Tensions between global neuronal workspace theory and recurrent processing theory have sparked much debate in the field of consciousness research. Here, we focus on one of the key distinctions between these theories: the proposed relationship between attention and consciousness. By reviewing recent empirical evidence, we argue that both theories contain key insights and that certain aspects of each theory can be reconciled into a novel framework that may help guide future research. Alternative theories are also considered, including attended intermediate-level representations theory, integrated information theory and higher order thought theory. With the aim of offering a fresh and nuanced perspective to current theoretical debates, an updated taxonomy of conscious and non-conscious states is proposed. This framework maps a wider spectrum of conscious states by incorporating contemporary views from cognitive neuroscience regarding the variety of attentional mechanisms that are known to interact with sensory processing. Whether certain types of attention are necessary for phenomenal and access consciousness is considered and incorporated into this extended taxonomy. To navigate this expanded space, we review recent ‘no-report’ paradigms and address several methodological misunderstandings in order to pave a clear path forward for identifying the neural basis of perceptual awareness.

This article is part of the theme issue ‘Perceptual consciousness and cognitive access’.

1. Introduction

Two leading theories regarding the neural basis of consciousness—global neuronal workspace theory (GNWT) and recurrent processing theory (RPT)—offer sharply contrasting views along several dimensions [1,2]. Conflicts between the two theories include the proposed neural correlates of consciousness (NCCs), the exact relationship between conscious perception and other psychological functions and the phenomenological quality of consciousness. On the one hand, GNWT argues that NCCs begin relatively late in time (more than 350 ms) after stimulus onset and rely on widespread cortical interactions, particularly involving fronto-parietal networks [3]. On the other hand, RPT posits that NCCs arise early in time (less than 150 ms) and involve localized recurrent processing within sensory cortex [4]. In terms of how consciousness relates to other psychological functions, GNWT postulates that attention is necessary for conscious perception, and that working memory is closely linked with global neuronal workspace activity [5]. RPT, however, suggests that conscious perception emerges at a more basic level and is independent from cognitive functions such as attention and working memory [6,7]. Finally, GNWT asserts that we experience one, or at best a few, conscious contents in any given moment, while RPT hypothesizes a phenomenologically ‘rich’ and multi-faceted conscious experience [8,9]. In other words, the discrepancies
between these theories permeate their scientific formulations from the microscopic to the experiential.

Here, we aim to provide a fresh perspective on the role of attention in conscious perception, reconciling certain aspects of GNWT and RPT. Current empirical evidence points towards a nuanced combination of the two theories and suggests areas ripe for future research. By highlighting other leading theories of consciousness, including attended intermediate-level representation (AIR) theory [10], integrated information theory (IIT) [11] and higher-order thought (HOT) theory [12], we explore how various pieces of existing theories may fit together in a coherent picture. On the methodological side, we critically examine experimental designs that include ‘no-report’ conditions [13] and outline common errors in the interpretation of neural contrasts between ‘perceived’ versus ‘not-perceived’ stimuli. By providing a new theoretical framework and recommended methodology, we hope to invigorate the search for the NCC.

2. The relationship between attention and conscious perception

While GNWT and RPT diverge on several levels, their pivotal point of conflict may boil down to the proposed relationships between attention and conscious perception. Almost as controversial as these two theories, however, are the very terms ‘attention’ and ‘consciousness’. It is important to operationally define these terms from the outset. Following Desimone & Duncan [14], Cohen et al. [15] and others, we refer to attention in the broadest sense as the process of selecting a subset of the available sensory information for preferential processing. This includes both top–down (endogenous) and bottom–up (exogenous) attention, as well as attention to spatial locations, sensory features, moments in time or entire perceptual objects. In some of the sections below, we discuss the importance of considering different varieties of attention and their interactions with perceptual representations, but in the current context, we stick to an all-encompassing definition. Following Koch et al. [16], we use the term consciousness to refer to a state in which contents can be subjectively experienced; ‘conscious perception’ and ‘perceptual awareness’ refer to the subjective experience of sensory-based content. This definition remains distinct from enabling states of arousal (e.g. coma versus wakefulness), which are necessary but not sufficient for experiencing conscious content.

If, as RPT suggests, conscious perception can occur in the absence of attention, then logically, NCCs could arise early in time (prior to the allocation of attention) in localized sensory cortices (without input from fronto-parietal attention networks). Additionally, as attention is necessarily a selective process, consciousness preceding that narrowing of content might indeed be phenomenologically rich and extensive. A double dissociation between attention and consciousness has been recently proposed [17] and defended [18]. This perspective asserts that attention can modulate sensory processing even in the absence of conscious perception, and conscious experience can and does occur in the absence of attention.

Contrary to this view, a single dissociation between attention and consciousness has been proposed [15]. According to this stance, while attention can modulate sensory processing regardless of whether the stimulus is consciously perceived, conscious perception requires attention. In other words, attention and consciousness are distinct psychological processes, and whereas attention can operate independently of consciousness, the reverse is not the case. Attention is necessary for consciousness. A third view has also been proposed, which argues against any type of dissociation—i.e. consciousness is attention [19].

Which of these views is more strongly supported by current empirical evidence? In 2012, when Cohen et al. [15] argued for the single dissociation view, there remained several viable counterarguments and evidence was still mixed [20]. Over the past 6 years, however, support for the single dissociation view has grown in the sense that some type of attention appears to be necessary for consciousness, even if awareness can arise in the absence of certain kinds of attention. For example, two separate studies demonstrated that even something as seemingly basic as perceiving the ‘gist’ of a scene or detecting animals in a photograph [21] can be disrupted by imposing sufficient demands on one’s diffuse attention [22,23]. Under these conditions, more than half of the subjects in these studies were inattentively blind to the essence of a scene or the presence of animals in natural images. Similarly, although claims have been made that certain visual summary statistics can be consciously experienced in the absence of directed attention [24,25], these same ensemble percepts went completely unnoticed by a majority of subjects when attention was more fully taxed [26]. Along the same lines, while Mack & Rock [27] reported reduced rates of inattentional blindness for certain stimuli, such as faces or one’s own name, subsequent studies have demonstrated robust (approx. 50%) inattentional blindness to photographs of faces, including one’s own face [28]. Even the commonly cited finding that subjects report being able to see all of the letters in the classic Sperling paradigm [29] has recently been challenged. Mack et al. [30] found that when attention was strongly focused on a distracter task, most subjects failed to notice that the entire Sperling letter array had been removed from the screen. In this case, subjects erroneously reported perceiving letters that were not physically present on the critical trial; it is currently debated whether such reports reflect confabulations during reflection or genuine reports of hallucinatory percepts [31–36]. Finally, while it has been claimed that attention and awareness exert opposite effects on the perception of afterimages [37], this finding has also recently been questioned [38]. In this recent study, both attention to and awareness of the adapting stimulus increased afterimage duration.

Critics of the single dissociation view would likely note that much of the above-cited evidence comes from studies on inattentive blindness, a method that measures what can be remembered and reported rather than what is perceived in the first place [4,6]. While we agree with the view that consciousness is distinct from memory and report, we disagree that inattentive blindness is due to a memory failure (inattentive amnesia) rather than a perceptual failure. Such a view would necessarily imply that all naïve observers consciously perceive the ‘invisible gorilla’ [39] and the ‘unicycling clown’ [40], but more than half fail to remember seeing the gorilla or the clown when questioned a few moments later. This seems very unlikely given the surprising and memorable nature of these stimuli and there is also recent empirical evidence supporting the view that inattentive...
blindness reflects a perceptual failure rather than a memory failure [41]. While it is almost certainly the case that we often perceive things and then rapidly forget about them (whether due to eye movements, shifts in attention, task interference, capacity limitations of working memory, etc.), it is also very likely that a complete lack of attention to objects/events results in a complete lack of conscious perception of those objects/events, in the first place. Other common arguments from critics of the single dissociation hypothesis rely on evidence from dual-task paradigms and studies of fragile visual short-term memory (fVSTM) [21,42–44]. In our view, dual-task paradigms cannot measure conscious perception without attention because assigning subjects a task (even if the task is secondary and deemphasized) necessarily requires that some amount of attention be devoted to these task‐relevant stimuli. Similarly, in the delayed-cueing change-detection studies of fVSTM, the best strategy is to initially diffuse attention across the entire display, because any of the items could later be cued. Again, because some form of attention is always allocated to the stimuli, this paradigm is unable to study perception in the absence of attention. A key advantage of inattentional blindness paradigms over dual-task and fVSTM paradigms is that the critical stimuli are unexpected and task-irrelevant. As we argue in a later section, it is vital to assess conscious perception of task-irrelevant stimuli (in addition to task-relevant stimuli).

Although the relationship between attention and conscious perception remains a key topic of debate and is open to future investigation, current evidence favours the single dissociation view [15,45]. Our working hypothesis postulates that (at least some type of) attention is necessary for all types of conscious perception. Importantly, however, this view is not necessarily synonymous with ‘access-only’ theories of consciousness [46], and certain aspects of RPT can still be integrated with this main tenet of GNWT. In 1995, Block [47–50] proposed a distinction between ‘access-consciousness’ and ‘phenomenal consciousness’. Access consciousness refers to experiences that are read-out by cognitive–behavioural systems and are therefore easily reported by subjects. Phenomenal consciousness, on the other hand, refers to the subjective experiences themselves, regardless of whether they are accessed, remembered, reported or not. Perhaps even a state as nebulous as phenomenal consciousness requires attention.

3. Is attention necessary for phenomenal consciousness, and if so, which type of attention?

In many cases, it is useful to think about attention as a monolithic resource or a mechanism that is invoked when task-relevant stimuli are selected for further cognitive processing. However, the contemporary understanding of attention in the cognitive neurosciences is much more nuanced, complex and multi-faceted [51–54]. The historic debate in cognitive psychology between ‘early’ and ‘late’ theories of attentional selection has essentially been settled by modern cognitive neuroscience [55–57]. The answer for top–down, endogenous attention is a resounding both (and everything in between). For example, visuospatial attention can modulate sensory processing as early as 80 ms post-stimulus [58,59], while feature-based attention can exert an influence at 100 ms [60,61], and object-based attention can affect visual processing as early as 150 ms [62–64]. Moreover, specialized neural systems for focusing attention on task-relevant target stimuli become engaged at around 200–250 ms after stimulus onset [56,57]. Importantly, attention can also modulate processing of task-irrelevant stimuli at unattended locations or time points as well as stimuli having unattended features or object properties [65–70]. Our main point here is that while it is useful to determine whether any type of attention (in the broadest sense of the term) is necessary for conscious perception, positing an attentional requirement for consciousness when so many varieties of attention exist still leaves open many questions. For example, even if attention is necessary for consciousness, when and where does conscious perception arise? Does consciousness of a visual event emerge at early, intermediate or late time points, and is it localized to posterior modality-specific brain areas, to widespread neuroanatomical networks or to a dedicated prefrontal mechanism? Positing a dependence of consciousness on (some type of) attention does not automatically imply an ‘access-only’ view of consciousness. In our current view, phenomenal consciousness is distinct from access consciousness, and each may depend on different attentional mechanisms. This view is consistent with specific aspects of both RPT and GNWT; phenomenal consciousness may be more basic, arise at earlier time-points and depend on more localized types of processing (consistent with RPT), while still being critically dependent on some variety of attention (consistent with GNWT).

What type of attention is necessary for conscious perception? In addition to the wide range of time-points and anatomical locations at which attention can bias processing, attention can be allocated in either a diffuse or focal manner and can be captured exogenously by salient stimuli or allocated endogenously according to one’s current goals [52,71]. For example, the phenomenon of ‘pop-out’ in visual search was once cited as an example of conscious perception prior to attention [17]. However, when one performs a visual search task, often the most effective strategy is to begin inspecting each array of items by diffusing attention broadly to see if the target stimulus can be easily located. Indeed, some type of attention appears to be required to allow salient stimuli to pop-out in the first place, as pop-out effects can be eliminated for arrays positioned during the attentional blink [72,73]. Along the same lines, nearly all claims that have been made regarding conscious perception in the absence of attention have included qualifiers, such as consciousness in the absence of ‘focal’ attention, or without ‘top–down’ attention, or in the ‘near-absence’ of attention or with ‘minimal’ attention [18,21,24,74]. If a particular variety of attention, or a minimal amount of attention, or some interaction between attention and perception is necessary for phenomenal consciousness (as our working hypothesis states), future research should focus on clarifying this relationship in detail (for a recent promising approach, see [75]).

It is well established that top–down attention can have suppressive effects on the processing of irrelevant stimuli as well as facilitatory effects on attended events (e.g. [14,76]). Such suppression is more likely to be engaged in situations where attention is narrowly focused on relevant stimuli,
which must be perceptually segregated from competing irrelevant stimuli. Thus, fewer processing resources will be allocated to a stimulus that is under active suppression than to stimuli presented when attention is more broadly diffused and suppression is not required. This suggests that to ensure a zero allocation of attentional resources, as specified in the taxonomy presented below, it may be necessary to employ a paradigm where the stimulus under study is being actively suppressed. A minimal allocation of processing resources may also be achieved in situations where the stimulus under study is irrelevant and fails to capture bottom–up attention.

4. An updated taxonomy of subliminal, preconscious and conscious states

In 2006, Dehaene et al. [5] proposed a testable taxonomy of various non-conscious and conscious states. This taxonomy has been extremely useful for generating new ideas for experiments, interpreting new and previous results and comparing findings across studies that have employed different manipulations of awareness. The two key dimensions of this taxonomy are bottom–up stimulus strength and top–down attention. Indeed, the majority of studies aimed at identifying NCCs have manipulated one (or both) of these dimensions. According to Dehaene et al. [5], when these two dimensions are crossed, they interact to form four basic categories of non-conscious and conscious states. For a physical stimulus that is weak (such as dim or low contrast) or interrupted (for example, by a mask or interocular suppression), two types of subliminal processing are possible (unattended and attended). For a stimulus with sufficient bottom–up strength, processing can be preconscious (if top–down attention is absent) or conscious (if top–down attention is present). In formulating and explaining this taxonomy, Dehaene et al. [5] employed an appropriately cautious strategy of restricting their claims to access consciousness, while leaving open the question of where (if at all) phenomenal consciousness fits into the picture. Of course, Dehaene et al.’s [5] taxonomy is intentionally consistent with the main tenets of GNWT [3], while contradicting RPT. For example, in the simplified diagrams depicting the types of neural processing involved in each of the four categories [5], localized recurrent processing is depicted in the preconscious category. RPT would presumably label this samecategory ‘phenomenally conscious’ because according to RPT, attention is not necessary for phenomenal consciousness [4, 6, 7].

While our current working hypothesis largely agrees with Dehaene et al.’s [5] taxonomy in terms of preconscious, access consciousness and the attentional requirements of consciousness, here we propose an updated taxonomy that includes both phenomenal and access consciousness, as well as an expanded gradient of subliminal processing (figure 1). While the vertical axis is left unchanged, we advocate for increased space across the horizontal axis in order to incorporate a larger variety of attentional manipulations. The key addition to this taxonomy, and the point at which it diverges from both the standard views of RPT and GNWT is a state of phenomenal consciousness that depends on (some type or amount of) attention.

Part of the inspiration for these additions to Dehaene et al.’s taxonomy (cells 2 and 5) comes from our own series of experiments in which the physical stimulus was equally strong in all conditions, while attention was manipulated in three steps [77–79]. During inattentional blindness (with little or no attention to the critical stimuli), we observed neural signatures of preconscious processing (cell 4). With minimal/partial attention, subjects became aware of the task-irrelevant stimulus and additional neural correlates were observed (cell 5), despite the absence of trial-by-trial reports. With full attention, when the same stimulus was made task-relevant and was reported on a trial-by-trial basis, a further series of neural events transpired (cell 6).

If this updated taxonomy is on the right track, the key contrasts that future studies should target are between preconscious and phenomenally conscious states (cells 4 and 5), and between moderate subliminal and phenomenally conscious states (cells 2 and 5; i.e. crossing the dashed lines in figure 1). So far, almost all existing studies that manipulate attention (employing inattentive blindness, attentional blink or change blindness) have made contrasts between preconscious (cell 4) and access conscious states (cell 6), skipping over the empirical investigation of phenomenal consciousness. On this horizontal axis, we lack data comparing preconscious to phenomenally conscious states (cells 4 and 5) as well as phenomenally conscious to access conscious states (cells 5 and 6). Meanwhile, on the vertical axis, paradigms that manipulate stimulus strength (via techniques such as masking or interocular suppression) face a similar limitation, and have almost always made contrasts between strong subliminal and access conscious states (cells 3 and 6); moving forward, it will be important to develop new methods for contrasting moderate subliminal versus phenomenally conscious states (cells 2 and 5 [80]) and weak subliminal versus preconscious states (cells 1 and 4). Future studies may also combine manipulations of stimulus strength and attention to allow for contrasts in both the horizontal and vertical dimensions of this space. For example, Fahrenfort et al. [81] developed a clever paradigm that involved both masking and the attentional blink to allow contrasts between strong subliminal and access conscious states (cells 3 and 6) as well as between preconscious and access conscious states (cells 4 and 6). Finally, paradigms that present stimuli at perceptual threshold, in which the stimulus is perceived 50% of the time [82], straddle the line between strong subliminal and access conscious conditions (cells 3 and 6), but could be expanded to straddle moderate subliminal and phenomenally conscious conditions (cells 2 and 5). Assisted by this new framework—and in conjunction with multiple methods—the interrelationships between attention, stimulus strength and consciousness could be more fully explored and better understood.

It is important to note that the cartoons depicting different patterns of brain activity in figure 1 are oversimplified, preliminary and serve only as temporary placeholders. It is certain that the true NCCs will prove to be more complicated than this depiction suggests. In particular, the patterns of neural activity portrayed are purely speculative at the time of this writing. Researchers may discover that the neural correlates of phenomenal consciousness include subcortical circuits [83], perhaps in conjunction with cortical circuits. Additionally, areas not depicted here, such as the claustrum [84], might play an integral role. It is also worth noting that
the patterns depicted here are relevant only to visual awareness. For other types of conscious content, the nodes depicted in occipitotemporal regions would move to cortical regions critical for formulating representations of that particular content. For example, nodes would most likely move to superior temporal regions for auditory awareness, to superior parietal regions for somatosensory awareness, to medial temporal regions for emotional awareness and so on. It will be interesting and important to determine whether the relationship between attention and conscious perception is similar for visual, auditory and somatosensory systems, as well as for integrated multi-sensory percepts [85–88]. Finally, the horizontal axis that depicts different ‘amounts’ of attention is also oversimplified. As we emphasized above, attention is a complicated, multi-faceted set of processes that undoubtedly varies along many more dimensions than are plotted here.

Despite these caveats, at this early stage of investigation, we hope that this updated taxonomy will prove useful both conceptually and practically as new experiments aimed at identifying NCCs are developed and refined. If our current hypothesis is accurate, and some type or amount of attention, or a particular interaction between attention and sensory processing is necessary for phenomenal consciousness, future research should attempt to clarify how this comes about. Exploring the space depicted in figure 1 seems like a reasonable path forward.

In figure 2, we present a few examples of how previous studies fit into the space outlined in figure 1. This list is not intended to be exhaustive, but provides key examples of studies that have employed the most common methods for manipulating perceptual awareness. Each coloured square represents one of the conditions used in the main experimental contrasts and the placements of these squares are rough estimates based on stimulus strength and attentional manipulations. While a few of the studies shown in figure 2 incorporate no-report conditions and/or task-irrelevant stimuli—an issue that we will unpack in a later section—most involve trial-by-trial reporting of task-relevant stimuli, leading to an overabundance of data for access consciousness (cell 6). It will be useful for future studies to more fully explore this space, particularly those regions most sparsely represented, such as moderate subliminal and phenomenally conscious states (cells 2 and 5). Finally, it is worth noting that this same space (minus the lines separating the six cells) may also be relevant for theories that view conscious perception as being more graded or continuous [95–102].

Figure 1. An updated taxonomy that combines key insights from global neuronal workspace theory [5] and recurrent processing theory [6]. The primary axes represent the two factors that are most commonly manipulated in experiments on perceptual awareness: stimulus strength and attention. Each cell represents a hypothesized state in which these two factors interact. Consistent with Lamme’s [6] proposal, as stimulus strength is increased, the amount of localized recurrent processing in the relevant sensory cortices (here, visual) increases. Consistent with Dehaene et al.’s [5] taxonomy, as the amount of attention increases, the amount of fronto-parietal activation and widespread information sharing increases. What distinguishes this framework from previous proposals is the incorporation of two key ideas: (i) attention is necessary for phenomenal consciousness (departing from Lamme [6]), and (ii) phenomenal consciousness relies on different attentional mechanisms from access consciousness (departing from Dehaene et al. [5]). Importantly, this framework is consistent with Prinz’s [10] AIR theory. According to the proposed framework, while most previous studies have compared cells 4 and 6 or cells 3 and 6, future studies should design experiments that enable contrasts between cells 4 and 5 and between cells 2 and 5 (indicated by the dotted lines). The star-shaped nodes indicate possible regions that might be linked with conscious visual perception.

5. Other leading theories of consciousness

So far, this paper has focused on only two of the many existing theories concerning the neural basis of consciousness. GNWT and RPT in many ways lie at extreme ends of the spectrum of existing theories, which may explain why they have been so hotly debated over the years. Several other leading theories, however, share some of the features of GNWT and RPT but combine them in unique ways. In this section, we briefly summarize three such theories: AIR theory [10], IIT [11] and HOT theory [12]. These perspectives may help
us navigate the subtleties of NCC research beyond the basic tenets of the controversial GNWT and RPT theories.

Prinz’s [10] AIR theory posits that phenomenal consciousness arises when perceptual representations at intermediate levels of sensory hierarchies are modulated by attention. From there, this perceptual information becomes available to various other neurocognitive systems. At first glance, this proposal may appear to mirror GNWT; however, AIR theory links phenomenal consciousness with transient accessibility rather than sustained global access. According to AIR theory, phenomenal consciousness critically depends on attention, but not incorporation into working memory, and is the neural equivalent of information that is broadcastable, while not necessarily being broadcasted or received by other systems. A consequence of this subtle, yet pivotal, distinction between AIR and GNWT is that conscious experience is ‘richer’ under AIR theory. Because attention can be focal or diffused, while working memory is inherently ‘focal’ (limited to only a few items at a time), we can experience more than we can report. For example, when viewing the 12 letters in the classic Sperling array, we initially diffuse our attention to the whole array (prior to the cue), thus rendering all of the letters accessible to working memory. But then, as soon as we focus our attention on a given row (after the cue), only 3–4 letters can be accessed for report. According to AIR theory, subjective experience is linked with the first step in this process, in which attention to perceptual representations makes this information available for potential selection into working memory and later stages of cognitive processing. It follows that if AIR theory is on the right track, NCCs should be found at earlier and more localized stages of processing than predicted by GNWT, but at later stages that involve more widespread cortical interactions than predicted by RPT. While AIR theory states that only intermediate-level representations can be consciously experienced, the current proposal of an expanded taxonomy remains neutral on this issue.

IIT takes a unique approach to the problem of consciousness by working from the phenomenology to neural activity rather than the other way around [11]. IIT posits that phenomenal consciousness has cause–effect power (meaning a physical substrate can both enact changes on itself and be changed by itself), is inherently integrated (irreducible to subcomponents), is structured (composed of several qualia, or subjective senses) and is distinct (each experience is differentiated from other experiences). According to IIT, the level of consciousness is related to the quantity of integrated information (denoted as $\Phi^{\text{max}}$), while the content of consciousness corresponds to the shape of the structure of integrated information [11]. Multiple structures may exist at any one time, but only the major complex (the structure with the maximum cause–effect power) forms the neural substrate of phenomenal consciousness. While IIT begins with axioms pertaining to experience rather than to neural mechanisms or cognitive processes, one of the main goals is to eventually link these core aspects of phenomenology to physical substrates in the brain. To date, IIT has focused more on explaining the neural difference between consciously experiencing anything versus nothing (termed the ‘total NCC’) rather than experiencing a particular thing (content-specific NCC) [11,16,103]. In terms of the current proposal, it may be the case that attention plays a crucial role in determining the shape of the structure of integrated information (what is in versus out of the major complex), and therefore the content that we consciously experience, while the more basic distinction between experience and no-experience may not depend on attention. Alternatively, these two aspects of consciousness may be intimately linked, because a common way to distinguish conscious from unconscious states is to assess whether any contents can be consciously experienced [96,104].

As a third example, HOTs are based on the idea that sensory representations themselves (first-order processes) are not sufficient for consciousness because such representations

Figure 2. Examples of previous studies that have explored the space depicted in figure 1. Each square represents an approximation of where one of the main conditions used in the primary neural contrasts would be located within this space. The studies included here have employed a variety of methods for manipulating awareness, such as masking [80,81,89], inattentional blindness [77], threshold detection [82], the attentional blink [81,90], binocular rivalry [91], change blindness [92] and interocular suppression [93,94]. This list of studies is not exhaustive and is biased due to the inclusion of most studies to date that have explored cell 5. If a more thorough review were conducted, the disparity between the number of studies with conditions falling in cell 6 versus cell 5 would be even greater.

previous studies

- Busch et al. [92]
- Eo et al. [94]
- Fahrenfort et al. [81]
- Frässle et al. [91]
- Herman et al. [82]
- Koivisto & Revonsuo [89]
- Pitts et al. [77]
- Pitts et al. [80]
- Sergent et al. [90]
- Wilke et al. [93]
are known to exist outside of awareness [12]. According to HOTs, an extra step of higher-order processing is required for conscious experience to arise. This extra step occurs when neural populations in prefrontal (and perhaps parietal) areas index lower-level perceptual states. In one version of HOT, this higher-order mechanism is proposed to carry out a computation analogous to perceptual reality monitoring in which the reliability of first-order representations is assessed to determine if they accurately reflect the external world in the present moment—in a type of ‘sensory metacognition’ (H. Lau, 2018, personal communication). Again, while HOTs are easily distinguished from first-order theories, such as RPT, such theories initially appear quite similar to GNWT. However, HOTs diverge from GNWT by remaining neutral about the possible relevance of higher-order processing to behaviour. While GNWT argues that consciousness serves critical functions related to cognitive control and global information exchange, HOTs only posit that awareness arises when a higher-order mechanism indexes lower-order information (which may or may not serve specific behavioural functions). In addition, HOTs predict a dedicated higher-order neural mechanism that is likely to be more spatially and temporally circumscribed compared to the sustained global ignition posited by GNWT. Finally, while HOTs typically propose a critical involvement of the prefrontal cortex in conscious awareness [105], IIT and AIR theory do not [10,11].

While consciousness researchers often focus on trying to challenge one or more of these leading theories, it remains possible at this early stage of empirical investigation that each of the major theories discussed here contains a piece of the larger puzzle. For example, phenomenal consciousness might arise at an early stage of processing (consistent with RPT), while critically depending on attention (consistent with GNWT and AIR). The interaction between attention and perceptual representations may be most closely linked with phenomenal experience (consistent with AIR), and such an interaction is inherently a second-order operation (broadly consistent with HOTs). It could even be the case that the interaction between fronto-parietal (or subcortical) attention networks and perceptual representations necessarily results in maximal complexes of integrated information (consistent with IIT).

Overall, in our view, current scientific research on the neural basis of consciousness is theory-rich but data-poor. This is especially true for non-visual sensory modalities [87]. Each of the current leading theories discussed here contains a piece of the larger puzzle. For example, phenomenal consciousness might arise at an early stage of processing (consistent with RPT), while critically depending on attention (consistent with GNWT and AIR). The interaction between attention and perceptual representations may be most closely linked with phenomenal experience (consistent with AIR), and such an interaction is inherently a second-order operation (broadly consistent with HOTs). It could even be the case that the interaction between fronto-parietal (or subcortical) attention networks and perceptual representations necessarily results in maximal complexes of integrated information (consistent with IIT).

6. Conscious perception of task-irrelevant stimuli: design details to consider

Recent proposals have advocated for the development of novel ‘no-report’ paradigms in order to more precisely distinguish NCCs from neural correlates of performing a reporting task [13,106]. This idea was partially motivated by an earlier proposal in which Aru et al. [107] warned researchers to avoid confusing neural prerequisites and consequences of conscious perception with the ‘NCC proper’ (see also [108]). The short-hand label ‘no-report paradigms’, however, is misleading because the problem is not with subjects pressing response buttons or giving verbal reports per se. Instead, the main issue is to avoid confusing NCCs with the sequence of neurocognitive events that occur after the perceptual experience itself in order to perform the task at hand. For example, a one-back task in which subjects must press a button whenever a stimulus is repeated on successive trials does not fulfill the purpose of no-report paradigms, even if stimuli repeat very infrequently and brain activity is only analysed for trials in which no report is made. The reason is that every stimulus in a one-back paradigm is task-relevant. Subjects must hold each stimulus in working memory to compare it with the next stimulus and decide whether to respond overtly or not; such additional cognitive processing is required to perform the task (whether or not there is a report), and neural correlates of this cognitive activity may be confused with the NCC proper. The challenge is to create new paradigms that allow for neural contrasts between consciously perceived versus not-perceived stimuli that are task-irrelevant.

One of the reasons why designing such paradigms is so difficult is that many of the common experimental designs for studying perceptual awareness rely on trial-by-trial reports from subjects in order to categorize the neural data into perceived versus not-perceived trial-types (such designs include threshold detection, attentional blink, change blindness, binocular rivalry and backward masking at threshold). Currently, one of the best experimental design strategies in NCC research is to keep the stimulus physically identical across perceived versus not-perceived trials, thus necessitating subjective reports to determine whether a stimulus was perceived or not. However, trial-by-trial reports tend to demand more attention and thus push the neural contrasts towards measurements of access consciousness rather than phenomenal consciousness (figure 1). To date, very few no-report paradigms have been developed [77–80,91,93,109,110] and earlier attempts were confounded by differences in stimulus competition between the seen and unseen conditions [111,112]. Future NCC research will require the invention of novel experimental designs that carefully control the stimuli for perceived versus not-perceived contrasts while enabling such contrasts for both task-relevant and task-irrelevant stimuli. In particular, if the updated taxonomy proposed in figure 1 is on the right track, task-relevant conditions may be indispensable for enabling contrasts between phenomenally conscious and preconscious processing (cell 5 versus 4) and between phenomenally conscious and subconscious activity (cell 5 versus 2). It is worth noting that report versus no-report methods do not uniquely map onto the space depicted in figure 1; it is not the case that report corresponds with access and no-report with phenomenal consciousness. While the far left-hand region of this space necessitates task-irrelevant, ‘no-report’ methods and the far right-hand region almost always correspond to task-relevant, trial-by-trial reporting methods, it may be possible to explore the middle regions both with and without reports.

But how exactly should no-report paradigms be designed, and which details should researchers consider when developing new no-report conditions? First, it may not be sufficient to present the same stimuli used in trial-by-trial reporting
conditions while simply taking away the response box and having subjects perform no task at all. Such no-task (passive viewing) conditions may result in subjects’ mind-wandering and paying no attention to the stimuli, and perhaps not perceiving the stimuli on some trials. Other subjects may create internal tasks in order to avoid boredom, such as subvocally labelling, counting or otherwise doing something cognitively with each stimulus as it is presented. No-task conditions, while potentially useful for research into resting states, mind-wandering and default mode network activity [113,114], may not be ideal for NCC research because of the lack of control over closely related processes such as attention and working memory.

If including some kind of task in a no-report paradigm is preferable, what kinds of stimuli and tasks should be considered? So far, our group has explored two different approaches. In a series of inattentional blindness experiments, in the no-report conditions, subjects performed a moderately difficult luminance detection task on separate stimuli presented concurrently with the critical stimuli [77–79]. When the critical stimulus patterns (shapes, faces or letter strings) were unannounced and thus unexpected, about half of the subjects failed to note their presence (inattentional blindness). After being queried about and shown examples of the critical stimuli, all subjects noted them during a subsequent phase of the experiment even though they were still performing the detection task on the separate stimuli. While this design avoided trial-by-trial reports, delayed reporting was required to determine whether or not the critical stimuli had been consciously perceived (reports were obtained after each phase of the experiment). Because the critical stimuli were presented hundreds of times in each phase of the experiment, it was not possible to determine how often the subjects had perceived these stimuli. Although subjects were asked to estimate how frequently they perceived the critical stimuli during the post-phase questioning, and most indicated 100+ times, such estimates are very rough and likely to be particularly inaccurate for task-relevant stimuli. An additional complication with this paradigm is that inattentional blindness only occurs for unexpected stimuli; therefore, the not-perceived condition always had to precede the perceived condition, thus creating a potential order-confound in the neural data. To circumvent this issue, we presented control stimuli (randomly oriented lines) within each phase of the experiment and always compared neural activity elicited by the critical stimuli (shapes, faces or words, formed by oriented lines) with activity elicited by these control stimuli, within each phase. This procedure was intended to subtract-out neural changes due to condition order. After this initial within-phase contrast, across-phase comparisons were made. Despite these limitations, we believe our modified inattentional blindness paradigm was a decent first attempt at enabling neural contrasts between perceived and not-perceived task-irrelevant stimuli.

In a follow-up experiment, we designed a different type of no-report paradigm [109]. Here, we presented two different categories of stimuli (shapes and colour) and presented stimuli of one or the other category (or a control stimulus with no shapes or colours) in a randomized sequence. Subjects were tasked with responding to either colour or shape on separate blocks of trials. In this experiment, there were no concurrent stimuli that subjects were required to attend to (all stimuli were randomized and presented in isolation). For blocks of trials in which shapes were relevant and colour was irrelevant, subjects had to press a button upon detecting vertical rectangle shapes and withhold responses for horizontal rectangle shapes. In other words, for the task-relevant category, stimuli required either a ‘go’ or ‘no-go’ response, whereas for the task-irrelevant category, none of the stimuli required a response. The main limitation with this experiment was that all stimuli were consciously perceived. Thus, even though neural activity associated with perceiving a task-relevant stimulus (report condition) versus a task-irrelevant stimulus (no-report condition) could be compared, we were unable to make the critical contrast between perceived versus not-perceived stimuli. This same limitation applies to two recent functional magnetic resonance imaging experiments that attempted to test the contribution of frontal cortical areas to conscious perception [110,115]. Farrowqui & Manly [115] reported deactivations of frontal cortical areas for non-target stimuli that were clearly seen. Wiegand et al. [110] showed that some (but not all) of the frontal activity observed in a perceived versus non-perceived contrast for task-relevant stimuli disappeared during a separate experiment when the same stimuli were always perceived, but were irrelevant to the task. While these studies are certainly suggestive, firm conclusions cannot be drawn regarding neural correlates of perceptual awareness without the contrast between perceived and not-perceived task-irrelevant stimuli.

To address this limitation, we designed a subsequent experiment in which we employed the same shape/colour task-relevancy manipulation as before [109], while also manipulating conscious perception via backward masking [80]. In an initial behavioural experiment, the interval between the stimulus and mask was systematically varied to determine a mask latency that would result in either 0 or 100% visibility for both types of stimuli. We then used these same mask latencies in the main experiment while manipulating task-relevancy (shape or colour) across blocks. Because the visible and invisible conditions were physically different (due to the different timing of the masks), we also included control, mask-only trials and subtracted brain activity elicited by mask-only from stimulus + mask trials prior to comparing visible versus invisible task-relevant versus task-irrelevant trials. The major limitation with this paradigm was that the stimulus + mask interaction may be nonlinear, which means that subtracting mask-only trials from stimulus + mask trials may not have completely controlled for the physical stimulus contrast of different mask timing between conditions. Also, the stimuli used in this study were not very salient (collinear shapes or individual red lines embedded within a larger grid of random white lines), and the heavy pattern masking likely disrupted early preconscious processing as well as later conscious processing, thus potentially overestimating the NCCs. Nevertheless, this paradigm was a step in the right direction as it did allow for neural contrasts between perceived versus not-perceived stimuli that were either task-relevant or task-irrelevant.

Other promising approaches for no-report paradigms have used indirect measures such as pupilometry or optokinetic eye movements to infer whether a stimulus was consciously perceived or not on a given trial. If such indirect measures can reliably distinguish between visible and invisible trials during task-relevant (report) conditions, it may then be possible to use them to make the same distinction during task-irrelevant (no-report) conditions. Frässle et al.
[91] used this approach to enable neural contrasts between binocular rivalry and a control condition (replay) during a no-report condition. Similarly, Wilke et al. [93] used generalized flash suppression to compare single-unit activity and local field potentials recorded from non-human primates while they viewed visible versus invisible targets in a no-report paradigm. Future studies should explore the possibility of using such eye-based or other physiological measures to categorize trials as perceived versus not-perceived, even for task-irrelevant stimuli.

Additional considerations that are likely to be important when designing new no-report paradigms pertain to the exact neural contrasts that are made. In contemporary consciousness research, written reports often gloss over the details about how the final data were derived. For example, several steps might be taken prior to a final figure that shows a neural difference between ‘consciously perceived’ and ‘not perceived’ stimuli. In many cases, neural activity was first contrasted between stimulus present and stimulus absent (blank control) trials prior to making perceived versus not-perceived contrasts [90,116]. In other studies, neural activity was first compared between intact and scrambled stimuli prior to the perceived versus not-perceived comparison [77,78]. Still other studies directly compared neural activity elicited by perceived versus not-perceived stimuli, both relative to baseline neural activity prior to stimulus onset [82]. These seemingly minor differences in neural contrasts prior to the key comparison between perceived versus not-perceived trials might be more important than initially anticipated, especially in terms of explaining discrepancies between experimental outcomes.

In studies of binocular rivalry, for example, there are at least three different types of contrasts that are possible. One can compare brain activity linked with percept A versus percept B, or between trials in which perception changed (reversals) versus stayed the same (stable), or between binocular rivalry versus physical alternation (replay). These three different types of contrasts are likely to reveal distinct patterns of neural activity, and it will be important to determine which (if any) are most relevant for identifying NCCs. It is also worth noting that binocular rivalry (and bistable figure) studies contain a potentially insurmountable confound in the not-perceived condition: when one stimulus (or interpretation) is not perceived, the alternate stimulus (or interpretation) is perceived. This issue may also be relevant to other manipulations of awareness that purport to achieve ‘not-perceived’ conditions. For example, with inattentional blindness, the neural contrasts are really between perceiving the distractor stimuli plus the critical stimulus versus perceiving the distractor stimuli alone. With the attentional blink, the neural contrasts are between perceiving a rapid stream of stimuli with target 2 versus a rapid stream of stimuli without target 2. With backward masking, contrasts entail perceiving the stimulus and the mask versus only the mask. In other words, even when the target is not perceived, the distractor stimuli and masks are still seen and still elicit neural activity linked with visual awareness. Thus, in many cases, even though it is tempting to label the key conditions ‘perceived’ and ‘not perceived’ or ‘aware’ and ‘unaware’, we may be subtracting out key elements of the NCC because conscious perception of something occurs during both the ‘perceived’ and ‘not-perceived’ conditions.

7. Conclusion

While many of the leading theories regarding the neural basis of consciousness appear to be irreconcilable, each may contribute key insights to the bigger picture. Determining the exact relationship between attention and consciousness may provide a path towards achieving a comprehensive theory. Future studies should fully explore the space created by various manipulations of attention and stimulus strength (figure 1), while considering a wider spectrum of theories as well as individual components of each theory. Methodologically, developing novel paradigms that assess conscious perception of task-irrelevant stimuli will likely bring us a step closer to enabling key neural contrasts between phenomenal consciousness and preconscious or subliminal processing. Such contrasts have the potential for revealing the true NCCs.

References


