Perceptual Consciousness and Cognitive Access from the Perspective of Capacity-Unlimited Working Memory

Steven Gross

Department of Philosophy, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218, USA

Abstract

Theories of consciousness divide over whether perceptual consciousness is rich or sparse in specific representational content and whether it requires cognitive access. These two issues are often treated in tandem because of a shared assumption that the representational capacity of cognitive access is fairly limited. Recent research on working memory challenges this shared assumption. This paper argues that abandoning the assumption undermines post-cue-based "overflow" arguments, according to which perceptual conscious is rich and does not require cognitive access. Abandoning it also dissociates the rich/sparse debate from the access question. The paper then explores attempts to reformulate overflow theses in ways that don't require the assumption of limited capacity. Finally, it discusses the problem of relating seemingly non-probabilistic perceptual consciousness to the probabilistic representations posited by the models that challenge conceptions of cognitive access as capacity-limited.

Keywords: perception; consciousness; cognitive access; working memory; probability

1. INTRODUCTION

When I glance out my back door, I take in a rich, detailed scene: the Japanese maple, flowers of diverse shapes and shades, my daughter sitting by the deck table playing with variously-sized pieces of colored chalk. Or do I? Change and attentional blindness phenomena [1, 2] suggest a sparser visual experience, with relatively rich and detailed representations only where gaze and attention are centered, and gist and generic representations elsewhere [3, 4]—perhaps accompanied by an implicit sense of how I may retrieve more information by shifting my focus [5]. On the other hand, there is Sperling's partial-report superiority and related post-cue phenomena [6-8]. A cue after stimulus offset can significantly affect performance on recall tasks. How could that be unless rich information from across the scene was encoded prior to offset and the cue? Perhaps then visual experience is as rich as it seems to be, suggestions to the contrary indicative rather of a subsequent attentional bottleneck that renders only some of what we see available for report [9-11].

These differing responses mark a major divide among theories of consciousness, which we may sum up according to their stances on two fundamental questions:

<u>Richness Question</u>
Is perceptual consciousness rich or sparse in its specific—as opposed to its gist, generic, or general (non-singular)—representational content?

Access Question

Does perceptual consciousness require cognitive access—that is, direct availability for the formation of thoughts and guidance of voluntary actions, including perceptual reports?

Positions on these questions tend to align because parties on both sides accept that the representational capacity of cognitive access is fairly limited. Thus, those who think perceptual consciousness requires cognitive access maintain that perceptual consciousness must have sparse content as well, while those who reject the access requirement maintain that the representational capacity of perceptual consciousness exceeds—"overflows"—the limited capacity of cognitive access. Positions on these questions are also tied to methodological debates. Proponents of the access requirement ask how, if there are conscious perceptual states to which *subjects* do not have access, *theorists* may access them [12]. Those who reject the access requirement argue that their opponents fail to distinguish mechanisms implicated in consciousness and those implicated in access, introspection, and report [13].

Recent research on working memory, however, challenges the shared assumption that the representational capacity of cognitive access is fairly limited. This paper examines the bearing of this challenge on arguments regarding consciousness. After reviewing the challenge, we argue, first, that abandoning the assumption undermines "overflow" arguments based on post-cue performance. We then show how it dissociates the rich/sparse debate from the access question. Next, we explore how one might try to reformulate overflow theses in ways that don't require the assumption of limited capacity. Finally, we discuss the problem of relating seemingly non-probabilistic perceptual consciousness to the probabilistic representations posited by the models that challenge conceptions of cognitive access as capacity-limited—and indeed posited by much research on perception and memory more generally.

This last topic is a specific instance of a larger theme: how to relate such notions as perceptual consciousness and cognitive access to psychological constructs and neurophysiological mechanisms so as to render them amenable, so far as is possible, to experimental inquiry. That this is not obvious for perceptual consciousness is not news; indeed, it is what these debates are about. But, as we will see, challenges arise for cognitive access as well. As in scientific inquiry generally, such challenges cannot be met simply by defining terms in advance. Instead, inquiry proceeds conceptually and empirically in tandem to uncover the most fruitful ways of delineating the subject matter. That said, the focus of this paper leans towards the conceptual—more specifically, towards laying out possibilities in conceptual space in light of recent empirical results.

2. OVERFLOW AND WORKING MEMORY

Research on working memory bears on these debates because it's invoked to provide substance to talk of cognitive access. For example, Block [9] argues that perceptual consciousness is associated with sensory memory (an earlier store supported by locally recurrent processing in sensory cortex [11]), and cognitive access is associated with later working memory (supported by globally recurrent processing connecting activity in sensory areas with activity in parietal and prefrontal cortices). The overflow claim thus becomes that sensory memory can represent a larger number of specific items than can working memory.[1]

Presence in working memory suffices for cognitive access because such representations are made available to cognitive systems—i.e., are poised for use by them—without need for further processing. The access is thus direct in not requiring intervening processes or mechanisms. This is not so, on Block's view, for representations in sensory memory, which must be transferred first to working memory in order to be available to cognitive systems (cf. [15]); they are thus merely access*ible*: each *could* enter working memory—though, given working memory's smaller capacity, not all can at once. Distinguishing accessibility and access removes a source of possible confusion in previous discussions of access consciousness and phenomenal consciousness [16, 17].

Is presence in working memory also *necessary* for cognitive access? It might seem that overflow arguments are committed to an affirmative answer; otherwise, that there's a capacity limit to working memory can seem irrelevant. But perhaps all that is needed is that presence in working memory is necessary for cognitive access *for the representations at issue in a particular case*, or that at least, in the case at issue, no other way of achieving cognitive access seems plausible. That said, the identification of working memory and cognitive access is not uncommon and simplifies initial exposition. Below we note a challenge from Carruthers [18] even to the weaker formulations.

We need a third term for a further status. Not all representations in working memory are used—or, as the vernacular would allow, accessed—by the cognitive mechanisms to which they are directly available. We avoid confusion by speaking of retrieval. Accessibility, access, and retrieval then correspond respectively to presence in sensory memory, presence in working memory, and use by cognitive mechanisms implicated in reasoning, guidance of voluntary action (including report), and the like.

This three-way division, in one form or another, is common among proponents and deniers of overflow—as is the claim of declining capacity. They differ rather in where to locate perceptual consciousness: in sensory memory, in working memory, or elsewhere. Perceptual consciousness then overflows cognitive access only if it's associated with an earlier, higher-capacity store. For example, whereas Block [9, 10] and Lamme [11] associate perceptual consciousness with sensory memory, Dehaene and Changeux [19] associate it with global broadcasting, including the globally recurrent processes that support working memory.[2] Higher-

---

[1] Some terminological notes: Block uses Neisser's [14] term 'iconic memory' for the earlier store. We use 'sensory memory' to remain neutral on issues of representational format. In speaking of specificity, we mean both that the representation is singular as opposed to general ('*That* is an F' vs. 'There is at least one F present') and that items are represented at a particular level of specificity held fixed across the capacity comparison (for example, as the letter *F*, as opposed to—more generically—as a *letter*). We assume that perceptions represent objects, events, or other such entities as having features, but to reduce verbiage we sometimes use the term 'item' to cover both the entities and the features.

[2] Recent results suggesting unconscious representations in working memory [20] have led Dehaene and colleagues to consider whether further conditions might need to be met [21]. Note also that Dehaene's [22] contention that only a single item can occupy consciousness at a time might suggest that, at least here, he identifies consciousness with retrieval.

order thought theorists require retrieval and specific further processing: one must form an appropriately caused thought about one's having that perceptual experience [23, 24].[3]

The claim that working memory is capacity-limited—able to represent about four items at a time, including items chunked from other items—has been a dominant view. Its main support comes from set-size effects. In a variety of tasks, performance implicating working memory remains more or less steady for stimulus set-sizes averaging four or below but rapidly deteriorates for set-sizes greater than four [26]. A common explanation is that working memory's representational limits reflect its containing a limited number of "slots" for representational vehicles, each capable of representing one item [27, 28].

The assumption of a "slot" model in post-cue overflow arguments is particularly clear in Lamme and colleagues' work (and Block's use of it). They combine elements of Sperling's post-cue paradigm with Luck and Vogel's [27] change-blindness task. In a typical task, participants are presented eight variously oriented rectangles circularly arrayed around a fixation point, followed by a probe array with orientation altered at most at one location [29]. Participants issue a change/no change report concerning the orientation of the relevant rectangle. Conditions vary with the timing of a cue that indicates the position of the rectangle at issue: the cue can appear either 10ms after target display offset, 1000ms after, or 1500ms after along with the probe display. Participants' hit rates decline with the latency of the cue. To calculate the number of items retained in memory in each condition, Lamme's group uses Cowan's formula for calculating capacity K from hit/miss rates (h and r) and set-size (n) in such tasks: $K = (h + r - 1) \times n$. They, and Block following them, interpret the resulting numbers—roughly 8, 6, and 4—as indicating a succession of three stores of declining capacity, the last corresponding to working memory. This interpretation relies on the slot model because it's what underwrites Cowan's formula. The calculation embodies the assumptions that items are either stored in memory or not in an all-or-nothing fashion; if they are stored in memory, reports will be accurate; but if they are not, then participants will guess. Thus, the hit rate $h = K/n + [(n-K)/n]g$ and the miss rate $r = K/n + [(n-K)/n](1-g)$—from which the formula for capacity falls out trivially.

The slot model's dominance, however, has waned over the past decade [30]. A switch from binary to continuous tasks results in findings difficult for it to account for. Recall variability, for example, increases with set size, with greater precision for salient items at a cost to the precision of recall for other items [31, 32]. Such results have motivated models in which a limited continuous resource is distributed among representational vehicles (or aspects of one vehicle). Set-size phenomena are then explained in terms of decreased precision and increased interference among probabilistic representations as the number of represented items increases, rather than in terms of a limited number of slots [33]. Some more recent models even drop the assumption of a limited vehicular resource [34-36]. For example, Oberauer & Lin [36] explain set-size phenomena and other results just in terms of interference and attention in memory (cf. [37]). In a range of tasks, limited resource and interference models outperform both simple slot models and their more sophisticated descendants devised to accommodate continuous task results. Implementations of non-capacity-limited models also provide a better fit to neural data [38].

The availability of these models challenges the argument from post-cuing to overflow by offering an alternative to the slot model on which the argument's capacity calculation depends.

[3] Some higher-order thought theorists may deny, like Carruthers, that retrieval of percepts requires passage through working memory. Cf. Brown [25].

Moreover, Gross and Flombaum [39] point to various ways, drawn from the literature, that Lamme and colleagues' results might be explained from the non-slot perspective. The timing of the cue can affect, for example, the prevalence of correspondence (or, swap) errors [40-42], the cue's ability to indicate inputs that can be treated as noise [43], and the redeployment of attentional resources that selectively protect representations from degradation and interference [44]. *Performance* limits may thus be explained without hypothesizing a *capacity* limit—a limit to the number of items working memory can represent. (Indeed, even if one adds a capacity limit to such models [45], their other resources for explaining performance continue to provide the basis for challenging overflow arguments, so long as the capacity is sufficiently high.)

The study of working memory continues to evolve (see [46, 47] for recent defenses of discrete slot models). But suppose some such alternative to slot models pans out, so that the overflow argument is undermined, not just challenged. We now discuss how these models dissociate the richness and access questions, whether they admit reformulated overflow theses, and how the probabilistic representations central to them relate to apparently non-probabilistic perceptual consciousness.

3. DISSOCIATING ACCESS AND RICHNESS

The richness and access questions are fundamental to debates concerning consciousness and often treated in tandem. Proponents of overflow attribute rich specific representational content to perceptual consciousness and deny that it requires cognitive access. Opponents of overflow require cognitive access for perceptual consciousness and attribute to it sparse specific representational content. Arguments on both sides assume limited representational capacity for working memory and thus cognitive access. But once this assumption is dropped, the richness and access questions dissociate. Working memory capacity may be high or unlimited, but it doesn't follow one sees all items presented; and, if one does see all items, it doesn't follow they are all encoded in working memory.

Some misunderstandings could obscure this point. One might object that 'rich' means *richer than something* [48]; and, in this case, it means *richer than what working memory represents*—so that questions of richness and access are not dissociable. But that's not how the richness question is usually posed. The question rather is whether perceptual consciousness is less rich in specific representational content *than one might have thought*. Thus, proponents of sparseness [49] often present their view as involving a cognitive illusion: a false belief *about* perceptual consciousness that, even after reason to the contrary is accepted, persists—if not as an endorsed belief, then at least as a disposition to believe or an "intellectual seeming." Others doubt that there is a pre-theoretical tendency to believe perception is rich, so that there would be no cognitive illusion should perceptual consciousness turn out to be sparse [5, 50, 51]. But even if this is so, we can parse 'than one might have thought' in terms of what proponents of richness think in comparison to what proponents of sparseness think.

A different source of doubt about dissociation is the following line of thought: if working memory is capacity-unlimited, then all items can be represented there, and therefore all items should be represented in perception as well—so, perception cannot be sparser than it's thought to be. But from capacity being a certain size, nothing yet follows about how much is used on some occasion. In particular, having a capacity-unlimited store doesn't guarantee that all items get encoded in that store. As the next section will underscore, on at least one formulation, overflow and the rejection of the access requirement require only that, on some occasion, more is in fact consciously seen than is accessed, whatever the capacities. Limited capacity comes into the arguments regarding access, richness, and overflow only as setting an upper bound on what's encoded into working memory. Abandoning the assumption of limited capacity removes the upper bound and thus undermines arguments that depend on it. But it does not tell us how many items are in fact encoded in any store. Unlimited-capacity working memory is in itself neutral on questions of access, richness, and overflow. Analogous remarks hold for sensory memory's capacity and what in fact gets encoded there on any particular occasion.

Consider Sperling's [6] results. Briefly presented with a 3x4 array of alpha-numeric characters, participants on average correctly recall about 4-5. But post-cueing a row within 300 ms of stimulus offset enables participants to correctly recall all, or almost all, the characters in that row. To proponents of overflow, this—along with participants' reports of having seen all of the characters, including those they do not correctly report (but see [52])—suggests that participants form conscious representations of all, or almost all, the characters as the specific characters they are, only some of which then get transferred to working memory for report. Previous critics countered that participants may have only *unconsciously* represented all, or almost all, the characters as the specific characters they are, with the cue serving to bring a subset to consciousness or to allow for their being consciously represented with that level of specificity [3, 4, 53, 54]. Gross and Flombaum [39] suggest instead that, since in this paradigm (unlike Lamme's group's), the post-cue occurs within the post-dictive window [53], the location of the post-cue may affect the integration of the sensory signal in a way that determines which characters and features get represented at all—consciously or unconsciously. (They note several ways this could happen—for example, by setting the starting point for a greedy algorithm. The more general point is that the exigencies of processing can dictate, in a task-relevant way, whether fewer items are encoded well rather than more encoded poorly.) Thus, while previous critics maintained that all items were encoded in sensory memory, Gross and Flombaum allow the possibility that only some are—even while *also* arguing that working memory may be capacity-unlimited.

4. REFORMULATING OVERFLOW?

Block [10] explicitly formulates his overflow claim in terms of the relative capacities of perceptual consciousness and cognitive access. Capacity-unlimited conceptions of working memory directly challenge this claim. A natural reply is to reformulate overflow in other terms, so that it does not depend on a capacity-limited conception of working memory [55-58]. The previous section contains an obvious suggestion: since capacity need not be exhausted, reformulate overflow in terms of the comparative number of items in fact encoded on some occasion. The challenge then is to find evidence that on some occasions not all items represented in perceptual consciousness are encoded in working memory. Note that Gross & Flombaum's [39] alternative explanations of post-cue effects do not invoke a decrease in number of items actually encoded across stores.[4] But even if there were a decrease across stores, it would not follow that perceptual consciousness should be identified with the earlier store. In particular, once the upper-bound on capacity is removed, the later store may still represent a sufficiently large number of items, even if on that occasion fewer than the earlier store, to support rich

[4] They allow that fewer items could be encoded in working memory *at the time of retrieval* than represented in perceptual consciousness; for the post-cue could cause non-attended representations to be *removed* from working memory [43]. But, since removal presupposes prior presence, this would not amount to an overflow of working memory.

perceptual consciousness. An appeal to other considerations would then be necessary to support an identification of the earlier store with perceptual consciousness [10].

A different reformulation strategy would argue that perceptual consciousness overflows something other than the number of items represented in working memory. The rest of this section canvasses some suggestions, offering commentary and noting possible challenges in each case. The suggestions are that perceptual consciousness overflows (a) report; (b) *genuine* report (as opposed to guessing) and thus retrieval; or (c) *correct* retrieval and report, or correct *access*. (Gross & Flombaum [59] discuss some other options.) The last suggestion returns us to the question of whether presence in working memory is necessary for cognitive access and also raises issues concerning probabilistic representations, developed further in section 5.

(a) ***Overflow of report***

Even if the representational capacity of perceptual consciousness does not exceed that of working memory, it may exceed participants' capacity to report what they see—even assuming normal capacity to report (for example, no aphasias). While this may be, it's important to note that neither Sperling's nor Lamme's group's results establish the claim—at least if what's meant is just entering a claim, regardless of accuracy. Sperling's full report condition required participants to fill in characters in all spaces on the recall grid, guessing where necessary, so that the number of items reported always equaled the number of items displayed. Lamme and colleagues' paradigm involves a choice about a single item. Neither shows that participants can report, in this weak sense, fewer items than they see. Moreover, reporting in this sense—that is, responding in some manner to the query—does not necessarily involve retrieving, as the possibility of guessing shows.

(b) ***Overflow of genuine report***

An overflow thesis that participants see more than they can *retrieve* bears more directly on theories of consciousness—even if it only threatens those that require retrieval. (Accepting larger capacity working memory might, for some, render retrieval a more attractive requirement for cognitive access and/or perceptual consciousness.) Moreover, it seems it could be assessed by considering whether perceptual consciousness overflows *genuine* report—where a genuine report is one guided by information actually retrieved from working memory, regardless of accuracy. How might genuine reports be distinguished from mere guessing? Participants' own sense of whether they guessed would provide a subjective basis. But participants may not have reliable introspective access to the causal source of their responses: a judgment of having guessed may actually reflect low confidence in the report itself (and the threshold for this judgment may be task-relative). More objective marks of guessing might be found in participants' responses themselves: all else being equal, guesses should be evenly distributed across candidates, unlike genuine reports. Indeed, slot models standardly mix genuine reports and guesses, as we saw with the calculation of Cowan's K. But in model comparisons for a variety of tasks, alternative non-slot models without guessing outperform or equal models that incorporate guessing [36, 60].[5]

(c) ***Overflow of correct access***

[5] However, if items are *removed* from working memory, then they cannot be retrieved. Note also that guessing does not entail that an item is *not* encoded in working memory. For example, low confidence may suggest that retrieval is not worth the cost.

More directly assessable is how many reports are *correct*, suggesting two further overflow claims, regarding correct retrieval and correct access respectively. That we can correctly perceive more than we correctly retrieve and report is plausible, since—if for no other reason—retrieval and report can introduce their own errors. But the overflow claim regarding correct access threatens a larger range of theories of consciousness, so we will focus on it: do we correctly consciously perceive more than is correctly represented in working memory—and thus more than the number of correct representations to which we have cognitive access? That representations in sensory memory are more precise than those in working memory, so that the latter can be said to be more error-prone than the former, is plausible and built into some models [36]. But this leaves open which store should be identified with perceptual consciousness. Moreover, there are several complications that arise regarding this "correctness" overflow claim and possible arguments for it—to which we now turn.

(i) *Perceptual accuracy and working memory*

Consider a straightforward way one might argue that perceptual consciousness is more accurate than working memory: participants surely are often more accurate when they report while perceiving as opposed to after offset. Why does that not suffice? Here are two issues. First, the better performance may reflect perceptual *exploration*, with gaze and attention shifting over time—that is, it may reflect, not one state of perceptual consciousness, but a succession of them. This raises delicate questions concerning the individuation and temporal extent of states of perceptual consciousness [61, 62]—questions that do not arise for typical tasks involving brief exposures. Second, on the views we have been considering, reports made while perceiving still reflect transfer to working memory, which is deemed necessary for cognitive access and thus retrieval and report. If the necessity of transfer into working memory is assumed, the difference in performance cannot reflect a greater tendency to error in working memory *per se*.

The assumption that perceptual reports require presence in working memory has been challenged—but in a way that does not aid the straightforward argument for the overflow of correct access. Carruthers [18] argues that cognitive access should be identified, not with presence in working memory, but with presence in the global workspace. The latter, on his view, encompasses not only working memory but perception as well. Presence in the global workspace is achieved when representations are broadcast to a wide range of cognitive mechanisms—that is, available to them for retrieval. This occurs when representations are sufficiently activated to generate recurrent loops among relevant regions of sensory, parietal, and prefrontal cortices. Perceptual representations and working memory representations differ, on Carruthers' account, in virtue of the source of their activation: percepts receive both bottom-up support from the signal and at least minimal top-down support from attention, while working memory representations are supported by top-down attention alone. Carruthers thus denies that cognitive access and retrieval requires percepts to pass into working memory. His explanation of post-cue performance therefore cannot rest solely on working memory's having a capacity limit. He adverts to the time-course and serial nature of report. Conscious percepts are directly available for report but, after offset, not present long enough to be reported. At that point, they can be reported only if they are maintained in working memory—and *then* working memory's limited capacity (which he accepts) enters the explanation.

Carruthers' identification of cognitive access with presence in the global workspace instead of presence in working memory can at first seem merely a redefinition or verbal maneuver [55]. But the various empirical commitments incurred in his development of the view show that this is not so. For example, Carruthers [63] assumes a sensory recruitment account for working memory ([64]—but see [65]); his activation account of working memory, as it stands, is challenged by evidence for "activity-silent" working memory [66]; and his view predicts that a greater number of items is in principle decodable from global recurrent loops than Block, Lamme, and Dehaene would predict.

Because Carruthers accepts capacity-constrained working memory, he agrees that perceptual consciousness overflows working memory.[6] But because he identifies cognitive access with presence in the global workspace, distinguished from working memory, he denies that perceptual consciousness overflows cognitive access.

Recall that the second issue raised above for the straightforward argument assumed that working memory is implicated even in perceptual reports made while perceiving. If cognitive access is not identified with some further store conscious perceptions must pass into, this issue is removed—but not in a way that supports the conclusion of the straightforward argument. For if conscious percepts are in the global workspace, then perceptual consciousness does not overflow cognitive access; and, for the same reason, correct conscious perceptions do not overflow correct representations to which we have cognitive access.

(ii) *Probabilistic accuracy*

Assessing the "correctness" overflow claim also faces another complication, concerning the probabilistic nature of the representations posited in non-slot models of working memory. When are such representations correct [67]? A representation to the effect that a certain item is red is correct just in case the item is red. But what of representations that assign credences across a space of candidate colors? Researchers in formal epistemology have explored measures of verisimilitude, or closeness to the truth, for subjective probability distributions [68, 69]. But these measures do not amount to correctness conditions. One might consider a probability distribution correct just in case the hypothesis to which it assigns highest value is correct. But, first, this threatens to treat the probabilistic representation as in fact a representation of just that hypothesis. Second, some theories of representational content take how a representation is used to partially determine its content [70]. But not all models use decision rules that extract the item assigned highest credence. (In principle, the same representation may be subject to different decision rules by different computations.)

Suppose probabilistic representations do not so much as have correctness conditions (unlike the representations they embed). Then it might seem trivially the case that perceptual consciousness can have more correct representations than working memory. For, if the representations in visual working memory are probabilistic and thus have no correctness conditions, then there are no correct representations in visual working memory. (This assumes having correctness conditions is necessary for being correct—indeed perhaps necessary for being

[6] A variant of Carruthers' view would replace his conception of working memory with one that is capacity-unlimited. Indeed, this provides one way of developing Gross and Flombaum's [39] remark that, not only does post-cue performance not provide evidence for a succession of stores of declining capacity, but perhaps one need not posit a succession of stores at all. On this view, there would be but one store with representations differing in their properties, including their functional roles, in virtue of what sustains their presence there. Carruthers [63], however, recognizes a visual short-term memory, distinct from working memory, capable of storing representations for 2 seconds without top-down attention.

the kind of thing that can be correct or incorrect, at least in the sense that is at issue.) So, having at least one correct representation in perceptual consciousness would suffice for "correctness" overflow.

This argument assumes that perceptual consciousness is not *itself* probabilistic. But if it is not, that would suggest a more fundamental reason—capacity considerations aside—for proponents of the alternative models not to associate perceptual consciousness with working memory. Exploring this requires a more general discussion of how to relate the apparently non-probabilistic representations of perceptual consciousness to the probabilistic representations posited in recent models of working memory and perceptual processing. The next section takes this up.

5. PERCEPTUAL CONSCIOUSNESS AND PROBABILISTIC REPRESENTATIONS

It is often remarked that perceptual consciousness seems to represent the world as populated by particular objects and features, not to present us with ranges of candidate objects and features each assigned some credence [71-74]. But the alternative models of working memory all trade in probabilistic representations. It might seem that this difference would by itself disqualify working memory as the locus of perceptual consciousness. But those who would locate it instead in earlier sensory stores should be wary of pressing this argument too quickly. For much recent work on perceptual processing posits probabilistic representations more generally, including in early stages of processing [75]. Perhaps then perceptual consciousness has a much more complex and even less obvious relation to the states posited by our best accounts of perceptual processing and memory than the simpler identifications dominating current debate would have it [58].

Here, we briefly lay out some alternatives for relating perceptual consciousness and probabilistic models of perception and short-term memory in light of this apparent mismatch. The strategies discussed would deny that perceptual consciousness is not probabilistic; deny that probabilistic accounts of perception and/or working memory require probabilistic representations; or show how non-probabilistic perceptual consciousness and models that trade in probabilistic representations can after all non-mysteriously co-exist. The aim is to identify options and obstacles, not to settle the matter.

(a) ***Probabilistic perceptual consciousness***

The first strategy embraces probabilistic perceptual consciousness and argues that resistance to this idea stems from a mistaken conception of what it would have to be like—how it would be phenomenologically—to perceptually represent a probability distribution (or density).

If someone whose default is non-probabilistic conscious perception presses the 'what is it like?' question for probabilistic consciousness, the presumption is that it must be phenomenologically different from the non-probabilistic conscious perception they take themselves to have. This suggests the principle Morrison [67] labels "confidentialism": if two experiences have the same phenomenology, they assign confidence (credence) in the same way. But what sort of phenomenological difference could assigning positive credence to more than one feature value make? One answer would require that each feature value assigned a positive credence show up independently in consciousness with its own distinct phenomenological correlate. It is phenomenologically implausible that our experience is like this. Suppose perception assigns credences across a color space. We should then in effect see rainbows at all locations. That we don't tells against probabilistic conscious percepts if this requirement is accepted.

But "confidentialism" needn't be developed this way. Representational aspects of a conscious percept may make a phenomenological difference without each represented feature having an *independent* phenomenological correlate. The variously represented colors can make a difference to how things look with respect to color without each color itself showing up in consciousness. The question remains what kind of difference they make. A natural worry is that, if multiple colors do not present themselves independently in consciousness, then just one does—perhaps the color assigned the highest credence, or perhaps the mean. But then the percept will be phenomenologically indistinguishable from a non-probabilistic representation of that color. (Indeed, an analogous worry arises even if multiple colors *do* show up.)

A reply is that this misconstrues the distinctive phenomenology of probabilistic representations. Morrison [67—see also 71, 76], for example, has us consider cases such as crowding in the periphery, where our experience is not of any specific number of bars. Similarly, a conscious probabilistic representation of color need not have the same phenomenology as a representation of any one color at that level of specificity. Examining Morrison's reasons for considering such representations probabilistic, however, points to another candidate phenomenologically-indistinguishable non-probabilistic perception, in apparent violation of "confidentialism."

Morrison argues that such cases are naturally *described* as involving perceptual assignments of credence—for example, as its *looking* more likely that there are five bars than three—and, moreover, that positing probabilistic perceptual consciousness best explains probabilistic *belief* based on such perceptions. The obvious alternative is that our characterization of how things look and our assignment of credences in belief both reflect post-perceptual processing (including possibly processing based on non-conscious probabilistic aspects of perceptual processing) rather than credences assigned in perceptual consciousness itself. But Morrison raises problems for a variety of ways one might develop this alternative. Suppose, for example, perceptual consciousness only represents that there is a number of bars within a certain range at that location—a generic representation indeterminate with respect to its more specific determinations. Morrison objects that, if we "completely trust" our experience (as we often do), then the resulting perception-based belief should not favor any determination over another—and yet we may, on the basis of our perceptual experience, believe that one or another more specific determination is more likely than others (5 bars is more likely than 3).

A reply is that belief formation may completely trust perceptual consciousness by being completely consistent with it, and yet it may go beyond it. In the case at hand, belief-forming processes may lead us to assign credences where none had been assigned in perception—with, for example, higher credences towards the center of the number range (whether owing to principles embodied—and not necessarily explicitly represented—in those processes themselves or drawing upon perceptual information not represented in perceptual *consciousness*). To argue that such assignments are *inconsistent* with the content of perceptual consciousness, one must maintain that perceptual consciousness assigns some credence to these hypotheses (which is what's at issue) or perhaps at least entails some such assignment. But principles from which this would follow are highly unobvious—for example, that if perception represents a generic property, then it assigns a credence to each of its more specific determinations. In the present case, it does not seem to follow, from perceptual consciousness' representing—let's suppose—that there is some number of bars between 3 and 7 there (or that the number of bars is either 3, 4, 5, 6, or 7), that it assigns a credence to each candidate number of bars. And to assume that the generic representation, in the absence of such an assignment or other relevant information, entails a flat distribution is to assume a controversial Principle of Indifference [77].

This constitutes a reply to Morrison's argument *for* probabilistic perceptual consciousness. To convert it into a consideration *against*, one suggests that a representation that *abstains* from assigning probabilities at the relevant level of specificity is a candidate non-probabilistic perceptual representation phenomenologically indistinguishable from the probabilistic representation he posits. The challenge, given "confidentialism," is to specify an overlooked phenomenological difference. We mention three possibilities.

One response would suggest that there are phenomenological differences associated with perceptual anticipations of what one would see were one, for example, to look more closely—where these perceptual anticipations would reflect the distribution of credences [78]. It would need to be the case that this phenomenology was perceptual and distinct from feelings of surprise or fulfillment experienced when such actions were in fact undertaken. A second response might appeal to the epistemic emotion sometimes dubbed the "feeling of certainty" [79, 80]. It is perhaps harder to maintain that one can simultaneously experience distinct subjective feelings of certainty regarding each of a range of candidates. But this might be an option if perceptual consciousness represents only *one* candidate feature with some credence assigned just to it. (Cf. Rahnev [81] on options intermediate between representations of full distributions and single point estimates. This option is not available to Morrison, who requires that multiple candidates be assigned credences; nor does it fit extant models that distribute credences across a range of hypotheses.) It is unobvious, however, that this subjective feeling should itself count as *perceptual*. Further, since a typical perception represents multiple items and features, it would remain questionable whether one could experience distinct feelings of certainty for each (even if one represented but one candidate for each). Finally, third, it is always open to maintain that it's simply the perceptual assignment of credences themselves that lends a distinctive phenomenological character, and it's a mistake to try to cash out this difference in other terms. Here the attempt to break down resistance to the very idea of probabilistic perceptual consciousness becomes an attempt to shift the burden of argument onto its opponents.

(b) *Probabilistic processing without probabilistic representations*

The second strategy would remove the mismatch by denying that such models really require probabilistic representations.

A first way of denying probabilistic representations is connected to a point from section 4. Distinguish contents and their functional role—that is, the "attitudes" subjects bear to them. (Varying the attitude, one might *believe* that it's cold out or *desire* it be cold out; varying the content, one might believe *that it is cold out* or believe *that it is warm out*.) Perhaps subjective probabilities in perception should be considered part of the attitude not the content: one perceptually represents with .6 credence that that's red, as opposed to perceptually representing that there's a .6 subjective probability it's red (an apparent perceptual meta-representation concerning one's perceptual state). This might account for the thought that the probability distributions do not have accuracy conditions, because, even if it is better for the distribution to be "closer" to the truth by some measure, the credences themselves do not function to represent something about how things externally are probabilistically, independent of the subjects' attitudes. The representations would thus not be *about* probabilities. This indeed provides *a* sense in which one can deny that perceptual representations are probabilistic—but not one that dissolves the mismatch problem. For it would remain the case that subjects bear a variety of probabilistic attitudes towards the various candidate hypotheses, and someone who denies that perceptual consciousness is probabilistic would deny that, regarding some item, perceptual consciousness presents us with a range of candidate features to which we bear varying perceptual attitudes. (Use of the term 'probabilistic representation' outside of this paragraph—above and below—can be understood as including representations towards which there is a probabilistic attitude.)

A second way of denying probabilistic representations can be extracted from recent debates concerning the aims and claims of Bayesian models [82, 83]. When confronted with worries about computational tractability, departures from optimality, etc., proponents of Bayesian models sometime reply that they are operating at Marr's computational level, specifying the problem that some system needs to solve. Even if the computational level characterization involves probabilistic representations, this is not required of the algorithmic implementation, so long as it's approximately input-output equivalent in ecologically valid circumstances to the Bayesian decision-theoretic model [84-86]. If the algorithm in fact does not involve probabilistic representations, then the denial of probabilistic representations follows so long as elements posited only at the computational level are not "psychologically real." This allows for real probability at the algorithmic level, but in such cases it would concern only the probabilistic transitions among the non-probabilistic representations. One might then speak of probabilistic transitions as "implicitly representing" probabilities [87], but those who argue that laws governing transitions among representations are not themselves representations would suggest that this courts confusion [88].

It's worth noting that this move rejects the common view that the algorithmic level provides an intensional characterization of the computational level's function-in-extension—and not only because the move allows for merely approximate extensional equivalence. For Bayesian inference in effect treats priors and likelihoods as inputs, along with the signal. But with algorithmic probabilistic processing *sans* probabilistic representations, priors and likelihoods are instead embodied in transitions among representations. Changes in priors and likelihoods thus manifest themselves as changes in the probabilities governing these transitions—that is, in a change of the function computed at the algorithmic level.

How might one resist this second way of denying probabilistic representations? We mention three possibilities. First, Rescorla [89, applying 90] argues, that it suffices for causal explanation if we can intervene on a variable to directly affect outcomes, whatever the underlying mechanism; and such interventions—for example, on priors—are characteristic of research investigating probabilistic models in perception. One might develop this idea to suggest that even elements occurring only at Marr's computational level—such as, perhaps, representations of priors—may nonetheless be psychologically real.[7] A second, less theoretically-based response would be to argue that, in any event, many successful probabilistic models—including those *not* constructed with an eye towards optimality—are simply not

---

[7] Some, including Marr, align the computational/algorithmic distinction with the competence/performance distinction [91—and, in the context of Bayesian models, 92]. Even just aligning the computational level with competence might provide another way, or at least motivation, to support computational-level psychological reality, as some are loath to deny psychological reality to competence models. The alignment, however, is contentious (see Franks [93—also 94, 95] who, however, requires full, not approximate, extensional equivalence).

presented as probabilistic in processing only. For instance, Oberauer and Lin's [36] model for visual working memory encodes binding among features implicitly in probabilistic activation propensities but encodes the features themselves as distributions of activation across the candidates; these distributions are most naturally interpreted as probabilistic representations. Finally, third, perhaps the strongest evidence for probabilistic representations comes from the use of information concerning variance, as exhibited in performance [96]. However, for this to be relevant to the mismatch worry, there must also be evidence that these probabilistic representations are among the candidates for perceptual consciousness—for example, that they don't only occur post-perceptually. Some evidence for this comes from the decodability of variance from visual cortex [97]—but see [74].

(c) *Reconciling non-probabilistic consciousness and probabilistic representations*

Suppose that there is good reason to posit probabilistic representations in perception and short-term memory. There remains a straightforward way to reconcile probabilistic representations with non-probabilistic perceptual consciousness. One need only deny that *all* perceptual representations are probabilistic. Perceptual consciousness might be identified with the result of a perceptual decision that settles on a candidate given a represented distribution across the hypothesis space. This could involve just sampling from the represented distribution, rather than a full implementation of Bayesian decision theory [39, 98, 99]. There are various options as to where this decision could occur, in accordance with various theories of consciousness—for example, prior to, at, or after encoding into working memory. (If prior to encoding into working memory, then encoding into working memory might then reintroduce imprecision in the form of a represented distribution.) Nor need the perceptual decision be seen as falling within a single linear path of information flow. Gross and Flombaum [39] mention the possibility of perceptual consciousness involving sampling that is independent of the information flow leading to retrieval from working memory. This might be compared with Jacobs and Silvanto's [100] argument, based on dissociable functional roles, that conscious introspection operates on a parallel "copy" of working memory traces.[8]

Finally, we mention one other way to reconcile probabilistic representations and non-probabilistic consciousness. In motivating the thought that perceptual consciousness is not probabilistic, we noted that various proposals for what it's like to experience a probability distribution yield results arguably indistinguishable from what it would be like to experience a certain non-probabilistic representation. The force of such considerations relies on a principle that two phenomenologically identical representational states cannot differ regarding assignment of credences. But what if this principle were abandoned? Representations could then be probabilistic even if their associated phenomenology did not determine that this was the case (since the phenomenology would also be consistent with a non-probabilistic representation). Perceptual consciousness might then be said to be non-probabilistic so far as its phenomenology is concerned. Because the phenomenology would be consistent also with non-probabilistic representational content (say, that of a single specific color), we might then also say that, in a

[8] There are also various options as to the kind of sampling. Two examples from the working memory models already cited: retrieval for Oberauer & Lin [36] involves one-off sampling where the probability of an item being selected equals its represented probability (in effect probability matching), while Schneegans & Bays [60] utilize drift-diffusion-style sequential sampling to a threshold.

broad sense, the phenomenology in some manner "samples" the representation: what it presents is consistent with one candidate from among the various assigned credences. But it would not be a sample in virtue of some further processing: we are speaking of the phenomenology associated with the probabilistic representation, not the phenomenology associated with a further, distinct representation derived from the probabilistic representation. This, along with the abandonment of an attractive principle relating phenomenology, content, and attitude, might suggest that this strategy fails to deliver a *non-mysterious* solution to the mismatch problem. But it's a question whether it introduces any new mystery beyond that of relating representations and phenomenology more generally—a core aspect of the "hard problem" of consciousness.

We have canvassed three responses to the apparent mismatch regarding probability between perceptual consciousness and some leading models of perception and short-term memory. The matter is of interest in its own right, but also the responses differ in their upshot for debates concerning perceptual consciousness and cognitive access. If perceptual consciousness is itself probabilistic, the trivial argument for "correctness" overflow is blocked, and the suggested fundamental difference with representations in working memory vanishes. If probabilistic models of processing in perception and short-term memory don't require probabilistic representations, then the worry about assessing the correctness of probabilistic representations does not arise. This may be so as well if some, but not all, of the representations need be probabilistic, depending on what representations get compared for accuracy. This in turn depends on which representations get identified with perceptual consciousness, something our discussion does not settle.

6. CONCLUSION

The relation of perceptual consciousness and cognitive access has been a central concern among consciousness researchers. Given how cognitive access has been associated with working memory, changing conceptions of working memory require that we re-examine our conceptions of cognitive access and the consequences for debates about consciousness. We have shown how these new conceptions reject the assumption—shared by many theories of consciousness—that cognitive access is capacity-limited, how they undermine overflow arguments, and how they dissociate debates about the richness of perceptual consciousness and about the necessity of cognitive access for perceptual consciousness. We also examined possible reformulations of overflow that drop the assumption of capacity-limited cognitive access, and we surveyed responses to the apparent mismatch between perceptual consciousness and the probabilistic representations posited by these models (and indeed more generally in currently dominant work on perceptual processing). Taking capacity-unlimited conceptions of working memory into account reconfigures important aspects of these debates, but many crucial questions remain open.

I have no competing interests.

Thanks to Ned Block, Cameron Buckner, Chris Fetsch, Jon Flombaum, John Morrison, Klaus Oberauer, Ian Phillips, Michael Rescorla, this journal's anonymous referees, and an audience at New York University.

REFERENCES

1. Jensen MS, Yao R, Street WN, Simons DJ. Change blindness and inattentional blindness. Wiley Interdisciplinary Reviews: Cognitive Science. 2011 Mar 1;2(5):529–46. (doi: 10.1002/wcs.130)
2. Dux PE, Marois R. The attentional blink: A review of data and theory. Attention, Perception & Psychophysics. 2009 Nov 1;71(8):1683–700. (doi: 10.3758/APP.71.8.1683)
3. Kouider S, de Gardelle V, Sackur J, Dupoux E. How rich is consciousness? The partial awareness hypothesis. Trends in Cognitive Sciences. 2010 Jul;14(7):301–7. (doi: /10.1016/j.tics.2010.04.006)
4. Cohen MA, Dennett DC, Kanwisher N. What is the Bandwidth of Perceptual Experience? Trends in Cognitive Sciences. 2016 May;20(5):324–35. (doi: 10.1016/j.tics.2016.03.006)
5. Noë A, Pessoa L, Thompson E. Beyond the Grand Illusion: What Change Blindness Really Teaches Us About Vision. Visual Cognition. 2000 Jan;7(1–3):93–106. (doi: 10.1080/135062800394702)
6. Sperling G. The information available in brief visual presentations. Psychological Monographs: General and Applied. 1960;74(11):1–29. (doi: 10.1037/h0093759)
7. Vandenbroucke ARE, Sligte IG, Lamme VAF. Manipulations of attention dissociate fragile visual short-term memory from visual working memory. Neuropsychologia. 2011 May;49(6):1559–68. (doi: 10.1016/j.neuropsychologia.2010.12.044)
8. Bronfman ZZ, Brezis N, Jacobson H, Usher M. We See More Than We Can Report. Psychological Science. 2014 May 9;25(7):1394–403. (doi: 10.1177/0956797614532656)
9. Block N. Consciousness, accessibility, and the mesh between psychology and neuroscience. Behavioral and Brain Sciences. 2007 Dec;30(5–6). (doi: 10.1017/S0140525X07002786)
10. Block N. Perceptual consciousness overflows cognitive access. Trends in Cognitive Sciences. 2011 Dec;15(12):567–75. (doi: 10.1016/j.tics.2011.11.001)
11. Lamme VAF. How neuroscience will change our view on consciousness. Cognitive Neuroscience. 2010 Aug 18;1(3):204–20. (doi: 10.1080/17588921003731586)
12. Cohen MA, Dennett DC. Consciousness cannot be separated from function. Trends in Cognitive Sciences. 2011 Aug;15(8):358–64. (doi: 10.1016/j.tics.2011.06.008)
13. Michel M. Methodological artefacts in consciousness science. Journal of Consciousness Studies. 2017 Jan 24(11-12): 94-117.
14. Neisser U. Cognitive Psychology. Prentice-Hall; 1967.
15. Baddeley A. Working memory and conscious awareness. In Theories of Memory. Lawrence Erlbaum Associates. 1993. p. 11-28.
16. Block N. On a confusion about a function of consciousness. Behavioral and Brain Sciences. 1995 Jun;18(2):227. (doi: 10.1017/S0140525X00038188)
17. Chalmers DJ. Availability: The cognitive basis of experience. Behavioral and Brain Sciences. 1997 Mar;20(1):148–9. (doi: 10.1017/S0140525X97240057)
18. Carruthers P. Block's Overflow Argument. Pacific Philosophical Quarterly. 2017 Dec 11;98:65–70. (doi: 10.1111/papq.12152)
19. Dehaene S, Changeux J-P. Experimental and Theoretical Approaches to Conscious Processing. Neuron. 2011 Apr;70(2):200–27. (doi: 10.1016/j.neuron.2011.03.018)
20. Soto D, Silvanto J. Reappraising the relationship between working memory and conscious awareness. Trends in Cognitive Sciences. 2014 Oct;18(10):520–5. (doi: 10.1016/j.tics.2014.06.005)

21. King J-R, Pescetelli N, Dehaene S. Brain Mechanisms Underlying the Brief Maintenance of Seen and Unseen Sensory Information. Neuron. 2016 Dec;92(5):1122–34. (doi: 10.1016/j.neuron.2016.10.051)
22. Dehaene S. Consciousness and the Brain. Viking; 2014.
23. Rosenthal D. Consciousness and the Mind. Oxford University Press; 2005.
24. Lau H, Rosenthal D. Empirical support for higher-order theories of conscious awareness. Trends in Cognitive Sciences. 2011 Aug;15(8):365–73. (doi: 10.1016/j.tics.2011.05.009)
25. Brown R. Consciousness doesn't overflow cognition. Frontiers in Psychology. 2014 Dec 4;5. (doi: 10.3389/fpsyg.2014.01399)
26. Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences. 2001 Feb;24(1):87–114. (doi: 10.1017/S0140525X01003922)
27. Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. Nature. 1997 Nov 20;390(6657):279–81. (doi: 10.1038/36846)
28. Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. Nature. 2008 Apr 2;453(7192):233–5. (doi: 10.1038/nature06860)
29. Vandenbroucke ARE, Sligte IG, Lamme VAF. Manipulations of attention dissociate fragile visual short-term memory from visual working memory. Neuropsychologia. 2011 May;49(6):1559–68. (doi: 10.1016/j.neuropsychologia.2010.12.044)
30. Ma WJ, Husain M, Bays PM. Changing concepts of working memory. Nature Neuroscience. 2014 Feb 25;17(3):347–56. (doi: 10.1038/nn.3655)
31. Bays PM, Husain M. Dynamic Shifts of Limited Working Memory Resources in Human Vision. Science. 2008 Aug 8;321(5890):851–4. (doi: 10.1126/science.1158023)
32. Bays PM, Gorgoraptis N, Wee N, Marshall L, Husain M. Temporal dynamics of encoding, storage, and reallocation of visual working memory. Journal of Vision. 2011 Sep 12;11(10):6–6. (doi: 10.1167/11.10.6)
33. Bays PM. Spikes not slots: noise in neural populations limits working memory. Trends in Cognitive Sciences. 2015 Aug;19(8):431–8. (doi: 10.1016/j.tics.2015.06.004)
34. Orhan AE, Ma WJ. Neural Population Coding of Multiple Stimuli. Journal of Neuroscience. 2015 Mar 4;35(9):3825–41. (doi: 10.1523/JNEUROSCI.4097-14.2015)
35. Swan G, Wyble B. The binding pool: A model of shared neural resources for distinct items in visual working memory. Attention, Perception, & Psychophysics. 2014 Mar 15;76(7):2136–57. (doi: 10.3758/s13414-014-0633-3)
36. Oberauer K, Lin H-Y. An interference model of visual working memory. Psychological Review. 2017;124(1):21–59. (doi: 10.1037/rev0000044)
37. Endress AD, Szabó S. Interference and memory capacity limitations. Psychological Review. 2017 Oct;124(5):551–71. (doi: 10.1037/rev0000071)
38. Bays PM. Reassessing the evidence for capacity limits in neural signals related to working memory. Cerebral Cortex. In press. (doi: 10.1093/cercor/bhx351)
39. Gross S, Flombaum J. Does Perceptual Consciousness Overflow Cognitive Access? The Challenge from Probabilistic, Hierarchical Processes. Mind & Language. 2017 Jun;32(3):358–91. (doi: 10.1111/mila.12144)
40. Bays PM, Catalao RFG, Husain M. The precision of visual working memory is set by allocation of a shared resource. Journal of Vision. 2009 Sep 1;9(10):7–7. (doi: 10.1167/9.10.7)

41. Bae GY, Flombaum JI. Two Items Remembered as Precisely as One. Psychological Science. 2013 Aug 12;24(10):2038–47. (doi: 10.1177/0956797613484938)
42. Bays PM. Evaluating and excluding swap errors in analogue tests of working memory. Scientific Reports. 2016 Jan 13;6(1). (doi: 10.1038/srep19203)
43. Souza AS, Oberauer K. In search of the focus of attention in working memory: 13 years of the retro-cue effect. Attention, Perception, & Psychophysics. 2016 Apr 20;78(7):1839–60. (doi: 10.3758/s13414-016-1108-5)
44. Pertzov Y, Bays PM, Joseph S, Husain M. Rapid forgetting prevented by retrospective attention cues. Journal of Experimental Psychology: Human Perception and Performance. 2013 Oct;39(5):1224–31. (doi: 10.1037/a0030947)
45. van den Berg R, Awh E, Ma WJ. Factorial comparison of working memory models. Psychological Review. 2014;121(1):124–49. (doi: 10.1037/a0035234)
46. Adam KCS, Vogel EK, Awh E. Clear evidence for item limits in visual working memory. Cognitive Psychology. 2017 Sep;97:79–97. (doi: 10.1016/j.cogpsych.2017.07.001)
47. Pratte MS. Iconic Memories Die a Sudden Death. Psychological Science. 2018 Apr 19;95679761774711. (doi: 10.1177/0956797617747118)
48. Graff D. Shifting Sands. Philosophical Topics. 2000;28(1):45–81. (doi: 10.5840/philtopics20002816)
49. Dehaene S, Changeux J-P, Naccache L, Sackur J, Sergent C. Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends in Cognitive Sciences. 2006 May;10(5):204–11. (doi: 10.1016/j.tics.2006.03.007)
50. Noë A. Is the world a grand illusion? Journal of Consciousness Studies. 2002 May;9(5-6): 1-12.
51. Cohen J. The grand grand illusion illusion. Journal of Consciousness Studies. 2002 May;9(5-6): 141-57.
52. Cova F, Gaillard M, and Kammerer F. Subjective data and the overflow argument. University of Geneva. Manuscript.
53. Phillips IB. Perception and Iconic Memory: What Sperling Doesn't Show. Mind & Language. 2011 Sep;26(4):381–411. (doi: 10.1111/j.1468-0017.2011.01422.x)
54. Stazicker J. Attention, Visual Consciousness and Indeterminacy. Mind & Language. 2011 Mar 15;26(2):156–84. (doi: 10.1111/j.1468-0017.2011.01414.x)
55. Shevlin H. Consciousness, Perception, and Short-Term Memory. Doctoral dissertation. CUNY; 2016.
56. Beck J. Some worries about the "no-overflow" interpretation of post-stimulus cueing effects. Mind & Language Symposium. The Brains Blog. 2017 Jun 12. http://philosophyofbrains.com/2017/06/12/symposium-on-gross-and-flombaum-does-perceptual-consciousness-overflow-cognitive-access-the-challenge-from-probabilistic-hierarchical-processes.aspx
57. Orlandi N, Franklin A. Getting clear on the challenge. Mind & Language Symposium. The Brains Blog. 2017 Jun 12. http://philosophyofbrains.com/2017/06/12/symposium-on-gross-and-flombaum-does-perceptual-consciousness-overflow-cognitive-access-the-challenge-from-probabilistic-hierarchical-processes.aspx
58. Phillips IB. Commentary on Steven Gross and Jonathan Flombaum's 'Does perceptual consciousness overflow cognitive access? The challenge from probabilistic, hierarchical processes.' Mind & Language Symposium. The Brains Blog. 2017 Jun 12. http://philosophyofbrains.com/2017/06/12/symposium-on-gross-and-flombaum-does-

perceptual-consciousness-overflow-cognitive-access-the-challenge-from-probabilistic-hierarchical-processes.aspx
59. Gross S, Flombaum J. Perceptual consciousness, short-term memory, and overflow: Replies to Beck, Orlandi and Franklin, and Phillips. Mind & Language Symposium. The Brains Blog. 2017 Jun 12. http://philosophyofbrains.com/2017/06/12/symposium-on-gross-and-flombaum-does-perceptual-consciousness-overflow-cognitive-access-the-challenge-from-probabilistic-hierarchical-processes.aspx
60. Schneegans S, Bays PM. No fixed item limit in visuospatial working memory. Cortex. 2016 Oct;83:181–93. (doi: 10.1016/j.cortex.2016.07.021)
61. Dainton, B. Temporal consciousness. The Stanford Encyclopedia of Philosophy. 2017 Fall. https://plato.stanford.edu/archives/fall2017/entries/consciousness-temporal/
62. Phillips IB. Consciousness, time, and memory. In The Routledge Handbook of Consciousness. In press.
63. Carruthers P. The Centered Mind. Oxford University Press; 2015. (doi: 10.1093/acprof:oso/9780198738824.001.0001)
64. Serences JT. Neural mechanisms of information storage in visual short-term memory. Vision Research. 2016 Nov;128:53–67. (doi: 10.1016/j.visres.2016.09.010)
65. Xu Y. Reevaluating the Sensory Account of Visual Working Memory Storage. Trends in Cognitive Sciences. 2017 Oct;21(10):794–815. (doi: 10.1016/j.tics.2017.06.013)
66. Stokes MG. "Activity-silent" working memory in prefrontal cortex: a dynamic coding framework. Trends in Cognitive Sciences. 2015 Jul;19(7):394–405. (doi: 10.1016/j.tics.2015.05.004)
67. Morrison J. Perceptual Confidence. Analytic Philosophy. 2016 Feb 25;57(1):15–48. (doi: 10.1111/phib.12077)
68. Joyce JM. A Nonpragmatic Vindication of Probabilism. Philosophy of Science. 1998 Dec;65(4):575–603. (doi: 10.1086/392661)
69. Pettigrew R. Accuracy and the Laws of Credence. Oxford University Press; 2016. (doi: 10.1093/acprof:oso/9780198732716.001.0001)
70. Millikan R. Language, Thought, and Other Biological Categories. MIT Press; 1984.
71. Maloney L. Statistical decision theory and biological vision. In Perception and the Physical World. 2001 Wiley. p. 145–189.
72. Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences. 2013 May 10;36(3):181–204. (doi: 10.1017/S0140525X12000477)
73. Denison RN. Precision, Not Confidence, Describes the Uncertainty of Perceptual Experience: Comment on John Morrison's "Perceptual Confidence." Analytic Philosophy. 2017 Mar;58(1):58–70. (doi: 10.1111/phib.12092)
74. Block N. If perception is probabilistic, why doesn't it seem probabilistic? Philosophical Transactions of the Royal Society B: Biological Sciences.
75. Kersten D, Mamassian P, Yuille A. Object Perception as Bayesian Inference. Annual Review of Psychology. 2004 Feb;55(1):271–304. (doi: 10.1146/annurev.psych.55.090902.142005)
76. Morrison J. Perceptual Confidence and Categorization. Analytic Philosophy. 2017 Mar;58(1):71–85. (doi: 10.1111/phib.12094)
77. Háyek A. Interpretations of probability. The Stanford Encyclopedia of Philosophy. 2012 Winter. https://plato.stanford.edu/archives/win2012/entries/probability-interpret/
78. Madary M. Visual Phenomenology. MIT Press; 2017.

79. Smith JD, Shields WE, Washburn DA. The comparative psychology of uncertainty monitoring and metacognition. Behavioral and Brain Sciences. 2003 Jun;26(3). (doi: 10.1017/S0140525X03000086)
80. Dokic J. Seeds of self-knowledge: noetic feelings and metacognition. In: Foundations of Metacognition. Oxford University Press; 2012. p. 302–21. (doi: 10.1093/acprof:oso/9780199646739.003.0020)
81. Rahnev D. The case against full probability distributions in perceptual decision making. Manuscript.
82. Jones M, Love BC. Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. Behavioral and Brain Sciences. 2011 Aug;34(4):169–88. (doi: 10.1017/S0140525X10003134)
83. Bowers JS, Davis CJ. Bayesian just-so stories in psychology and neuroscience. Psychological Bulletin. 2012;138(3):389–414. (doi: 10.1037/a0026450)
84. Griffiths TL, Chater N, Kemp C, Perfors A, Tenenbaum JB. Probabilistic models of cognition: exploring representations and inductive biases. Trends in Cognitive Sciences. 2010 Aug;14(8):357–64. (doi: 10.1016/j.tics.2010.05.004)
85. Griffiths TL, Chater N, Norris D, Pouget A. How the Bayesians got their beliefs (and what those beliefs actually are): Comment on Bowers and Davis (2012). Psychological Bulletin. 2012;138(3):415–22. (10.1037/a0026884)
86. Sanborn AN, Chater N. Bayesian Brains without Probabilities. Trends in Cognitive Sciences. 2016 Dec;20(12):883–93. (doi: 10.1016/j.tics.2016.10.003)
87. Icard T. Subjective Probability as Sampling Propensity. Review of Philosophy and Psychology. 2015 Aug 22;7(4):863–903. (doi: 10.1007/s13164-015-0283-y)
88. Burge T. Origins of Objectivity. Oxford University Press; 2010. (doi: 10.1093/acprof:oso/9780199581405.001.0001)
89. Rescorla M. An interventionist approach to psychological explanation. Synthese. In press.
90. Woodward J. Making Things Happen: A Theory of Causal Explanation. Oxford University Press; 2003. (doi: 10.1093/0195155270.001.0001)
91. Marr D. Vision. MIT Press; 1982.
92. Feldman J. Bayesian Models of Perceptual Organization. Wagemans J, editor. Oxford Handbooks Online. Oxford University Press; 2014. (doi: 10.1093/oxfordhb/9780199686858.013.007)
93. Franks B. On Explanation in the Cognitive Sciences: Competence, Idealization, and the Failure of the Classical Cascade. The British Journal for the Philosophy of Science. 1995;46(4):475–502. (doi: 10.1093/bjps/46.4.475)
94. Patterson S. Competence and the Classical Cascade: A Reply to Franks. The British Journal for the Philosophy of Science. 1998 Dec 1;49(4):625–36. (doi: 10.1093/bjps/49.4.625)
95. Franks B. Discussion. Idealizations, competence and explanation: a response to Patterson. The British Journal for the Philosophy of Science. 1999 Dec 1;50(4):735–46. (doi: 10.1093/bjps/50.4.735)
96. Ma WJ, Jazayeri M. Neural Coding of Uncertainty and Probability. Annual Review of Neuroscience. 2014 Jul 8;37(1):205–20. (doi: 10.1146/annurev-neuro-071013-014017)
97. van Bergen RS, Ji Ma W, Pratte MS, Jehee JFM. Sensory uncertainty decoded from visual cortex predicts behavior. Nature Neuroscience. 2015 Oct 26;18(12):1728–30. (doi: 10.1038/nn.4150)
98. Vul E, Hanus D, Kanwisher N. Attention as inference: Selection is probabilistic; responses

are all-or-none samples. Journal of Experimental Psychology: General. 2009;138(4):546–60. (doi: 10.1037/a0017352)
99. Moreno-Bote R, Knill DC, Pouget A. Bayesian sampling in visual perception. Proceedings of the National Academy of Sciences. 2011 Jul 8;108(30):12491–6. (doi: 10.1073/pnas.1101430108)
100. Jacobs C, Silvanto J. How is working memory content consciously experienced? The "conscious copy" model of WM introspection. Neuroscience & Biobehavioral Reviews. 2015 Aug;55:510–9. (doi: 10.1016/j.neubiorev.2015.06.003)