

Probabilistic Representations in Perception: Are There Any, and What Would They Be?

Steven Gross
Department of Philosophy
Johns Hopkins University

address for correspondence:
Department of Philosophy
Johns Hopkins University
Baltimore, MD 212198
USA

email:
sgross11@jhu.edu

Abstract

Nick Shea's *Representation in Cognitive Science* commits him to representations in perceptual processing that are about probabilities. This commentary concerns how to adjudicate between this view and an alternative that locates the probabilities rather in the representational states' associated "attitudes". As background and motivation, evidence for probabilistic representations in perceptual processing is adduced, and it is shown how, on either conception, one can address a specific challenge Ned Block has raised to this evidence.

keywords: perception, probability, representation, attitudes, content

1. Introduction

Much research in perception science traffics in probabilistic representations. Should we adopt a realist attitude towards any of this work? Is there evidence for at least some probabilistic representation in perceptual processing? Nick Shea's (2018) *Representation in Cognitive Science* provides resources towards a positive answer. Below, I develop this by deploying Shea's conception of representation in response to a challenge drawn from Ned Block's (2018) recent defense of instrumentalism. However, I lay out an alternative response to this challenge as well. The second response adverts to a conception of probabilistic representation distinct from that to which Shea is committed. In addition to providing a challenge to Shea, it thus raises a fundamental question concerning what probabilistic representations are (or would be), at least of the sort possibly found in perceptual processing. I close by outlining considerations that might bear on this question, including some drawn from elsewhere in Shea's book.

2. Bayesian Modeling in Perception Science

First, some background. A familiar approach sees the task of perceptual systems as transitioning from a transduced signal to a representation of the signal's distal cause. But the relation between stimuli and signals is many-many, and each step from stimulus to signal to percept is noisy. A natural idea is to conceive of perceptual systems, faced with these facts, as carrying out, or approximating, Bayesian reasoning. This requires that perceptual processing has, or operates as if it has, the following elements: an assignment of prior probabilities to candidate hypotheses about the stimulus; a likelihood function assigning conditional probabilities to signals given hypothesized stimuli (in effect, a model of its own sensory response to stimuli); the ability to compute—in accordance with Bayes' theorem—the posterior probability of the hypotheses given the signal; and (assuming there is then a transition to a non-probabilistic percept) the ability to “decide” among the candidate hypotheses on the basis of their posterior probabilities. The Bayesian framework has proven immensely fruitful and arguably dominated perception science in recent decades (Knill & Richards 1996—see Bowers & Davis 2012 for dissent).

Consider a parade case. Drivers have a surprising tendency to speed up in the fog. Why? Because in low contrast conditions, motion is perceived as slower than it actually is, leading drivers to compensate by speeding up (Snowden et al. 1998). By why is speed perceived as slower than it actually is in low contrast conditions? One well-known answer is that the visual system embodies a prior that favors slower speeds and a likelihood function that flattens in low-contrast conditions (Weiss et al. 2002). That is, in low-contrast as compared to high-contrast conditions, the range of speeds that have a good chance of giving rise to the signal is wider, so that the likelihood function provides less differentiation among them, leaving more room for the prior to do so—which leads to the perception of a slower speed in low-contrast (e.g., foggy) conditions. This answer would explain, not only why drivers speed up in the fog, but also a variety of other previously disparate phenomena in the motion perception literature.

Behavioral results of this sort suggest that the output of perceptual processing, at least on these tasks, is more-or-less in accord with that of an ideal Bayesian decision-maker. But it is a further claim that perceptual systems achieve this accord via computations that parallel the ideal Bayesian decision-maker's. Indeed, the computational costs of updating probabilistic representations across a wide range of hypotheses provide reason to doubt that they do. Thus, a common suggestion is that perceptual systems may be Bayesian at Marr's computational (or, functional) level, but not at the algorithmic level. That is, Bayesian decision theory, by characterizing how an ideal reasoner should deal with noise and uncertainty, may help characterize the job of perceptual systems, supposing their function is to arrive at the likely cause of the signal under such conditions; but non-ideal perceptual systems may at best approximate this ideal by implementing algorithms approximately input-output equivalent (in relevant conditions) to what an ideal Bayesian decision-theoretic model would yield, or as close to such equivalence as resource constraints and opportunity have allowed. To the extent that they are roughly input-output equivalent to the ideal, approximating algorithms would cover the sort of behavioral data adverted to above (speeding up in the fog). To the extent that they depart from the ideal, they might explain as well the range of sub-optimal performance that has also been found (Rahnev & Denison 2018; cf. Won et al. ms.).

To be sure, one can reject full Bayesian up-dating without rejecting computations over probabilistic representations: an approximating algorithm might, for example, up-date the probabilities of just a few hypotheses from a partially represented space. But in fact many of the approximating algorithms explored assume only probabilistic transitions between states, not probabilistic representations (Sanborn and Chater 2016). If we grant that only what is found at

Marr's algorithmic level is psychologically real (but cf. Gross 2018), and if we suppose that it is a real possibility that algorithms over only non-probabilistic representations cover the data (however difficult it may be to discover what those algorithms are), then nothing so far said provides reason to posit probabilistic representations in perceptual processing.

3. Possible Evidence of Probabilistic Representations

Some recent work, however, suggests that some aspects of visual processing may indeed involve probabilistic representation. For example, van Bergen et al. (2015) had human subjects indicate the orientation of Gabor patches while fMRI targeted V1-V3, areas of sensory cortex widely agreed to be implicated in early visual processing. Their experiments thus combined behavioral and neurophysiological data—and they did so in an attempt to identify a causally relevant probabilistic state intermediate between signal and perceptually-driven behavior. Using a probabilistic population model for how probabilities may be encoded in patterns of neural firing (Ma et al. 2006), they decoded from the fMRI data a likelihood function and asked whether the variance decoded on a trial-by-trial basis correlated with aspects of performance.¹ Their main results included that more precise encodings (but not the width of estimated neural response) correlated with less performance variability, and that greater decoded variance correlated with greater oblique bias.² The first results obtained even considering only trials with the same stimulus orientation, eliminating the possibility that the results reflected some stimulus feature. That the results were found for variance decoded on a trial-by-trial basis is important because probabilistic properties can be found in data aggregated from non-probabilistic data points.

In subsequent work, Walker et al. (ms.) obtained similar results with macaques, using electrode-array recordings and a different orientation task. The macaques had to classify drifting grating stimuli into one of two classes defined by normal distributions with the same mean orientation but different standard deviations. A deep neural network trained on 80% of the electrode-array data was used to decode the likelihood function for the remaining 20%, so that Walker et al. relied on fewer theoretical assumptions than van Bergen et al. In addition, they performed a model comparison and found that a model making use of the full shape of the likelihood function outperformed one that used a point-estimate only.

These results—henceforth, the orientation results—are of course very far from dispositive. But my aim is neither to establish the case for probabilistic representations nor to assess the full range (or even a wide range) of currently available considerations. It is rather to address a particular philosophical challenge, which in turn will lead to a fundamental question concerning what probabilistic representations in perceptual processing would be. For my purposes, it suffices that the orientation results seem to provide some evidence of a causally-relevant state in sensory cortex—intermediate between transduction and behavior—that encodes probabilities. Suppose the results hold up. Would we then have evidence for probabilistic representations?

¹ Simplifying, the probabilistic population code relates likelihood and neural firing as follows: $p(r|s) \propto \exp[\sum_i r_i h_i(s)]$, where s is the stimulus and, for each neuron i , r_i is the neural response and $h_i(s)$ is the log of its tuning curve. Note that the relation is between the algorithmic level and the neurophysiological level; it is not something computed at the algorithmic level. Further theoretical assumptions are required to decode the likelihood from fMRI data.

² The oblique bias is a tendency to judge orientations just off vertical and horizontal as more oblique than they are. There exist various models of how this arises, but all would predict greater bias with greater variance in the likelihood function.

4. The Mere Sensitivity Challenge

Block (2018) offers two responses to van Bergen et al. (2015).³ I am unmoved by the first, as I will briefly explain; but it is the second I wish to draw out in order to raise my main question for Shea.

Block's first point is that one can redescribe the decoded variance in terms of the degree of competition among non-probabilistic representations. However, Block faces a dilemma. On the one hand, the redescription may retain the form of the probabilistic population coding model, including its account of computational transitions among states—as may be suggested by Block's claim that his alternative “gives essentially the same explanation” (2018, p. 6). But then the redescription just amounts to replacing talk of probabilities with talk of degrees computed over in the same way.

On the other hand, if the competition model is a genuine alternative, a comparison requires details Block does not provide. To underscore why this matters, consider a model Block alludes to earlier in his paper (2018, p. 2). Arguing for a mismatch between alleged probabilistic representations and the content of conscious perceptual experience, Block notes that neurons preferentially tuned to a 30° orientation fire to a substantial degree when presented with a vertical line, though the line does not look to have some probability of being tilted 30°. This assumes the encoded probability of a feature is proportional to the firing rate of neurons that prefer that feature—an assumption rejected by probabilistic population codes, which factor-in activation levels of neurons with other tuning preferences as well, preserving information highly relevant in a noisy system (cf. fn. 1 above, also Shea 2014). Probabilistic population codes in fact yield tighter curves in such cases (Ma et al. 2008—cf. Anderson et al. 2000 for empirical support), so that substantial firing of neurons that prefer 30° need not entail that a significant probability is assigned to that orientation.⁴ Thus, if Block's alternative competition model were cashed in terms of levels of activation among neurons whose preferred tuning determined the orientations they represent, it is not at all obvious the model would fit the orientation data as well as the probabilistic population coding model. Indeed, that the estimated width of neural response did not correlate with performance variability in van Bergen et al.'s experiment suggests it would not.⁵

Block's second, independent point—henceforth, the mere sensitivity challenge—is that sensitivity to a feature does not entail representation of that feature. This is a general point about representation, applied here to putative probabilistic representation. To drive it home, Block recounts an experiment by Dener et al. (2016) in which pea plant root growth displays a sensitivity to comparative nutrient variability in different regions (NB: not just to local changes in nutrient amounts). It would be quite controversial to conclude that the plants represent this variability (though cf. Gagliano 2017 and Shea 2018, pp. 213-4, on representation in plants). Similarly—the challenge runs—subjects' sensitivity to variability in their sensory system's response to stimuli does not of itself entail representation of that variability.

³ Block (2018) does not discuss Walker et al. (ms.), which was not then available. But the mere sensitivity challenge can be raised to it as well.

⁴ Block notes that population codes—presumably probabilistic population codes—might provide a reply to his mismatch argument. But he does not explain that this crucial difference in variance is why. His response is that we lack evidence for cortical decoding of population codes, which requires a mechanism of combination. But this is true for any mechanism implemented across a population of neurons, whether the code is probabilistic or not—including, presumably, Block's competition alternative.

⁵ Depending on the details of Block's model, it is also possible that the replies offered below to the mere sensitivity challenge would entail that Block's model contains probabilistic representations after all.

Adverting to the plants might seem to miss the crucial feature of the orientation results, since the plant experiment involves only behavioral data (the plants' root growth): no effort is made to exhibit a causally-relevant intermediate state that is sensitive to—for example, carries information about—nutrient variability. We have already noted that behavioral data of this sort constrains, but need not determine, the process by which the behavior was produced. The orientation experiments' behavioral and neurophysiological results are intended to take us deeper inside the mechanism.

But the mere sensitivity challenge need not be just that behavior's exhibiting sensitivity to a feature does not entail representation of that feature. We can render the challenge more relevant to the orientation results by extending it to causally-relevant intermediate states. Suppose the plant has such a state sensitive to—for example, carrying information about—nutrient variability. The challenge remains: it would be quite controversial to ascribe representational content to that state. Representationality is a more demanding notion than such sensitivity.

This is the point on which we will focus. Suppose the orientation results provide some evidence of a causally-relevant state in sensory cortex—intermediate between transduction and behavior—that is sensitive to probabilities. Is there reason to think there is probabilistic representation and not mere sensitivity to probabilities?

5. First Response: Sufficient Condition on Representationality Met

An obvious question is: what does it take for a mere sensitivity to rise to the level of representation? On a plausible conception of representation, would the probabilistic states suggested by the orientation results make the grade? Here is where we may wheel-in Shea's work. Shea (2018, p. 85) argues for a conception of "subpersonal" representation, according to which a component R of a system with some task function(s) represents a condition C if the component carries exploitable, unmediated explanatory, correlational information about C. 'Subpersonal' is here used as a short-hand for, roughly, non-doxastic, non-conscious states (Shea 2018, pp. 8, 26); it is left open whether the content of other states are otherwise constituted. And, even for subpersonal states, this is just a sufficient condition. Shea offers a distinct sufficient condition for representation based on structural correspondence. He thus characterizes his approach to content as pluralist.

But let us unpack the provided sufficient condition, the one relevant to our discussion. In the basic case, a state carries correlational information about C just in case the probability of C is different from the probability of C given that one is in that state (Shea 2018, p. 76). This information is exploitable just in case the correlation generalizes, for a univocal reason, to collections of conditions and states with elements thus correlated—for example, if there is not only a state that carries information about distal redness, but a range of states each correlated with a color (Shea 2018, p. 77). The explanatory condition requires that this type of state's carrying this information, as part of an algorithmic process, helps explain how the task function(s) has become stabilized through evolution, learning, or contribution to the organism's persistence and/or help explain how such states are produced robustly in response to a range of different inputs and relevant external conditions (a function that is stabilized and robust in these senses is a task function—Shea 2018, pp. 83-4). Lack of mediation rules out explanations based on correlations that depend on explanatorily more basic correlations, and thus reduces content indeterminacy (Shea 2018, p. 85). For example, it rules out explanations of frog fly-catching

behavior based on a correlation between a neural state and proximal light patterns, as opposed to one based on a correlation between the state and distal flies (Shea 2018, pp. 151-2).

The presumably subpersonal probabilistic states suggested by the orientation results fare well on this conception. As Shea (2018, p. 107) notes, accommodating probabilistic representation requires extending his conception of exploitable correlational information to joint probability distributions relating a state and a condition. But, that done, it is plausible that the states posited in light of the orientation results carry exploitable correlational information about the probability of the sensory transducers going into the state they are in given various candidate stimuli; and that they're doing so, as part of perceptual processing, helps explain in an unmediated way how the functions this processing serves became stabilized and robust. The probabilistic states thus rise to the level of representation on Shea's view.⁶

Incidentally, other conceptions of representation may likewise support this conclusion—for example, Burge's (2010). This is of particular interest because the mere sensitivity challenge echoes Burge's distinction between mere (functional) sensory registration of information and genuine representation, and Block endorses central elements of Burge's conception of representation (Block 2015, forthcoming). According to Burge, the mark of genuine representation in perception—a sufficient, perhaps necessary condition (Burge 2010, p. 413)—is the role of constancy mechanisms in generating the state. Might there be constancy mechanisms for probability assignments? That is unclear. But Burge's views do not actually require that the relevant constancy mechanism be a constancy mechanism *for the feature at issue*. In some cases, information rises to the level of representation in virtue of being “harnessed” (Burge 2010, p. 546) to a state that represents *something else*. Burge argues that this is so for numerical and temporal features. For example, though we lack constancy mechanisms for temporal features, we perceptually represent them insofar as a state carrying functional temporal information also carries other information supported by a constancy mechanism for it. Thus, given color constancy, if a perception as of *red there* carries the information that it is there *now*, and if this information is used downstream, it amounts to a representation as of *red there now* (but cf. Gross 2017). Since orientation constancy is established by V1 (Finn et al. 2007), one might likewise argue that probabilistic information attached to orientation representations rises to the level of genuine representation of probability.

Of course, one might reject Shea's (or Burge's) conception of representation. But Shea's arguments and case studies render it a plausible candidate: rejecting it puts pressure on the proponent of the mere sensitivity challenge to present an alternative conception of representation on which the posited probabilistic states would *not* make the grade.

6. Second Response: Probabilistic Representation Without Representation of Probability

But there is also an alternative response to the mere sensitivity challenge. The first response deploys a sufficient condition for representationality to argue that the orientation results provide evidence of genuine representation of probabilities. The second suggests that perceptual processing involves probabilistic representation *without* representation of probability.

The idea is to generalize the familiar distinction from conception between attitudes and content, evident in how one can vary one while holding fixed the other (believing vs. desiring that it is cold; believing it is cold vs. believing it is warm). This distinction allows us to locate

⁶ Shea (2018, pp. 106-10) discusses a model from Bogacz (2015) for assigning probabilities that candidate actions will lead to reward. See Shea (2014) for discussion of probabilistic population codes.

two places where probabilities might arise in conception: in the content or as an aspect of the attitude. Thus we might distinguish a graded doxastic attitude towards a proposition (e.g., assigning .7 credence to its being cold) and believing a proposition about the probability of something's obtaining (e.g., believing there is a .7 probability of its being cold). The latter requires greater conceptual sophistication than the former: the subject must possess and deploy the concept of probability. Arguably, young children and some non-human animals may have graded doxastic states but not beliefs about probabilities as such.⁷

Does an attitude-content distinction exist as well in perceptual processing?⁸ Initial motivation might come from our wanting to distinguish perceptual representations and conceptual representations, especially those with the same, or appropriately related, contents. Similarly, within perception, we want to distinguish visual and haptic representations, especially to the extent they agree in content. A difference in attitude can capture the differences in these representations' functional role.⁹ Generalizing the attitude-content distinction opens the possibility that, in perception as in conception, probabilities (if there are any) may be an aspect of attitudes—a possibility explicitly allowed for in recent discussions, including Block's (cf. Morrison 2016, Gross & Flombaum 2017, Gross 2018, Rescorla forthcoming).¹⁰

If so, what are the consequences for whether there is evidence for probabilistic representations in perception? Of course, if probabilities in perception are conceived of as aspects of attitudes, then the orientation results provide no evidence of representations *of* probability—that is, no evidence of probabilities in representational content. But the question remains whether there are perceptual probabilistic representations in the broader sense that includes representational states with probabilistic attitudes. This is an interesting question in its own right and, relatedly, a perfectly reasonable use of the label 'probabilistic representation' (it is so used in the discussions just cited). Block (2018), for example, is addressing a felt *prima facie* tension between probabilistic perceptions and conscious perceptual experience. As his title asks: If perception is probabilistic, why does it not seem probabilistic? Block's answer rejects the antecedent: perception is not probabilistic. To support this conclusion, it would not suffice to argue only that probabilities are not represented in perception. For the *prima facie* tension remains if we posit probabilities as aspects of attitudes: experience does not seem to present a range of candidate features to which we bear various probabilistic attitudes.¹¹

⁷ I bracket contents that mix conceptual and non-conceptual constituents (where perhaps probabilistic content could be non-conceptual). Also, to keep things simple, I use 'aspect of the attitude' to cover both views that would incorporate probabilities into the attitude and views that would instead treat probabilities as a third relatum distinct from both the attitude and the content.

⁸ Substitute some other term if one wants to reserve 'attitude' for conceptual states. Chalmers (2004) suggests 'manner of representation' (albeit with conscious states in mind) as preferable to Crane's (2003) 'intentional modes' (see also Searle 1983—also Shea's 2018 'modes of representation'), which might suggest Fregean modes of presentation, an aspect of content.

⁹ One can challenge these motivations. Perhaps differences in content—including broadly structural differences such as whether predicational attribution is involved (Burge 2010, Block ms.)—suffice to capture the perceptual-conceptual difference. Perhaps the visual-haptic difference is also a difference in content (Rescorla forthcoming argues they involve different modes of presentation). Or perhaps a difference in non-content aspects of the vehicles could be invoked (though why would these not count as differences in manner of representation?). I lack space to pursue these suggestions here. We turn to considerations specific to probabilistic representation in section 7.

¹⁰ Block and Rescorla are explicit that they do *not* have in mind representations *of* probability.

¹¹ Morrison (2016) and Gross & Flombaum (2017) provide alternative responses to the tension. Morrison argues for probabilities in perceptual experience; Gross & Flombaum suggest there is a perceptual decision—a transition from probabilities to a point estimate—prior to perceptual experience. Cf. Gross (2018).

But if representations with probabilistic attitudes are a species of probabilistic representation, we have an alternative reply to the mere sensitivity challenge. For the distinction between mere sensitivity and genuine representation is *irrelevant* if we are talking about a state's attitude, not its content. Given that a state has representational content at all, the attitude associated with it is determined by its functional role. There is no distinction—parallel to the mere sensitivity/genuine representation distinction—that we must also bear in mind and so no further question concerning whether there really is such an attitude associated with the state.

7. What Are Probabilistic Representations in Perception?

We have two replies to the mere sensitivity challenge, corresponding to two views of where to locate the probabilities posited in light of the orientation results. If the attitude account is preferable, either the clauses of Shea's sufficient condition must not be satisfied in this case, or the condition itself must be rejected (possibly just restricted or otherwise tweaked). Of course, Shea's book presents an extended general argument for the conception of representation of which the proposed sufficient condition is a part. This argument features a variety of case studies—including a probabilistic case—intended to show that the proposed conception captures the explanatory role of representations in cognitive science. Here, however, I want to indicate some considerations specific to probabilistic representation that we might bring to bear. The considerations raise questions too large to fully explore here, but we can say enough to show that it is not obvious which conception they favor.

7.1 representational sophistication

The attitude view requires less representational sophistication, perhaps to its advantage. Recall the comparison with probabilities in conception. Conceptual representations of probability are representationally more sophisticated than graded attitudes. Unlike graded attitudes, conceptual representations of probability require possession of, and the ability to deploy, the concept of probability. Relatedly, in the conceptual case, graded attitudes would seem to be developmentally prior to representations of probability: we can believe to varying degrees before we can entertain thoughts about probability. This might suggest that the capacity to represent probabilities is more demanding than the capacity for graded attitudes and that, when deciding between attribution of one or the other, we should default to the latter.

It is unclear, however, that the considerations relevant to the conceptual case carry over to the perceptual case. At least for “low-level” features like orientation, the representational capacity of perception—unlike that of conception—is arguably innately determined or strongly constrained,¹² so that developmental achievement is not a concern: evolution might have taken care of the achievement for us.

7.2 objective probability

A second candidate advantage of the attitude view is that it need not posit objective probabilities (however conceived) that representations of probability are supposed to get right. But perhaps the content view need not posit objective probabilities either, only representations *as of* objective probabilities: representations as of F do not entail the existence of F. Agnosticism about objective probabilities allows the possibility of an error theory, whereby objective probabilities would have something like the status of colors according to eliminativists who

¹² Even mad-dog conceptual nativists can accept the disanalogy with conception by formulating it in terms of when representations first become available for various processes—cf. Gross and Rey (2012).

nonetheless maintain we have representations as of color. Of course, it would then be a question how states could have this empty content when there is nothing for them to carry information about.

In any event, the consideration carries no weight if one is not bothered by objective probabilities. Taken at face value, the Bayesian framework's motivation can seem to take them for granted. The whole idea of the likelihood function is that the perceptual system models its own stochastic response to external stimuli (on an instrumentalist construal: *in effect* models).¹³ Some of the Bayesian methodology can likewise seem to take objective probabilities for granted. For example, the natural scene statistics approach tries to establish plausible perceptual priors by uncovering actual frequencies in natural scenes (Geisler 2008). Moreover, Shea is committed to objective probabilities independent of considerations concerning probabilistic representations: he characterizes informational correlation in terms of the raising and lowering of probabilities objectively construed. It may not be obvious how best to understand objective probabilities. (Perceptual posterior probabilities—the probability of a stimulus feature given the signal—provide a nice example of Humphreys' (1985) Paradox for propensity interpretations.) But one may be committed to objective probabilities in advance of knowing how best to understand them.

Still, the existence of objective probabilities does not entail that all assignments of probability are best construed as attempts to get objective probabilities right. How this might bear on our question might depend on just what probabilistic representations are posited. Suppose, for example, what is decoded are not likelihoods or posteriors, but rather the up-dated probabilities of stimuli (i.e., after conditionalization). It is unclear what objective probabilities these might be aimed at—unless they are notional after all.

7.3 one state with multiple probabilities

A potential consideration in favor of the content view comes from a variant of Jackson's (1975) well-known objection to adverbialism. Regarding what to place in content and what in the attitude, adverbialism occupies one extreme: everything is in the attitude, nothing in content; instead of having a perception as of red, one perceives redly. Jackson objects that adverbialism has trouble handling scenarios containing multiple particulars with distinct features. Whether there was a green square and red circle or a red square and green circle, adverbialism would have us perceiving redly, squarely, greenly, and circle-ly (but see Tye 1984). A representational view has no problem with this insofar as some sort of combinatorics in representational content allows features bound to particulars.

What if, then, we need to assign—or bind, as it were—distinct probabilities to distinct aspects of a single state? It may seem plausible that we do: on the basis of perception, I may be very confident of the color of an object, less confident of its shape; or very confident of the color of one object, less of another. But, even if we grant that the basis of these judgments are single perceptual states, it is less clear that they reflect probabilities in perception as opposed to post-perceptual assignments of confidence (cf. Morrison 2016 and Munton 2016 for confidences in perceptual experience, but also Gross 2018 and Beck forthcoming).

What do we find in experimental and modeling work on probabilities in perceptual processing? I am not aware of a clear case of a model that assigns distinct probabilities to distinct

¹³ Perhaps the conditional distribution of r given s (where r is varied and s is constant) is better thought of as modeling this, rather than the likelihood function (where r is held constant and s varied). But it suffices for our purposes that perceptual processing has the capacity to represent the conditional probabilities each incorporates.

aspects of a single state. (Granted, this could just reflect the tendency of scientific work to start with the simplest, most tractable cases.) Of course, one state may assign distinct probabilities to the distinct candidates of the hypothesis space; but we would be looking for one state with multiple distributions. Also, distinct states may be the basis for a transition to a new state, as with cue combination. But this does not yet get us one state with distinct probabilities for distinct aspects of that state.¹⁴ Similarly, models that would take probability distributions for distinct features and bind them by creating a joint distribution would not count. This involves a transition from two simpler hypothesis spaces to one more complex space, but still with one probability assigned to each hypothesis. (Tractability worries can become especially acute here, but this might just suggest that probability representations might only occur early in processing for low level features.)

One might think it is easy enough nonetheless to find what we are looking for: Take the state that encodes the probabilities for one feature and the state that encodes the probabilities for another, and consider the complex state that has those two states as constituents. But this would be the analogue of the complex state that consists of believing P and believing Q, not the analogue of believing the conjunction P and Q. We need reason to posit something like the latter, with probabilities assigned within content.

7.4 attitudes and reusability

Shea (2018, p. 178) remarks that the attitude-content distinction is useful to mark a representation's reusability in distinct functional roles. Differences in content likewise correspond to differences in functional role. So, reusability would provide a necessary, not sufficient, condition (unless a kind of difference in functional role particular to attitudes were specified). If the condition were just that the same representational content be attachable to distinct attitudes, it would certainly be met from the proponent of probabilistic attitudes' perspective. Within perceptual processing there may be, say, a .6 representation as of that's being red or .8 representation; and, after a perceptual decision, one may have an outright visual experience as of that's being red. But the condition Shea proposes is rather that the vehicle be reusable.¹⁵ This requires a distinct vehicle corresponding to the non-probabilistic content. Would probabilistic representations on the attitude view meet this condition? If not, this might favor the content view. (Unless the content view is committed to a kind of content combinatorics that likewise requires reuseability. If it is, reusability *per se* does not provide a neutral way to adjudicate between the alternatives.)

Probabilistic representations may well satisfy reusability by Shea's lights—though the answer turns on delicate questions concerning the nature of vehicles and their individuation. He notes (2018, p. 193, fn. 8) that, in probabilistic representation, a vehicle's level of activation can represent the probability that a particular world state obtains. The suggestion seems to be that a vehicle representing a world state may persist across changes in level of activation, where the latter represents the probability assigned to that state. (On the probabilistic population coding scheme, where it is unclear there are distinct vehicles for distinct candidate world states, we

¹⁴ With cue combination, there are also the distinct weights assigned to each cue based on reliability. But they simply reflect the difference in variance (and, again, are not assigned to distinct aspects of one state).

¹⁵ This could be understood as requiring that vehicles of the same *type* be attachable to distinct attitudes or as requiring reusability of vehicle *tokens*. Some of Shea's wording seems to suggest the latter. This may be objectionably strong: is that *required* in the conceptual case? But the responses to which I now turn could be given regardless.

might rather say that the vehicle's level of activation represents the distribution of probabilities across the hypothesis space—where the representation of the space is what persists across changes in level of activation.) But now the proponent of the attitude view can recast this in her terms: there is a persistent vehicle that alters its relation to something (level of activation) in virtue of which its associated manner of representation varies.

Additionally, perhaps one may question Shea's condition in the first place. Reusability comes up when Shea (2018, p. 178) considers the descriptive-directive distinction (C obtains vs. produce C). In the subpersonal case, Shea opts to locate this distinction within content, as opposed to manner of representation, in part to avoid the assumption of reusability, which—he argues—is not satisfied in all cases where one wants to mark that a state is directive or descriptive. He accordingly provides criteria for descriptive vs. directive content in terms of whether the state carries (exploitable, unmediated explanatory) information about condition C because the state produces C. But Shea (2018, p. 187) acknowledges that there can be subpersonal systems with descriptive-directive reusability to which therefore an attitude-content distinction applies.¹⁶ If the attitudinal aspect of a state is distinguished from its content, then in such cases it seems the descriptive or directional aspect both counts as content in virtue of Shea's descriptive-directive criteria and does not so count in virtue of reusability. The tension is superficial: Shea's discussion marks a broader and narrower conception of content. But then why not contemplate a broader and narrower conception of attitudes as well? The narrower would require reusability, the broader would not. In favor of the broader attitudinal view would be that its associated narrower conception of content would not require special clauses to accommodate the descriptive-directive distinction within content.

Shea (2018, p. 178) characterizes his locating the descriptive-directive distinction within content as a terminological choice—albeit one reflecting a preference for which he gives reasons. The first reason, already mentioned, is not having to assume reusability, and a reply is to drop reusability as condition on there being attitudes. The second is that otherwise the theory of content would be incomplete. Here one might reply that it need not be incomplete so far as content goes, though—to be sure—there would be further work for a theory of representational states that covered both their contents and their manners of representation. It would be reasonable to ask, if reusability is dropped as a necessary condition, what might take its place. But recall that, when it comes to theorizing about content, Shea is a pluralist; something similar could be in the cards concerning manners of representation, which might have a cluster or prototype structure. If the reasons supporting each choice are balanced (or, if they are not, but the deciding reasons could be considered merely pragmatic), might we have a case of indeterminacy, at least so far as the subpersonal goes? That is possible, but it is also reasonable to suspend judgment.

Probabilistic representations in perceptual processing—if there any—thus present an intriguing but difficult case. It would be good to gain further clarity about them.

Acknowledgements

These remarks are sparked by Shea (2018), but also provide a partial reply to Block (2018). Thanks to Nick Shea and Ned Block for conversation, and to Nick for his excellent book.

¹⁶ Shea also notes the possibility of a descriptive-suppositional distinction, to accommodate hypothetical reasoning in off-line planning (for example when place cells are deployed to plan navigation). Visual perception vs. visual imagery might provide another case.

Because bits and pieces have shown up in various talks, my acknowledgements are disproportionate to the length and substance of this note. Thanks to Peter Achinstein, Austin Baker, Gary Ebbs, Chris Fetsch, Chris Gauker, Justin Halberda, Kevin Lande, Weiji Ma, Larry Maloney, Jorge Morales, John Morrison, Bence Nanay, Casey O’Callaghan, Nico Orlandi, Ian Phillips, Michael Rescorla, Georges Rey, J. Brendan Ritchie, and Susanna Siegel—as well as audiences at Antwerp, Cambridge, Columbia, Johns Hopkins, Oslo, SPP, SSPP, and York. My apologies for omissions.

References

- Anderson, J.S., Lampl, I., Gillespie, D.C. & Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science*, 290, 1968-72.
- Beck, J. (forthcoming). On perceptual confidence and ‘completely trusting your experience’. *Analytic Philosophy*.
- Block, N. (2015). Consciousness, big science and conceptual clarity. In J. Freeman & G. Marcus (Eds.), *The future of the brain: Essays by the world’s leading neuroscientists*. Princeton, NJ: Princeton University Press.
- Block, N. (2018). If perception is probabilistic, why doesn’t it seem probabilistic? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373. <https://doi.org/10.1098/rstb.2017.0341>
- Block, N. (forthcoming). *The border between seeing and thinking*.
- Bogacz, R. (2015). Optimal decision-making in the cortico-basal-ganglia circuit. In B. U. Forstmann & E.-J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience*. Berlin: Springer.
- Bowers, J. & Davis, C. (2012a). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, 138, 389-414.
- Burge, T. (2010). *Origins of objectivity*. Oxford: Oxford University Press.
- Chalmers, D. (2004). The representational character of experience. In B. Leiter (Ed.), *The future of philosophy*. Oxford: Oxford University Press.
- Crane, T. (2003). The intentional structure of consciousness. In Q. Smith and A. Jokic (Eds.), *Consciousness: New philosophical perspectives*. Oxford: Oxford University Press.
- Dener, E., Kacelnik, A. & Shemesh, H. (2016). Pea plants show risk sensitivity. *Current Biology*, 26, 1763-1767.
- Finn, I., Priebe, N. & Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron*, 54, 137-52.
- Gagliano, M. (2017). The minds of plants: Thinking the unthinkable. *Communicative & Integrative Biology*, 10, e1288333.
- Geisler, W. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, 59, 167-92.
- Gross, S. (2017). Perception and the origins of temporal representation. *Pacific Philosophical Quarterly*, 98(S1), 275-92.
- Gross, S. (2018). Perceptual consciousness and cognitive access from the perspective of capacity-unlimited working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373. <https://doi.org/10.1098/rstb.2017.0343>
- Gross, S., and Flombaum, J. (2017). Does perceptual consciousness overflow cognitive access? The challenge from probabilistic, hierarchical processes. *Mind & Language*, 32, 2017, 358-91.

- Gross, S., and Rey, G. (2012). Innateness. In E. Margolis, R. Samuels & S. Stich (Eds.) *The Oxford handbook of philosophy of cognitive science*. Oxford: Oxford University Press.
- Humphreys, P. (1985). Why propensities cannot be probabilities. *Philosophical Review*, 94, 557-70.
- Jackson, F. (1975). On the adverbial analysis of visual experience. *Metaphilosophy*, 6, 127-35.
- Knill, D. & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge: Cambridge University Press.
- Ma, W., Beck J., Latham P. & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9, 1432-8.
- Ma, W., Beck J. & Pouget, A. (2008). Spiking networks for Bayesian inference and choice. *Current Opinion in Neurobiology*, 18, 217-22.
- Morrison, J. (2016) Perceptual confidence. *Analytic Philosophy*, 57, 15–48.
- Munton, J. (2016). Visual confidences and direct perceptual justification. *Philosophical Topics*, 44, 301-26.
- Rahnev, D. & Denison, R. (2018). Suboptimality in perceptual decision making. *Behavioral and Brain Sciences*, 41, 1-107.
- Rescorla, M. (forthcoming). Perceptual co-reference. *Review of Philosophy and Psychology*.
- Sanborn, A. & Chater, N. (2016) Bayesian brains without probabilities. *Trends in Cognitive Sciences*, 20, 883–93.
- Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Shea, N. (2014). Neural signaling of probabilistic vectors. *Philosophy of Science*, 81, 902-913.
- Shea, N. (2018). *Representation in cognitive science*. Oxford: Oxford University Press.
- Tye, M. (1984). The adverbial approach to visual experience. *The Philosophical Review*, 93, 195-225.
- Snowden, R., Stimpson, N. & Ruddle, S. (1998). Speed perception fogs up as visibility drops. *Nature*, 392, 450.
- van Bergen, R., Ma, W., Pratte, M. & Jehee, J. (2015) Sensory uncertainty decoded from visual cortex predicts behavior. *Nature Neuroscience*, 18, 1728–30.
- Walker, E., Cotton, R., Ma, W. & Tolias, A. (manuscript). A neural basis of probabilistic computation in visual cortex.
- Weiss, Y., Simoncelli, E. & Adelson, E. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5, 598-604.
- Won, I., Firestone, C. & Gross, S. (manuscript). Impossible somatosensation.