected information from the pacman-like shapes is thus integrated with amodally supplied information to yield an experience of an integrated whole consisting of four black discs occluded by a white rectangle.

The process of amodal completion also seems to integrate visually detected information with information provided amodally in hemineglect patients (Vuilleumier and Landis 1998). Hemineglect patients bisected Fig. 8 as if it were a rectangle covering four black discs, even though they denied experiencing the two black pacman-like shapes on the left (see Fig. 9). Information from the neglected side (i.e., information that the subjects deny experiencing) thus seems to contribute to how the figure is consciously seen.

This suggests that hemineglect patients are able to unconsciously detect the features of stimuli presented to the neglected field and integrate them with consciously detected features. This integration results in a conscious percept of the sort depicted in Fig. 9.

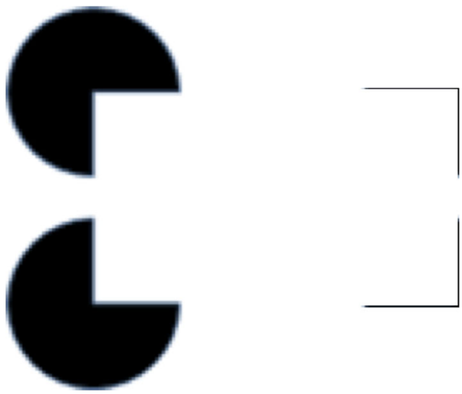

Fig. 9 The part of the image from Fig. 8 as consciously experienced by hemineglect patients. We have included the thin lines on the right in order to illustrate how a hemineglect patient seems to be experiencing the illusion despite not consciously seeing the pacman-like shapes on the right

Unconscious stimuli in the contralesional hemifield thus make a direct contribution to the conscious experience that emerges. Again, if the axioms fail to adequately capture the phenomenology of experience, then the postulates will also fail to capture how they are implemented by the brain.

One possible objection here is that since there is a phenomenological difference between Figs. 8 and 9 (that is, the right-hand side discs are experienced in the former but not in the latter), these experiments (Vuilleumier and Landis 1998) show the opposite of what we are claiming.[19] Namely, that the lack of integration (of information about the right-hand side discs) results in the corresponding lack of phenomenal experience (of the right-hand side discs).

However, our contention is that the way the right-hand side discs contribute to experience is not merely by showing up on the right-hand side of the image. They make another contribution as well: they are inducing the illusory contours of the white rectangle. In other words, it seems to us that there is something it is like to see the illusory rectangle hovering above all the discs and, whatever that experience is like, the discs on the right play a constitutive role in determining its phenomenology. That indicates that had the discs not been there, it is possible that the experience of the rectangle would not be the same (in fact, it would not be an experience of a figure hovering above the discs at all). This can be easily seen by covering the two discs on the right—no rectangle is visible then. If this is line of thought is on the right track, then the discs on the right *do* play a role in the experience of the rectangle or in what it's like to seem to see the rectangle.

Results of the bisection experiment on the hemineglect patients, by contrast, show something unexpected. This is because the patients are able to pinpoint the midpoint of the illusory rectangle despite not reporting seeing the right-hand side of the image. This is evidence that there is something in the subjects' experience in virtue of which the white rectangle appears to them in some manner, which enables them to bisect it correctly. As mentioned in the preceding paragraph, the "something" appears to be the discs on the right. It follows that the discs somehow determine what the subjects' experience is like even though they are not a part of the subjects' (reported) experience.

[19] We are grateful to an anonymous reviewer for raising this objection.

One natural way of expressing this would be to say that the information about the discs on the right is processed together or is integrated with the information about the discs on the left yielding an experience whose phenomenology we try to capture in Fig. 9. At the very least, what we try to show with Fig. 9 is that the phenomenology of the hemineglect subjects' experience of the display is not simply that of seeing Fig. 8 without the discs on the right. In addition, whatever the phenomenal difference is, it is constituted by the contribution from the two discs to the overall phenomenology. If this interpretation of the experimental results is on the right track, then we have a case where unconsciously processed information is integrated with consciously processed information to yield an experience that *couldn't have been what it is like* without integrating unconscious information.

In sum, the aggregate of the evidence presented above suggests that, contrary to what IIT says, information integration can occur in the absence of consciousness. Hence, information integration is not sufficient for consciousness.

6 Conclusion

Empirical psychology offers a wealth of insights into philosophical theories of mind and consciousness. Here we considered only a small fraction of empirical data relevant to such theories. We discussed research that sheds light on the nature of consciousness and examined what consequences the results of this research have for three target theories of consciousness that all put great emphasis on integration. The empirical evidence, we argued, shows that information integration is neither necessary nor sufficient for consciousness. Our arguments do not, however, aim to show that integration doesn't play a central role in theories of consciousness. Indeed, our conclusion is consistent with various hypotheses about consciousness including the hypothesis that the function, or role, of consciousness is to integrate information as well as the hypothesis that consciousness facilitates integration (which is what Shea and Frith (2016) suggest in light of their criticism of IIT). But while the exact nature of the interrelation between consciousness and information integration is a fascinating subject worthy of further exploration, the promise of a philosophical explanation of consciousness in terms of information integration remains unfulfilled.

References

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge, MA: Cambridge University Press.

Baars, B. J. (1997). *In the theater of consciousness*. New York, NY: Oxford University Press.

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences, 6*(1), 47–52.

Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research, 150*, 45–53.

Baars, B. J., & Franklin, S. (2009). Consciousness is computational: The LIDA model of global workspace theory. *International Journal of Machine Consciousness, 1*, 23–32.

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Science, 7*, 166–172.

Baddeley, A., & Wilson, B. (1988). Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition*, *7*(2), 212–230.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1994). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*, 66–80.

Bayne, T. (2010). *The unity of consciousness*. Oxford: Oxford University Press.

Bayne, T. (2018). On the axiomatic foundations of the integrated information theory of consciousness. *Neuroscience of Consciousness*. https://doi.org/10.1093/nc/niy007.

Bayne, T., & Chalmers, D. J. (2003). The unity of consciousness. In A. Cleermans (Ed.), *The unity of consciousness: Binding, integration, dissociation*. New York: Oxford University Press.

Bisiach, E., & Rusconi, M. L. (1990). Break-down of perceptual awareness in unilateral neglect. *Cortex, 26*, 643–649.

Blackmore, S. (2002). There is no stream of consciousness. *Journal of Consciousness Studies, 9*(5–6), 17–28.

Block, N. (1995). On a confusion about the function of consciousness. *Behavioral and Brain Sciences, 18*, 227–287.

Block, N. (1997). On a confusion about a function of consciousness. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 375–415). Cambridge, MA: MIT Press.

Brogaard, B. (2011a). Are there unconscious perceptual processes? *Consciousness and Cognition, 20*(2011), 449–463.

Brogaard, B. (2011b). Conscious vision for action vs. unconscious vision for action. *Cognitive Science, 35*(2011), 1076–1104.

Brogaard, B. (2015). Type 2 blindsight and the nature of visual experience. *Consciousness and Cognition, 32*, 92–103.

Brogaard, B., & Gatzia, D. E. (2017). Unconscious imagination and the mental imagery debate. *Frontiers in Psychology, 8*, 799. https://doi.org/10.3389/fpsyg.2017.00799.

Brüntrup, G., & Jaskolla, L. (Eds.). (2016). *Panpsychism*. New York: Oxford University Press.

Carruthers, P. (1996). *Language, thought and consciousness*. Cambridge: Cambridge University Press.

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

Chalmers, D. J. (1997). Availability: The cognitive basis of experience? *Behavioral and Brain Sciences, 20*, 148–149.

Chalmers, D. J. (1998). The problems of consciousness. In H. Jasper, L. Descarries, V. Castellucci, & S. Rossignol (Eds.), *Advances in neurologe: Consciousness: At the frontiers of neuroscience*. New York: Lippincott-Raven.

Chalmers, D. J. (2004). The representational character of experience. In B. Leiter (Ed.), *The future for philosophy*. New York: Oxford University Press.

Chalmers, D. J. (2015). Panpsychism and panprotopsychism. In T. Alter & Y. Nagasawa (Eds.), *Consciousness in the physical world: Perspectives on Russellian Monism* (pp. 246–276). New York: Oxford University Press.

Cohen, A., & Ivry, R. (1989). Illusory conjunctions inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 650.

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*, 200–227.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204–211.

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

Dretske F (1993) Conscious experience (pp. 263–283). Mind, CII

Driver, J., & Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition, 79*(1), 39–88.

Franklin, S., & Graesser, A. C. (1997). Is it an agent, or just a program?: A taxonomy for autonomous agents" In *Intelligent agents III*. Berlin: Springer.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*, 20–25.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349*, 154–156.

Hardcastle, V. G. (1994). Psychology's 'binding problem' and possible neurobiological solutions. *Journal of Consciousness Studies, 1*(1), 66–90.
Hu, Y., & Goodale, M. A. (2000). Grasping after a delay shifts size-scaling from absolute to relative metrics. *Journal of Cognitive Neuroscience, 12*, 856–868.
Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition, 8*(3–5), 489–517.
Kanne, S. (2002). The role of semantic, orthographic, and phonological prime information in unilateral visual neglect. *Cognitive Neuropsychology, 19*(3), 245–261.
Konow, A., & Pribram, K. H. (1970). Error recognition and utilization produced by injury to the frontal cortex in man. *Neuropsychologia, 8*, 489–491.
Köhler, S., & Moscovitch, M. (1997). Unconscious visual processing in neuropsychological syndromes: A survey of the literature and evaluation of models of consciousness. In M. D. Rugg (Ed.), *Cognitive neuroscience* (pp. 305–373). Cambridge, MA: The MIT Press.
Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition, 10*, 294–340.
Lycan, W. (1996). *Consciousness and experience*. Cambridge, MA.: MIT Press.
Macpherson, F. (2015). The structure of experience, the nature of the visual, and type 2 blindsight. *Consciousness and Cognition, 32*, 104–128.
Manson, N. (2000). State consciousness and creature consciousness: A real distinction. *Philosophical Psychology, 13*(3), 405–410.
Marr, D. (1982). *Vision*. San Francisco: Freeman.
Marshall, J. C., & Halligan, P. W. (1988). Blindsight and insight in visuospatial neglect. *Nature, 336*, 766–767.
McGlinchey-Berroth, R., Milberg, W. P., Verfaellie, M., Alexander, M., & Kilduff, P. (1993). Semantic priming in the neglected field: Evidence from a lexical decision task. *Cognitive Neuropsychology, 10*, 79–108.
Michotte, A., Thinès, G., Costall, A., & Butterworth, G. (1991). *Michotte's experimental phenomenology of perception*. Hillsdale: L. Erlbaum Associates.
Milner, A., & Goodale, M. (1995). *The visual brain in action*. Oxford: Oxford University Press.
Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., et al. (1991). Perception and action in 'visual form agnosia'. *Brain, 114*((Pt 1B)), 405–428.
Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.
Neri, P., & Levi, D. M. (2006). Spatial resolution for feature binding is impaired in peripheral and amblyopic vision. *Journal of Neurophysiology, 96*(1), 142–153.
Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Computational Biology, 10*(5), 1–25.
Pins, D., & Ffytche, D. (2003). The neural correlates of conscious vision. *Cerebral Cortex, 13*(5), 461–474.
Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences USA, 91*, 7398–7403.
Posner, M. I., & Dehaene, S. (1994). Attentional networks. *Trends in Neuroscience, 17*, 75–79.
Prinz, J. (2000). A neurofunctional theory of visual consciousness. *Consciousness and Cognition, 9*(2), 243–259.
Prinz, J. (2012). *The conscious brain*. Oxford: Oxford University Press.
Raffone, A., & Pantani, M. (2010). A global workspace model for phenomenal and access consciousness. *Consciousness and Cognition, 19*, 580–596.
Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience, 4*(2), 93.
Robertson, L., Treisman, A., Friedman-Hill, S., & Grabowecky, M. (1997). The interaction of spatial and object pathways: Evidence from Balint's syndrome. *Journal of Cognitive Neuroscience, 9*(3), 295–317.
Rosenthal, D. M. (1997). A theory of consciousness. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), *The nature of consciousness: philosophical debates (a Bradford Book)* . Cambridge, MA: MIT Press.
Schacter, D. L., Buckner, R. L., & Koutstall, W. (1998). Memory, consciousness and neuroimaging. *Philosophical Transactions of Royal Society London B, 353*, 1861–1878.
Shea, N., & Frith, C. D. (2016). Dual-process theories and consciousness: The case for 'Type Zero' cognition. *Neuroscience of Consciousness* 1–10. https://www.philosophy.ox.ac.uk/sites/default/files/philosophy/documents/media/shea_frith_type_0_cogn_nconsc16_oa.pdf.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during real-world interaction. *Psychonomic Bulletin and Review, 5*(4), 644–649.
Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLoS One, 3*(e1699), 1–9.
Strawson, G. (2006). Realistic materialism: Why physicalism entails panpsychism. *Journal of Consciousness Studies, 13*(10–11), 3–31.
Szczepanowski, R., & Pessoa, L. (2007). Fear perception: Can objective and subjective awareness measures be dissociated? *Journal of Vision, 7*, 1–17.
Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience, 5*(1), 42.
Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin, 215*, 216–242.
Tononi, G. (2010). Information integration: Its relevance to brain function and consciousness. *Archives Italiennes de Biologie, 148*, 299–322.
Tononi, G. (2012). Integrated information theory of consciousness: An updated account. *Archives Italiennes de Biologie, 150*, 290–326.
Tononi, G., & Koch, C. (2015). Consciousness: Here, there and everywhere? *Philosophical Transactions of the Royal Society, B370*(1668), 20140167.
Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*(2), 171–178.
Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition, 14*(4–8), 411–443.
Turnbull, O. H., Driver, J., & McCarthy, R. A. (2004). 2D but not 3D: Pictorial-depth deficits in a case of visual agnosia. *Cortex, 40*, 723–738.
Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 92–114.
Vuilleumier, P., & Landis, T. (1998). Illusory Contours and spatial neglect. *NeuroReport, 9*(11), 2481–2484.
Vuilleumier, P., Armony, J. L., Clarke, K., Husain, M., Driver, J., & Dolan, R. J. (2002). Neural response to emotional faces with and without awareness: Event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia, 40*(12), 2156–2166.
Westwood, D. A., & Goodale, M. A. (2003). Perceptual illusion and the real-time control of action. *Spatial Vision, 16*, 243–254.

Affiliations

Berit Brogaard[1,2] · Bartek Chomanski[1] · Dimitria Electra Gatzia[3]

Berit Brogaard
brit@miami.edu

Bartek Chomanski
b.chomanski@gmail.com

[1] Department of Philosophy, University of Miami, Coral Gables, USA

[2] University of Oslo, Oslo, Norway

[3] Department of Philosophy, The University of Akron, Akron, USA