Gorillas in the missed (but not the unseen): Reevaluating the evidence for attention being necessary for consciousness

Benjamin Kozuch
University of Alabama
Philosophy Department

Forthcoming in Mind and Language
Please do not quote without permission

Abstract: The idea that attention is necessary for consciousness (the “Necessity Thesis”) is frequently advocated by philosophers and psychologists alike. Experiments involving inattentional and change blindness are thought to support the Necessity Thesis, but they do so only if subjects failing to notice the target stimulus are also not conscious of it. This article uses commonsense phenomenological observations supplemented with empirical data to argue that some subjects failing to notice the target stimulus nonetheless experience its color. Since subjects not noticing the target are commonly assumed to be not attending to it, these scenarios would be instances of consciousness without attention: Instead of inattentional and change blindness supporting the Necessity Thesis, they would present counterexamples to it.
1. Introduction

Attention and consciousness surely share a close association. Consider the vivacity with which the qualities of attended objects enter the mind, and their retreat to the background when attention is removed. Many philosophers and psychologists, however, have argued the association to be tight indeed, in that there is no consciousness without attention. Key evidence for this claim comes from experimental phenomena where subjects appear to lack both attention to and consciousness of a stimulus, one example being inattentional blindness. In a well-known experiment (Simons & Chabris 1999), subjects performing the attentionally demanding task of counting basketball passes in a dynamic scene sometimes failed to notice a woman in a gorilla suit wandering through their visual field. Also thought to provide support for attention being necessary for consciousness is change blindness, experiments in which subjects are slow to detect apparently obvious changes to a scene, such as a large jet engine appearing and disappearing in a picture of an airfield (Rensink, O’Regan & Clark 1997).

Call the thesis that attention is necessary for consciousness the “Necessity Thesis,” or “Necessity” for short. If inattentional and change blindness support Necessity, it is because subjects in these experiments lack both attention to and consciousness of the target stimulus, since this co-absence could be taken to suggest that attention is necessary for consciousness. These experiments, however, also pose risk to Necessity: Since subjects not noticing the target are presumed to be not attending to it, if any such subjects were conscious of the target, it would be an instance of consciousness without attention—a counterexample to Necessity. This article argues that such counterexamples exist.

Here is the general idea: When advocates of Necessity claim that subjects are not conscious of the unattended stimulus, they often say things such as the subject is “blind” to it, or that it is “invisible.” But is this meant literally? In inattentional blindness, for instance, is it that the subject experiences none of the gorilla’s properties? Not its color, its shape, its motion? This

---


2 Quoted words from Marchetti 2012:2 and Prinz 2012:85, respectively.

3 Many passages suggest that Necessity Theorists advance a claim as strong as this: Simons and Chabris argue that “in absence of attention, visual features of our environment are...not consciously perceived”
interpretation of inattentential blindness has been described as “quite radical” (Wu 2014:157),
probably because it is at odds with our phenomenology. Consider the following: If one reflects
upon visual experience, it appears to include an experience of color throughout the visual field.
Now here is a hard question for the Necessity Theorist: If the inattentively blind subject does
not experience the gorilla’s color, what does she experience, color-wise, in the part of the visual
field where the gorilla is located? A different color? (But which color, and why that one?) No
color? (There is an absence of color in that part of the visual field?) Prima facie, neither option
seems plausible, at least no more so than one where the subject experiences the gorilla’s color.
But if inattentional blindness is to be taken to support attention being necessary for
consciousness, it is the onus of the Necessity Theorist to provide a convincing argument for one
of the other two options obtaining. Below, I argue that this cannot be done, and that it is
therefore a live possibility that inattentually “blind” subjects are sometimes conscious of an
unattended object, providing a counterexample to Necessity. I develop a similar argument in the
case of change blindness.

At this point, savvier readers are brandishing empirically based objections to the claim I
made above about the ubiquity of color experience (e.g., that visual experience is not as “rich” as
it seems; that we are subject to a “refrigerator light illusion”—such objections are addressed in
due course. In the meantime, here is the article’s layout: Sect. 2 provides an in-depth description
of inattentional and change blindness, elaborates the argument of the last paragraph, and
examines an objection to the argument. Sect. 3 considers alternatives to the idea that subjects in
these experiments are sometimes conscious of an unattended object’s color, arguing that they are
(at least) no more probable than the alternative according to which they are. And Sect. 4 offers
considerations meant to amplify the effects of the article’s main argument against Necessity.

Before beginning, a preliminary issue: As often observed (Block 1995; Chalmers 1995;
Lycan 1996), the term “consciousness” is understood in multiple ways. According to some ways

(1999:1060); Mack and Rock state that “unless certain kinds of objects are presented…the rule seems to
be that nothing is perceived consciously” (1998:163); Prinz, speaking of brain-damaged subjects who
cannot attend to left-located items, says “phenomenologically, these patients seem to be blind on the left”
(2012:82); and Dehaene and colleagues state that “a mildly masked stimulus…becomes invisible when
attention is diverted” (2006:205).

Wu refers to this interpretation as phenomenal blindness, “or the absence of any visual experience: one
is blind to features, objects, and space” (2014:156). He contrasts it with other interpretations, such as
category blindness, “where one is blind to the gorilla’s being a gorilla…that is, one does not see it as a
gorilla” (2014:158, italics mine).
of using it, subjects in the above experimental paradigms are clearly not conscious of the unattended object: Sometimes being conscious of X is equated with being able to report on X (e.g., Baars 2003; Dehaene & Changeux 2004), something that these subjects clearly cannot do. However, the thesis evaluated in this article is whether attention is necessary for experiential consciousness, where some mental state is experientially conscious if and only if there is something it is like for the subject to be in that mental state (Nagel 1974; Chalmers 1995; Block 1995). Experiential consciousness is the explicit target of some Necessity Theorists (O’Regan & Nöe 2001; Prinz 2012), arguably the target of others (Mack & Rock 1998; Simons & Chabris 1999; De Brigard & Prinz 2010; Marchetti 2012), with it being unclear in the case of other Necessity Theorists (Merikle & Joordens, 1997; Dehaene et al. 2006). Regardless of the idea’s advocacy, knowing whether or not experience occurs in absence of attention seems essential to our understanding of consciousness, and so this is the issue on which this article is meant to shed light.

2. Inattentential and change blindness reexamined

In this section, I first describe inattentional and change blindness, and why they are thought to support Necessity. Then I argue that some subjects in these experiments who fail to notice or attend to the target stimulus nonetheless experience its color. Finally, I consider a view (Rensink 2013, 2015) according to which subjects who fail to notice the target stimulus in an inattentional or change blindness paradigm are nonetheless conscious of it.

Note that the goal of this section and the next is modest: My hope is not to provide a decisive argument for some subjects experiencing an unattended object’s color, but rather for this scenario being at least as likely as its alternative (i.e., that subjects in these experiments never experience an unattended object’s color). As I explain later (4.2), establishing even this modest

---

5 More specifically, the thesis to be evaluated is whether attending to X is necessary for being conscious of X, where subject S is conscious of X if and only if S bears an experientially conscious representation that has X as its content.

6 For some textual evidence suggesting that these commentators have experiential consciousness as their target, see fn. 3.

7 For most part, I will be able to to steer clear of the tricky issue of how “attention” should be understood (see, e.g., De Brigard & Prinz 2010): My argument only requires that subjects not noticing the target also be not attending to it, something that nearly everyone in this debate agrees upon; one exception is Rensink (2013, 2015), whose view is discussed in 2.3.
conclusion substantially weakens whatever reason there is for believing Necessity, since it means that there is a notable chance that Necessity is falsified (through counterexample).

This also explains why it is of no great detriment to this article that it closely examines only two of the many experimental paradigms employed in inattentional and change blindness (one in the case of each), and that it includes no examination of other lines of evidence offered in support of Necessity (e.g., attentional blink, hemispatial neglect). Naturally, a more comprehensive (and much lengthier) critique of the Necessity Thesis would do all this. But the present article—even with its narrow scope—nonetheless stands to be quite informative: If the experiments examined below produce counterexamples to Necessity, then whatever support these other paradigms and experimental phenomena might have provided for it is nullified, since Necessity would already have been shown to be false.

2.1 Inattentional blindness, change blindness, and the Necessity Thesis

Inattentional blindness (hereafter, “IB”) occurs when a subject, preoccupied with an unrelated task, fails to notice an unexpected (and often contextually bizarre) object appearing in her visual field.\(^8\) In the experiment we consider here (Simons & Chabris 1999), subjects watched a video in which students in white or black shirts dynamically moved in a circle, throwing basketballs back and forth. Subjects were given the attentionally demanding task of counting only those passes made between white-shirted players. At one point, a woman in a gorilla suit (hereafter, “the gorilla”) walks through the circle, pausing in the center to beat her chest. After the video, subjects were asked whether they noticed anything unusual. Amazingly, only half reported seeing the gorilla.\(^9\) It is commonly (and mostly uncontroversially) assumed that those subjects not reporting the gorilla have also not attended to it; it is also commonly (but more controversially: see Wolfe 1999; Block 2001; Lamme 2003) assumed that those subjects not reporting the gorilla were also not conscious of it. Because of this apparent co-absence of attention and consciousness, IB is thought to support Necessity (Mack & Rock 1998; Simons & Chabris 1999; Mack 2003; Prinz 2007a, 2007b, 2012; Marchetti 2012).

In experiments involving change blindness (hereafter, “CB”), subjects will show difficulty in detecting highly visible and apparently obvious changes in a visual scene.\(^10\) There

---

\(^8\) For reviews, see Mack & Rock 1998; Rock 2003.

\(^9\) For experiments thought to be precursors to the Simons and Chabris experiments, see Neisser & Becklen 1975.

\(^10\) For review, see Simons & Ambinder 2005.
are numerous CB paradigms, but the one that we look at here is the *flicker task* (Rensink, O’Regan & Clark 1997). In this paradigm, subjects are presented with two alternating images that differ in some significant way (e.g., in a lake scene, the post on a pier is tan in one image, gray in the other), and which are separated by a short blank screen (the “flicker”). Subjects are asked to identify the difference between the two images. This seemingly simple task turns out rather difficult, with some subjects taking upwards of 45 seconds to recognize the change. Like in IB, it is commonly assumed that subjects in CB experiments who have not noticed the changing object have not attended to it. It is also commonly assumed that such subjects also lack consciousness of the changing object (but see Dretske 2007), though there is variation in exactly how (cf. Rensink 2010). Some commentators seem to claim that subjects who have not identified the target stimulus fail to experience it at all (O’Regan & Nöe 2001; Nöe 2005; Dehaene et al. 2006; Marchetti 2012); however, many other Necessity Theorists do not take the idea that subjects experience the object itself to be controversial, instead advancing a weaker claim, which is that such subjects do not experience the *change* in the target (Rensink et al. 1997; Simons & Rensink 2005; De Brigard & Prinz 2010). Under either interpretation, the co-absence of attention and consciousness is thought to provide support—of one kind or another—to Necessity. Below, I argue that subjects experience the target stimulus itself, leaving aside the question of whether they experience the change in it.

2.2 *Some IB and CB subjects experience the unnoticed object’s color*

The idea that subjects in an inattentional or change blindness (hereafter IB/CB) paradigm are not conscious of the unnoticed object appears belied by the character of visual experience. Consider how visual experience seems expansive and detailed (Gregory 1966/1978; Siewart 1998, Chap. 7; Carruthers 2000, Chap. 8), with phenomena such as edges, shapes, and colors appearing in each part of one’s visual field. If the visual experience of a subject in an IB/CB paradigm is like this, then she must experience *something* in the part of the visual field where the unnoticed object is. Whatever this something is, it might as well be the unnoticed object. Thus subjects in IB/CB studies experience the unnoticed object after all.

The argument just given is a good start, but is vulnerable to empirical and philosophical objections. A similar but more resilient argument can be constructed with more limited claims. It looks as follows:
P1. Some subjects in an IB/CB paradigm have the unnoticed object appear in central vision

P2. Some subjects in an IB/CB paradigm who have the unnoticed object appear in central vision are not attending to its color (Call one such subject the “Naive Subject”)

P3. The Naive Subject experiences color in each part of central vision

P4. The color that the Naive Subject experiences in each part of central vision is the color of whatever object appears in that part of the visual field

P5. In the part of visual field where the unnoticed object appears, the Naive Subject experiences the unnoticed object’s color

C. The Naive Subject experiences an unattended color

The conclusion, of course, is a counterexample to Necessity. Now I offer support for each premise.

P1: Some subjects in an IB/CB paradigm have the unnoticed object appear in central vision

One might wonder whether a subject’s failure to notice the target stimulus always results from her not fixating on it (i.e., having it appear in central vision). However, the eye movements of subjects in IB studies who notice the unexpected object are not significantly different from those who do not (Koivisto, Hyönä & Revonsuo 2004; Memmert 2006). In addition, subjects in CB studies sometimes fail to identify the changing object when looking directly at it (Mack & Rock 1998; Henderson & Hollingworth 1999; O’Regan et al. 2000; Hollingworth, Schrock & Henderson 2001), this sometimes occurring even when the subject is visually tracking its motion (Triesch et al. 2003). It thus seems that, of those subjects yet to notice the target in an IB/CB paradigm, some have it appear in central vision.

---

11 “Noticing” can be understood operationally, so that S has noticed target T iff S reports having seen T.
12 Here, to say that the object “appears” in central vision is just to say that the object is located in central vision; it does not imply that the subject consciously experiences it.
13 By “central vision,” I mean foveal and parafoveal parts of the visual field; this comprises roughly twenty degrees of our central visual field (Palmer 1999). About four degrees of the visual field fall on the eye’s fovea, the part of the retina densely packed with color-sensitive (“photopic”) cells. Parafoveal areas
P2: Some subjects in an IB/CB paradigm who have the unnoticed object appear in central vision are not attending to its color

It is commonly assumed about IB subjects that those not noticing the target stimulus are also not attending to it, whether it appears centrally or not. This is why IB is thought to support Necessity: It purportedly involves not just a lack of consciousness, but also attention. If this is right, some IB subjects have the unnoticed object appear in central vision while also not attending to it.

Initially, it might be unclear whether the same is true of CB subjects. One might think that if a CB subject fixates the target and yet fails to notice its changing, this means that she must not have attended it. However, one might argue that she could have fixated on and attended to the target and yet fail to identify it as changing, so long as she fixated on and attended to it just before or after the change, but not both; since attention plausibly must be maintained throughout the change for it to be detected. And so perhaps it is the case that, though some CB subjects fixate on and attend to the target without identifying the change, such subjects attend to the target each time that they do so.

We can, however, rule out this possibility. Consider that, if a subject attended to the target throughout the change, we would expect the subject to notice the change; this is especially probable if the subject maintained fixation on the target throughout the change. This means that, if a subject fixates on the target throughout the change without noticing the change, she probably was not attending to the target at least before or after the change, and therefore there was a moment in which the subject fixated on but did not attend to the target. The upshot here is that, if we find examples in which a subject fixates on the target throughout the change without noticing the change, then we can get the desired conclusion, which is that sometimes subjects in a CB paradigm fixate on the target without attending to it.

There are in fact such examples: In (Triesch et al. 2003), subjects were asked to manipulate virtual blocks, moving them from a countertop to a conveyer belt. On some trials, the block abruptly changed size while the subject moved it. Even when following the block with their eyes, subjects would sometimes fail to notice the change in size. And in (O’Regan et al. 2000), subjects who maintained fixation on the target throughout its change nonetheless reported the change only 40% of the time. And so it seems that it is true not just of IB, but also CB experiments, that subjects sometimes fixate on the target without attending to it.

also contain large numbers of color-sensitive cells, though their density drops off with increasing eccentricity.
For convenience, let us call one such subject the “Naive Subject.”

**P3: The Naive Subject experiences color in each part of central vision**

Reflecting upon visual experience, it seems that it very frequently (always?) includes an experience of color throughout central vision. Certainly, when attending to centrally located objects, we seem to experience color in central vision. This also seems true when attention is removed from central vision: If one “covertly scans” a peripherally located object, attending to it without fixating on it, one seems to continue to experience color in central vision. Indeed, this experience of color arguably continues even in reverie, when one is focused on events far removed in time and space (e.g., remembering a bright, balmy day in the Caribbean); perhaps the vivacity of color experience is attenuated under these conditions, but it never seems to go away. Finally, the experience of color in central vision seems to persist when one carries out the IB/CB tasks: When focused on counting the number of passes made by the white-shirted players in an IB paradigm, or when searching the alternating images for the changing object in a CB paradigm, it seems that one continues to experience color in central vision. From these phenomenological observations, we can infer that the Naive Subject likely experiences color in each part of central vision.

Certainly these observations might be mistaken. (3.1 is dedicated to thoroughly scrutinizing them.) Nonetheless, that our visual experience appears this way, both in general, and when viewing IB/CB stimuli, is some evidence for its being this way. Consider these observations to be starting assumptions, ones that might be discarded if—and only if—adequate reason is found for doing so.

---

14 Wolfe appears to have something similar in mind when saying that “even at stimulus onset, conscious perception does not seem to be restricted to the object of attention” (1999:2).

15 One might wonder whether making the above phenomenological observations would require attending to centrally located colors. However, it seems unlikely that centrally located items are being attended to when one is scanning covertly, is lost in reverie, or is carrying out the kind of attentionally demanding tasks involved in IB or CB. (But see fn. 18, where related issues are addressed.)

16 The Necessity Theorist might object here that the feeling that one continuously experiences color in central vision results from gist perception (the ability to use just a quick glance to recover things like a scene’s meaning and its distribution of low-level properties [Oliva 2005]), since this might create the illusory impression of a rich experience of color throughout the visual field. This line of argument is promising, but at this point ineffective: There is, as yet, no reason to think that gist perception does not provide an actual experience of color throughout the visual field, let alone in central vision.
**P4:** The color that the Naive Subject experiences in each part of central vision is the color of whatever object appears in that part of the visual field

This premise is based upon the following principle: Typically, the color that one experiences in each part of central vision is the color of whatever object appears in that part of the visual field. For instance, when a red ball appears in central vision under normal lighting conditions, one typically experiences redness in the part of the visual field where it appears. Or if a green light appears there, one typically experiences greenness. And so on. Naturally, there are exceptions. For instance, in *continuous flash suppression* (e.g., Koch & Tsuchiya 2006), stimuli of different colors (one static, one dynamic) are presented to the subject’s eyes, the result being that only the dynamic one is consciously perceived; in which case the color that the subject experiences in the the suppressed eye’s visual field does not correspond to the color that is actually there. Nonetheless, the principle cited above is still a good default assumption, meaning that, in any given case—including the one concerning the Naive Subject—we can assume that there is such correspondence unless we have reason to think otherwise. Whether such reason can be found in the case of the Naive Subject is a matter taken up in 3.2.

**P5:** In the part of visual field where the unnoticed object appears, the Naive Subject experiences the unnoticed object’s color

Follows from P4.

**Conclusion:** The Naive Subject experiences an unattended color

I have argued that, for some subjects in IB/CB experiments, the unnoticed object appears in central vision while they are not attending to it; that such subjects experience some color where the object is located; and that the color they experience there is that of the unnoticed object. This all adds up to subjects in IB/CB experiments sometimes experiencing colors to which they are not attending. If this is correct, then not only do IB and CB fail to function as evidence for

\[\text{17} \]

This raises the question as to the nature of this color experience. Rensink (2013, 2015) hypothesizes the experience of the unnoticed object’s color to be part of a “fragmentary” experience, a low-level visual experience consisting of colors and shapes unbound to objects (similar to an “Impressionist painting”). Wu considers both the scenario in which the color is unbound to an object (“individuation blindness”), and one in which it is bound to an object that fails to be categorized as a gorilla (“category blindness”) (2014, Chap. 5). (For another suggestion as to how the unnoticed object’s properties might be
Necessity, they instead provide counterexamples, since they present instances of consciousness without attention.  

Shortly we consider objections. But first some remarks. I acknowledge that the argument just given is probably not decisive. Knock-down arguments are typically not on the offer once we get to a point in the dialectic where phenomenological intuitions must be taken into account. Nonetheless, it is crucial to see how the gauntlet is thrown down: The Necessity Theorist will want to deny that the Naive Subject experiences the unnoticed object’s color. However, doing so convincingly requires not only providing objections to the idea that this scenario obtains, but also offering a viable alternative; that is, if the Necessity Theorist is to successfully argue that the Naive Subject does not experience the unnoticed object’s color, she must say what color the Naive Subject does experience in the part of the visual field where the unnoticed object appears, and why this scenario is more likely than the one in which it is the unnoticed object’s color. Whether the Necessity Theorist can do this is the subject of the next section. First, though, we consider an objection that tries to undermine the argument just given by denying that the Naive Subject lacks attention to the target.

---

18 Here, the Necessity Theorist might object that the Naive Subject only lacks focused attention for the unnoticed object, but not distributed attention (for this distinction, see Srinivasan et al. 2009), or that she lacks object attention for the unnoticed object, but not spatial attention for the region in which it appears. It seems unlikely, however, that subjects engaged in the attentionally demanding tasks involved in IB/CB distribute their attention across the visual field throughout the entire trial. And the idea that spatial attention is responsible for the Naive Subject experiencing the unnoticed object misaligns with a commitment of some Necessity Theorists (Prinz 2012, Chap. 3), which is that spatial attention alone is not sufficient for experiencing an object. Finally, using either line of objection probably means giving up on IB/CB supporting Necessity, since a co-absence of attention and consciousness would no longer be established.

19 An exception to this is Rensink (2013, 2015), who allows that the subject experiences the color of the target, but also claims that there is a low-level form of attention being paid (see fn. 17, and just below [2.3] for details).

20 It is worth noting that it seems possible to empirically support one of the ideas upon which the above argument leans heavily, this being that we continue to experience color in central vision throughout a
2.3 Is there attention to the unnoticed object?

It is commonly assumed—by advocates and detractors of Necessity alike—that IB/CB subjects that have not noticed the object are also not attending to it. The present article intends to ultimately join in this assumption. Nonetheless, not everyone accepts this idea, and it is worth digressing briefly to consider how such a view might affect the argument that I presented above.\footnote{I owe gratitude to an anonymous referee at \textit{Mind and Language} for pointing out that these distinctions in types of attention might present a threat to the argument given in this article.}

According to Rensink (2013, 2015), even when the target is unnoticed, it is nonetheless sometimes subject to what he calls “attentional filtering.”\footnote{Thus Rensink takes there to be more than one type of attention; see his 2013, 2015 for details. For another view in which multiple types of attention are hypothesized, see Allport 1993.} This low-level mechanism uses criteria like color or location to sift through the massive amount of visual information received at any moment, keeping only what is relevant to the task at hand. The information that is retained (“gated”) gives rise to “fragmented” visual experience, one consisting of low-level properties like colors and shapes, where these properties are not bound to objects.\footnote{Some might want to classify color representations as being “mid-“ rather than “low-level,” given that processes such as color constancy are computationally complex. In present context, this seems merely semantic.} However, merely gating an object’s properties is not sufficient for reporting the type of object it is (e.g., a gorilla)

wide range of circumstances, including when covertly scanning or being focused inward. Such experiments might look (roughly) like this: Subjects would maintain fixation while performing an attentionally demanding task, such as reporting whether two peripherally presented shapes are identical, or multiplying numbers “in one’s head”; meanwhile different colors are added and removed throughout the visual field, including at and near fixation. At certain intervals, all of the colors are removed, and the subject is asked what color was located at some spot cued in central vision. (For conditions closer to IB, subjects could be asked to carry out something like the pass-counting task before the colors are removed, with an eye-tracker being employed to determine where to put the centrally located cue.) It seems that if subjects are largely successful at reporting the colors, this at least \textit{lends support} to the idea that we continue to experience the color of centrally located objects throughout a wide range of circumstances; this in turn would support the idea that the the Naive Subject experiences the color of the unnoticed, centrally located target.

\footnote{Because the unnoticed object is \textit{experienced}, Rensink refers to this as Type-2 IB; this is to be distinguished from Type-1 IB, a type of IB in which Rensink takes there to be no experience of the unnoticed object, since it fails to be gated. Rensink takes Type-1 IB to be possible with only very small stimuli (i.e., of less than one degree eccentricity).}
or whether it is undergoing a change (e.g., from one color to another); for this, a more sophisticated kind of attention is required (e.g., “attentional holding”). The reason IB/CB subjects cannot report on the target is because this more sophisticated kind of attention is dedicated to task-related stimuli. What is important here is that, were Rensink right, then each IB/CB subject that experiences the unnoticed object might also be attending to it, in which case IB/CB cannot produce a counterexample to Necessity.  

In evaluating this proposal, it should first be noted that the existence of multiple types of attention in the way that Rensink describes is not yet widely accepted; indeed, the taxonomy that he presents is not intended to be definitive, but rather “to provide a basis of comparison for any proposed improvement” (2015:8). Still, he does provide evidence in support of his taxonomy, and so it is worth examining what it would mean for how IB/CB is interpreted if it were roughly correct. However, Rensink’s proposal might not be as helpful to Necessity as it first seems.  

In IB, for instance, there is reason to think that the gorilla’s color would not be gated. I will explain. According to Rensink’s schema, visual information is filtered according to feature (some simple property like color), location (some region of space), or eye (which eye the information came in). In the case of IB, the filtering is most plausibly done by feature, with properties such orangeness and roundness (the ball) and whiteness (the players’ shirts) being gated. On the other hand, among those properties filtered out will probably be blackness, since this is the color of the players’ shirts whose passes the subject must ignore (cf. Most et al. 2001). And so it seems that, even if there were multiple types of attention (as Rensink holds), there is still reason to doubt that a subject experiencing but not noticing the gorilla would nonetheless have attended to its color.  

---

25 It is worth noting that the argument that I gave in the last subsection could be used to provide support for Rensink’s interpretation of IB (2013, 2015), insofar as it gives reason for thinking that there is an experience of the unnoticed object.  

26 Here, the Necessity Theorist might object that such filtering is not always complete (see Treisman 1964); if so, there would be some attention paid to the experienced color, and the scenario would not count as a counterexample to Necessity. Consider, however, (a) that it would need to be the case that every time the target appeared in central vision, there was only partial filtering (if not, we would have a counterexample to Necessity), and (b) in any case where there was only partial filtering, there would now be a co-presence of attention and consciousness, which means that such a scenario would no longer function as evidence for Necessity (see discussion of CB in the next paragraph in the body of the article).  

This does not rule out, however, that there might be a co-absence of, say, consciousness of the gorilla as a gorilla (Wu’s category blindness; see fn. 17) and one of the more sophisticated kinds of
In CB, it looks more likely that the unnoticed object’s color is gated. Subjects typically search for the change by serially fixating on different parts of the stimulus (Caplovitz, Fendrich & Hughes 2008), in which case the filtering is probably done by location, with information in central vision being gated. And so it is reasonable to think that the target’s color would be gated when the subject fixates on it. Still, this is no substantial victory for Necessity. While it prevents CB from serving as a counterexample to it, it also means that a line of evidence formally thought to provide strong support for Necessity no longer does: The idea that CB supports Necessity is based on its purportedly demonstrating a co-absence of attention and consciousness, but now there is a co-presence of the two.

In summary, Rensink’s interpretation of CB looks like it might prevent it from serving as a counterexample to Necessity (if not preserving CB as evidence for Necessity), but this interpretation as yet relies on an idea not yet widely accepted, this being that there is an attentional filtering mechanism (as Rensink describes it). In any event, this approach does not seem to help in the case of IB, so we are still left with a potential counterexample to Necessity.

We just considered the possibility that, in IB/CB, there is attention to the unnoticed object, finding it of limited help to the Necessity Theorist. This was in large part because if there is attention for the unnoticed object, IB/CB no longer provides strong support for Necessity. A potentially better strategy would be to maintain that there is a lack of attention for the unnoticed object, instead contesting the idea that the target’s color is experienced. Whether this can work is what we examine next. In any event, this article henceforth adopts the widely held assumption that subjects not noticing the target are also not attending to it, so as to see what follows from this.

---

attention posited by Rensink (e.g., attentional holding); suffice to say that there are interesting issues here deserving future attention.

27 However, if Rensink’s schema, is adopted, there is a way in which CB can be interpreted so that it still provides support for Necessity; the idea here would be to hypothesize that the change in the target is neither experienced nor subject to attentional holding, meaning that there is a co-absence of attentional holding and consciousness, i.e., support for Necessity. But, as seen above, IB still seems to offer a counterexample to Necessity (even in Rensink’s schema), in which case Necessity is still subject to counterexample.
3. Alternative Scenarios

In the last section, I gave an argument concerning what the Naive Subject experiences color-wise in the part of the visual field where the unnoticed object is located. There are three possibilities:

(A) the color of the unnoticed object  
(B) no color  
(C) a color different from that of the unnoticed object

(the option argued for above)  
.reject P3)  
.reject P4)

In this section, we consider whether either of the latter two options plausibly obtain, and whether they do so in a way that itself does not involve consciousness without attention. I will argue that, at best, IB/CB is at least as likely to provide a counterexample to Necessity as it is to not.

3.1 The no-color option

Above, it was observed that our visual experience appears rich and detailed, with phenomena such as edges, shapes, and colors appearing in each part of the visual field. But some commentators argue that this impression of richness is illusory, and the Necessity Theorist might hope to use such arguments to reject the idea that the Naive Subject experiences color throughout central vision.

Dennett (1991, Chap. 11), for example, has asked us to consider the shortcomings that our visual system has, such as poor image resolution for peripheral parts of the visual field, and the blind spot that results from the optic nerve exiting from the front of the retina. Given these and other deficiencies, how is it that our visual system constructs the rich, comprehensive representations of our environment that we find in visual experience? Dennett’s answer: It doesn’t. This would be pointless: For whose benefit would the visual system be doing this “filling in,” and to what end? So, though consciousness has the appearance of unity and continuity (Dennett does not deny this), it is actually “gappy” and “sparse” (see also O’Regan 1992; Blackmore 2002).

However, this argument speaks not to the present issue. Dennett’s argumentative strategy is to point out how the visual system lacks some kind of input, then to claim that the visual system would not bother filling in the missing details (but see Ramachandran 1992, 1993; Pessoa, Thompson, & Nöe 1998). Dennett argues, for example, that we should not expect the visual system to try to make up for the low resolution of peripheral vision; more specifically, we
should not expect it to fabricate representations of the peripheral visual field equal in detail to the more fine-grained ones received from the fovea. But in the case of the Naive Subject, the visual system does not lack the unnoticed object’s color as part of its input, since the unnoticed object’s image is falling on the fovea, where the eye is dense in color receptors. And so Dennett’s arguments against the richness of experience do not call into question the idea that the Naive Subject experiences color where the unnoticed object is located.

Other commentators appeal to change blindness itself to argue against the richness of visual experience (Blackmore et al. 1995; Rensink 2000; Dehaene et al. 2006). According to them, that subjects would miss such prominent changes in a visual scene belies the visual system constructing detailed representations of one’s environment (Blackmore et al. 1995; Dehaene et al. 2006), at least ones lasting for more than a few hundred milliseconds (Rensink 2000). The apparent richness of visual experience is instead explained by the way in which information from any part of our visual field is instantaneously available (one need only perform a quick eye movement to retrieve it). So easily available is it, one gets the mistaken impression that the item retrieved was represented in experience all along, a “refrigerator light illusion” (Thomas 1999): The act of checking whether we are conscious of an item brings it into visual experience.

However, CB might result not from a lack of low-level representations—the kind that apparently compose our rich visual experience—but rather a lack of higher-level and/or conceptual representations, since it is the latter types of representation that might be required for change detection (Rensink 2000; Block 2001). Alternatively, perhaps low-level representations could potentially be used to detect changes, but they are not preserved from moment to moment; instead, they are overwritten each time the visual field changes, making difficult the kind of comparison needed for change detection (Simons & Ambinder 2005). Overall, it appears that CB can be reconciled with our visual experience being rich, at least in low-level representations like those of color (see also Nöe, Pessoa, & Thompson 2000; Cohen 2002).

Stepping back now, I think it important that we not lose sight of the paramount issue here, which is not how strong the CB-based theoretical arguments against richness are in and of themselves, but rather their strength relative to the kind of phenomenological intuitions discussed above (2.2, P3): Reflecting upon what one’s visual experience is like during the flicker task, there certainly does not seem to be an absence of color in the part of the visual field where, e.g., the gray then tan post is located (especially when looking directly at it). Indeed, it is probably because of observations like this that Necessity Theorists usually make relatively limited claims as to the kind of consciousness lacking in CB, saying that subjects fail to be conscious only of the change in the object, but at the same time allowing that they “have qualitative visual
experiences of…the image both before and after the change” (De Brigard & Prinz 2010:54; see also Prinz 2012, Chap. 3). On balance, whatever conclusions we might draw from the difficulty CB subjects show when trying to find a change in a visual scene seems unable to overpower the phenomenological intuition that there is nevertheless an experience of color in the part of the visual field where the change is occurring, especially when centrally located.

In conclusion, arguments against the richness of visual experience do not provide reason to think that the Naive Subject fails to experience color in central vision: First, Dennett’s arguments infer a lack of experience from a lack of visual input, but the relevant visual input is available in the case of the Naive Subject; second, the experimental phenomenon of CB itself can be reconciled with visual experience being rich in low-level representations; finally, CB-inspired arguments according to which the Naive Subject does not experience the unnoticed object’s color seem unable to overpower the phenomenological intuition that she does.28

3.2 The different-color option

Next we consider whether P4 might be rejected. Here, the objector allows that the Naive Subject experiences color where the target is located, but argues that it is some color other than that of the unnoticed object. There are two initially promising ways to do this.

The first is to appeal to theories of vision according to which the goal of the visual system is to produce a detailed model of one’s visual environment (McConkie & Rayner 1976; Marr 1982; Feldman 1985; Tsotsos 1987; Trehub 1991), i.e., to “create an accurate representation of the three-dimensional world and its properties” (Aloimonos & Rosenfield 1991:1250). In the case of IB, the idea would be along these lines: Prior to the gorilla entering the scene, the Naive Subject’s visual system consciously represents some object, say the back wall, as being in the place that the gorilla later occupies. When the gorilla enters, the Naive Subject’s visual system fails to detect it, and continues to consciously represent the wall as being where the gorilla is. As a result, the Naive Subject experiences the wall’s color instead of the

28 While the no-color interpretation is unlikely in the case of the Gorilla study, there are other IB studies in which it might be more plausible: Mack and Rock, for example, have found that a black circle of less than one degree eccentricity presented at fixation goes unnoticed under certain conditions (1998, Chaps. 6 & 7). While it is not clear that such subjects do not experience any color in central vision (they might very well, for instance, experience the white of the background), this is one possible interpretation of the study.
gorilla’s.\textsuperscript{29} With CB, we can picture a scenario in which the Naive Subject’s visual system has not yet registered, e.g., that the jet engine is disappearing and being replaced by the sky, meaning the visual system continues to consciously represent the engine’s color in that part of the visual field.

The second way to argue for the different-color option is by appealing to that fact that low-level visual mechanisms are thought to sometimes “fill in” a portion of the visual field with colors or patterns neighboring the area in question (Ramachandran 1992). In IB, perhaps the visual system consciously fills in the gorilla’s location in the visual field with the wall’s yellow color, this being the color that the Naive Subject experiences. And in CB, perhaps the visual system fills in the location of the jet engine with the color of some feature contiguous to it, such as the sky or tarmac.

However, in the case of the first argument, the model of the visual system it assumes—according to which the visual system builds a comprehensive, high-level representation of what appears in the visual field—is now in disfavor in vision science (see, e.g., Churchland, Ramachandran & Sejnowski 1994), in part because of empirical evidence against it (Irwin 1996; Rensink \textit{et al.} 1997; Simons & Levin 1997; Nöe, Pessoa & Thompson 2000; Rensink 2000).\textsuperscript{30} The scenario described in the second argument also seems unlikely, since filling-in is a phenomenon that mostly occurs with static, informationally simple stimuli (Ramachandran 1992), and often requires that the subject have a lengthy, uninterrupted view of the stimulus (Ramachandran & Gregory 1991); but the IB/CB stimuli are dynamic and complex, and in the case of IB, the stimuli’s background components (such as the walls) frequently become occluded.\textsuperscript{31} But I dwell not upon the considerations I just raised, as there is a more expeditious way to show that these scenarios cannot help Necessity.

\textsuperscript{29} I note that this does not imply that the gorilla and its color are not represented at all, since it could be unconsciously represented. Indeed, in cases of Type-1 IB (cases where the target arguably is actually not experienced; see fn. 24 and fn. 28), the unnoticed target is sometimes still capable of influencing the subject’s behavior (Mack & Rock 1998).

\textsuperscript{30} Note that what is being rejected here is not a visual system that produces many low-level representations, such as one that merely represents whatever color actually appears in any given part of the visual field; what is being rejected, rather, is a visual system that goes through the trouble of substituting alternative colors, doing so according to what is consistent with its high-level interpretation of the scene (e.g., that there is a wall, and not a gorilla, in some given part of the visual field).

\textsuperscript{31} In addition, it has been questioned whether the phenomenon of filling-in ever even occurs (Dennett 1992; Cornelissen \textit{et al.} 2006).
Assume that the Naive Subject experiences, not the gorilla’s color, but rather the color of whatever object the gorilla occludes. The problem is that those objects that the gorilla occludes are frequently ones to which the Naive Subject is not attending: The Naive Subject is attending to the basketball and white-shirted players, not the walls or elevator doors behind them. And in CB, if the visual system fills in (e.g.) the sky’s color where the engine is located, it is filling in the color of something to which the Naive Subject is unlikely to be attending. Overall, the different-color option probably ends up still describing a scenario involving consciousness without attention, a counterexample to Necessity.

Now, the Necessity Theorist might insist here that the Naive Subject is actually attending to the object whose color she experiences instead of the gorilla. After all, such a subject might later report that she had seen, e.g., the wall, and this could provide reason to think that she attended to it. However, subjects simply having reported seeing the wall falls far short of showing that they were attending to it. Such reports could be the result of “gist perception,” a type of visual perception that allows the subject to consciously perceive, in just a quick glance, things like the overall meaning of a scene, some of the objects within it, and low-level properties such as color (Oliva 2005). Importantly, gist perception is sometimes thought to be able to occur in absence of attention (Li et al. 2002; Koch & Tsuchiya 2007; Tsuchiya & Koch 2008; but see Cohen et al. 2011; Mack & Clarke 2012; Jackson-Nielsen et al. 2017). In addition, even if the subject attended to the wall, it is highly likely that any such attention would have occurred before the trial started. At this point, the Naive Subject’s attention might have been distributed enough to encompass many objects simultaneously, including the wall. But once the trial began, and the subject began concentrating on the challenging task of counting just white-shirted players’ passes, the subject’s attention would go from distributed to focused, at which point she probably would not continue attending to the wall. But if this response by the Necessity Theorist is to work, what is needed is not just for the wall to be attended at some point, but rather while the task is being performed; more specifically, at some time after the gorilla has appeared. And so, the current response notwithstanding, it is improbable that there is attention for a background item like the wall, especially at the moment when the target is present.

In summary, arguments for the Naive Subject experiencing a color other than that of the unnoticed object seem to either assume an antiquated view of the visual system, or to

---

32 I would like to thank an anonymous referee at Mind and Language for pointing out this possible response to be made on behalf of the Necessity Theorist.

33 For the distinction between distributed and focused attention, see fn. 18.
hypothesize color to be filled in under unlikely circumstances; and even if one of these arguments succeeds, it probably ends up establishing a case of consciousness without attention.

In Sect. 2, I argued that some IB/CB subjects experience the unnoticed object’s color. I also pointed out how, if IB/CB is to be regarded as supporting Necessity, it is not enough for the Necessity Theorist to merely offer objections to the scenario in which the Naive Subject experiences the color of the unnoticed object; we also need reason to think that one of the alternative scenarios (i.e., no-color, different-color) likely obtains, and that it obtains in a way where it does not itself involve consciousness without attention. I think that I have demonstrated that—at very least—these alternatives are no more probable than the one in which the Naive Subject experiences the unnoticed object’s color; furthermore, in the case of the different-color option, it likely still involves consciousness without attention. I conclude that IB/CB is at least as likely to present a counterexample to Necessity as it is to support it.

4. An argument and an observation

In this section, I present two ideas meant to further chip away at the plausibility of Necessity. I first argue that considerations of parsimony speak against the hypothesis that subjects in IB/CB lack consciousness of the unnoticed object. Then I point out how the seemingly neutral result that I settled for above—that there is an equal or better chance that IB/CB provides a case of consciousness without attention—is much worse for Necessity than might first appear.

4.1 An argument: Hypothesizing a lack of visual experience in IB/CB is theoretically unmotivated

The Necessity Theorist claims that subjects in IB/CB experiments who have not noticed the target stimulus are not conscious of it. However, the phenomena of IB/CB can be fully explained without introducing this controversial hypothesis, meaning that it should be left out of explanations of IB/CB, on grounds of parsimony. But then no co-absence of attention and consciousness is established, and these experimental phenomena do not support Necessity. I will explain.

When the Necessity Theorist appeals to IB/CB, something like the following reasoning is used: First, the claim is made that IB/CB subjects are both not attending to and not conscious of
the target; then there is an inference to best explanation, which is that IB/CB subjects are not conscious of the target because they are not attending to it, supporting Necessity.34

However, on what grounds does the Necessity Theorist hypothesize the subject to be not conscious of (e.g.) the gorilla? Presumably it is because the subject fails to report the gorilla. But this failure to report could just as well be considered to result from a lack of attention: Plausibly, failing to attend to the gorilla precludes the subject’s reporting its presence, regardless of whether she visually experiences it or not (cf. Wu 2014, Chap. 5). Given this, appealing to a lack of consciousness to explain the failure to report is theoretically unmotivated: Since the lack of attention on its own explains the failure to report, we need not introduce an additional assumption to our explanation (viz., that the subject does not experience the unnoticed object), an assumption that is, in addition, more controversial than the assumption that the subject is not attending to the gorilla.

To clarify, we compare the two explanations. First, the Necessity Theorist’s explanation:

In IB/CB, the subject is not attending to the target. Also, the subject fails to report the target. We can explain this failure to report by the subject’s not being conscious of the target. (Then this co-absence of attention and consciousness is used to support Necessity.)

The alternative, more parsimonious explanation:

In IB/CB, the subject is not attending to the target. Also, the subject fails to report the target. We can explain this failure to report by the subject’s not attending to the target. (No co-absence of attention and consciousness is established.)

Here is what is important about these two explanations: First, they both require hypothesizing the subject to be not attending to the target. But this is unproblematic, since (nearly) everyone agrees that IB/CB involves a lack of attention (see, e.g., Rensink 1997; Simons & Chabris 1999; Prinz 2012, Chap. 3). Second, only the Necessity Theorist’s explanation requires the controversial

34 This passage from Prinz nicely encapsulates the reasoning: “We can move toward a theory of consciousness by asking what is preventing conscious experience in neglect [a disorder that prevents the subject from noticing left-located items]. The inferior parietal brain areas that usually cause the disorder are known to play a role in allocating attention. The answer is that neglect is an attention deficit…Patients with neglect cannot consciously perceive things on the left, because they can’t attend to them” (2012:83).
hypothesis that the lack of report results from a lack of experience. It looks, then, as if the alternative explanation is to be favored for its parsimony. But if it is adopted, then there is no co-absence of attention and consciousness, and IB/CB can no longer be taken to support Necessity.

4.2 An observation: For the Necessity Theorist, a tie is a loss

Above, I settled for the modest conclusion that IB/CB is at least as likely to present a counterexample to Necessity as it is to support it. But this apparent stalemate is actually of bad consequence for Necessity. I will explain.

Consider the two possibilities between which the stalemate is occurring. In the first, the Naive Subject in the IB/CB paradigm does not experience the unnoticed object. Since the subject is also not attending to it, there would be the co-absence of attention and consciousness that Necessity predicts. This confirms Necessity, and therefore provides a little more evidence for it, perhaps increasing the probability that Necessity is true. But now consider the second possibility, according to which the Naive Subject does experience the unnoticed object (or some other unattended object). Since the subject is not attending to the unnoticed object, and since Necessity predicts that wherever there is no attention there is also no consciousness, this disconfirms Necessity; or—putting it more strongly (but no less accurately)—it falsifies Necessity (Popper 1959XX): It shows, once and for all, that attention is not required for consciousness. Furthermore, this result obtains if a subject is ever conscious of an unattended object in either IB or CB.

Putting things together: We saw earlier that in the case of both IB and CB, there is at least as much reason to think that subjects are sometimes conscious of an unattended object as to think that they never are. Combining this with observations just made, we get the result that, in both the case of IB, and of CB, there is at least as much reason to think that it falsifies Necessity as there is to think that it acts as a little more evidence for it. This asymmetry in implied consequences should significantly weaken our confidence in the Necessity Thesis.
5. Conclusion

The phenomena of inattentional and change blindness are often taken to support the thesis that attention is necessary for consciousness, but they do so only if subjects in these experiments are not conscious of the unnoticed object. In this article, I appealed to commonsense phenomenological observations combined with empirical data to argue that sometimes subjects in these experiments who are yet to notice the target stimulus nonetheless experience its color. Alternative possibilities, ones in which such subjects experience no color or a different color in the part of the visual field where the target is located, looked no more likely than the option in which it is the unnoticed object’s color that they experience, and in any event might end up presenting another case of consciousness without attention. We also considered a novel but somewhat contentious explanation of these experiments, according to which there is a type of low-level attention being paid to the unnoticed object, finding that this strategy might help block change blindness, but not inattentional blindness, from being a counterexample. We also learned that considerations of parsimony provide a reason to reject the idea that subjects not noticing the target are also not conscious of it. Finally, I argued that even the relatively neutral result that I argued for in this article, according to which subjects in these experiments are at least as likely to be conscious of the unnoticed object as they are to be not, should shake our confidence in Necessity, given this means that Necessity is likely as not falsified.
References


Marchetti, G. (2011). Against the view that consciousness and attention are fully dissociable. *Attention and consciousness in different senses, 23*.


