Variance misperception under skewed empirical noise statistics explains overconfidence in the visual periphery

Authors: Charles J. Winter[1] & Megan A. K. Peters[1]

Affiliations:
[1] Department of Cognitive Sciences, University of California Irvine, Irvine, CA 92617

Correspondence should be addressed to:
Charles J. Winter
2201 Social & Behavioral Sciences Gateway Building
Irvine, CA 92697-5100
cjwinter@uci.edu

Megan A. K. Peters
2201 Social & Behavioral Sciences Gateway Building
Irvine, CA 92697-5100
megan.peters@uci.edu

Abstract

Perceptual confidence typically corresponds to accuracy. However, observers can be overconfident relative to accuracy, termed 'subjective inflation'. Inflation is stronger in the visual periphery relative to central vision, especially under conditions of peripheral inattention. Previous literature suggests inflation stems from errors in estimating noise, i.e. 'variance misperception'. However, despite previous Bayesian hypotheses about metacognitive noise estimation, no work has systematically explored how noise estimation may critically depend on empirical noise statistics which may differ across the visual field, with central noise distributed symmetrically but peripheral noise positively skewed. Here we examined central and peripheral vision predictions from five Bayesian-inspired noise-estimation algorithms under varying usage of noise priors, including effects of attention. Models that failed to optimally estimate noise exhibited peripheral inflation, but only models that explicitly used peripheral noise priors incorrectly showed increasing peripheral inflation under increasing peripheral inattention. Our findings explain peripheral inflation, especially under inattention.

Keywords: confidence; metacognition; perception; peripheral inflation; Bayesian ideal observer; hierarchical inference; natural scene statistics; empirical priors

Variance misperception under skewed empirical noise statistics explains overconfidence in the visual periphery

A. INTRODUCTION

Great progress has been made towards understanding the internal statistical models that guide our perceptual decision-making and corresponding confidence ratings. When we make perceptual decisions about the world around us, they are accompanied by a metacognitive sense of certainty, or confidence, in whether those percepts are correct. Confidence ought to depend on the strength (signal) and reliability (noise) of the evidence to make the decision, i.e., should correlate with these measures (Fleming and Daw, 2017; Green and Swets, 1966; Macmillan and Creelman, 2005; Pouget et al., 2016); however, sometimes it does not (Fetsch et al., 2014; Koizumi et al., 2015; Maniscalco et al., 2020, 2019, 2016; Morales et al., 2020; Peters et al., 2017a; Rahnev et al., 2015, 2012a, 2012b; Rounis et al., 2010; Samaha et al., 2017, 2016; Zylberberg et al., 2016, 2012) or does so but not in a Bayesian optimal way (Adler and Ma, 2018; Denison et al., 2018); these cases suggest a need for innovation in models of perceptual confidence. Here, we sought to determine how natural scene statistics about noise distributions in the visual field might explain why confidence does not always perfectly track performance -- beyond a supposition that metacognition might simply involve "more noise" over and above noise in first-order (Type 1) estimates (Maniscalco and Lau, 2016).

A particularly intriguing case in which confidence and accuracy can dissociate is the phenomenon of subjective inflation, which is particularly evident in the visual periphery: subjective feelings of decisional confidence for unattended stimuli in the visual periphery are higher than those for stimuli in the center of the visual field for both complex and simple stimuli when task accuracy is matched (Knotts et al., 2020; Li et al., 2018; Odegaard et al., 2018; Rahnev et al., 2011; Solovey et al., 2014). In other words, observers appear 'overconfident' in their peripheral percepts relative to central percepts of the same signal reliability (Ehinger et al., 2017; Gloriani and Schütz, 2019; Hess et al., 2008; Rosenholtz, 2016). One explanation given for these findings is that observers' Type 2 (metacognitive) criteria are fixed, such that under conditions of increased noise they do not adapt to the changing noise statistics of the environment, leading to a higher proportion of "high confidence" responses under a signal detection theoretic framework (Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014). However, the degree to which such Type 2 criteria are fixed remains open to some debate, as others have suggested that confidence judgments *are* sensitive to changing environmental or attentional conditions (Adler and Ma, 2018; Denison et al., 2018), albeit not in a Bayes-optimal manner. Importantly, it also remains unexplored why such explanations would provide stronger subjective inflation due to attentional manipulations in the visual periphery over central vision.

Other series of studies have suggested a conceptually analogous explanation: that illusions of confidence can be bidirectional, with both over- and under-confidence exhibited by the same system, due to a similar "fixed" estimate of noise, i.e. a "misperception" of variance (Gorea and Sagi, 2001; Peters et al., 2017a; Zylberberg et al., 2014). That is, one element of the metacognitive system's capacity to judge decisional accuracy is its ability to judge noise, i.e. signal reliability; by knowing a signal's or representation's noisiness, the observer can then set confidence criteria or engage in Bayesian inference appropriately. In cases of variance misperception, rather than directly measuring the reliability of its internal signals and using such an optimal measure in metacognitive calculations, the observer instead uses a heuristic -- a fixed, 'typical' level of noise applied unilaterally to metacognitive evaluations in the current task regardless of actually varying noise conditions (Gorea and Sagi, 2001; Peters et al., 2017a; Zylberberg et al., 2014) or otherwise possesses mistaken beliefs about noise levels ((Drugowitsch et al., 2014; Fleming and Daw, 2017); see also (Adams et al., 2013)). It has been shown that such variance misperceptions can explain bidirectional confidence errors relative to decisional accuracy, and are related to the fixed metacognitive (Type 2) criteria also mentioned above (Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014).

Despite these potential answers due to fixed confidence criteria or variance misperception, a core explanation for why peripheral inflation occurs remains elusive for at least two reasons. First, it remains unclear clear how such heuristic 'typical noise level' judgments might be achieved: by what mechanism does the system "pick" a typical level of noise, and how is it used to set confidence criteria? And second, as with fixed criteria, variance misperception alone doesn't necessarily explain why illusions of confidence appear to be easier to achieve in the periphery relative to central vision: on average, variance misperception (i.e., errors in noise beliefs) should lead to similar over- or under-confidence in general for both central and peripheral vision, unless the system were differentially applying misperceived variance in peripheral versus central vision. Instead, we see two main trends: subjective inflation in the periphery happens more than in the center under normal visual conditions (Knotts et al., 2020; Li et al., 2018; Odegaard et al., 2018; Rahnev et al., 2011; Solovey et al., 2014) (i.e., not in the dark; (Gloriani and Schütz, 2019)), and on average there is no inflation or deflation in central vision (Zylberberg et al., 2014).

Here, we sought to develop a simple explanation for these phenomena by combining these theoretical explanations with hypotheses about their potential source in empirical distributions of internal noise, in the same vein as the well-known influence of natural scene statistics on visual perception in general. Generally, the visual system and its perceptual computations are sensitive to external natural scene statistics (Adams et al., 2004; Girshick et al., 2011; Peters et al., 2015; Stocker and Simoncelli, 2006) and human observers adapt their internal models to such external environmental statistical properties (Seriès and Seitz, 2013) -- making use of knowledge that cardinal directions are more comment, light typically comes from above, or motion tends to be slow and smooth, for example. Knowledge of these external natural scene statistics drives the formation of prior beliefs in a Bayesian decision-making framework, which are combined with incoming sensory information to form Bayes-optimal percepts of motion speed, orientation, convexity, and so on (Adams et al., 2004; Girshick et al., 2011; Peters et al., 2015; Stocker and Simoncelli, 2006).

We extend this idea to suppose that the system is sensitive to internal 'natural scene statistics' or empirical priors about *its own noise*, i.e. that the visual system has learned typical distributions of internal signal noisiness as a function of a stimulus' location in the visual field. The system has likely learned that peripheral vision is noisier than central vision, forming the basis for an empirical prior over the reliability of internal representations of these differing parts of the visual field. However, while such empirical priors are based in fact, they can sometimes lead to inaccurate perceptual estimates when the system is required to use them in Bayes-optimal computations. For example, empirical priors about man-made versus natural objects' densities are used by the system in judging visuo-haptic weight estimates (Peters et al., 2015), but in ways that lead to illusory percepts of heaviness due to simplification of complex, continuous priors over density into mixtures of Gaussians (Peters et al., 2018, 2016) -- perhaps due to biological restrictions on information coding (Heng et al., 2020). It has also been shown that highly asymmetric (skewed) priors over stimulus location which are initially learned correctly can be relied upon as if they were Gaussian when the observer is required to use them in a cue combination task (Acerbi et al., n.d.), and that incorrect priors in general can lead to illusions across many areas of visual perception in both healthy and atypical perception (Flanagan et al., 2008; Geisler and Kersten, 2002; Teufel et al., 2013; Valton et al., 2019; Weiss et al., 2002). Thus, incorrect empirical priors can play a large role in ultimate percepts, even in first-order visual tasks.

We therefore reasoned that inflation of confidence in the visual periphery might likewise occur due to incorrect empirical priors about *noise* in the visual system as a function of visual field location, or at least incorrect usage of such priors. Specifically, we hypothesized that the true distribution of visual noise (variance or standard deviation of sensory signals) in the visual periphery is significantly positively skewed relative to central noise, due to mathematical properties of variance (i.e., it cannot be less than zero) and observation that peripheral vision is on average noisier (less reliable) than central vision. However, the metacognitive or even perceptual system may not be able to use this true prior to optimally estimate momentary noise levels, instead using a simplified, non-skewed prior for Bayesian noise estimation or even simply the mode of the prior, i.e. the most likely noise level to be encountered in the periphery. If the visual system uses incorrect priors, or otherwise incorrectly estimated the differences in reliability between central versus peripheral vision, such suboptimal introspection (Adler and Ma, 2018; Denison et al., 2018; Peters et al., 2017a) could lead to variance misperception (Zylberberg et al., 2014) and thus over- as well as underconfidence.

In this project, we used five different Bayesian-inspired model observers to systematically explore the potential consequences of such a process. Our findings demonstrate how such an environmental statistic coupled with certain metacognitive errors in estimation or usage of this prior -- even under near-optimal Bayesian inference about noise -- could explain peripheral subjective inflation, especially under inattention.

B. RESULTS

B.1. Conceptual description of models

We formalized the above intuition about the relative impact of different natural statistics of noise in central versus peripheral vision using signal detection & Bayesian decision theory (SDT/BDT) (Figure 1a).

B.1.i Type 1 decisions

We simulated an observer that engages in a simple discrimination task. On each trial, an observer samples the environment. The sample the observer receives may have come from a stimulus that has identity S_1, or one with identity S_2. Assuming Gaussian internal noise, these samples form Gaussian internal response distributions centered at the true mean for S_1 and S_2 with standard deviations governed by the combined internal noise of the observer and the external noise of the stimulus. For simplicity here we assume equal variances between the S_1 and S_2 distributions, following standard convention (Green and Swets, 1966; Macmillan and Creelman, 2005).

We modeled the distributions of these internal response signals as two Gaussian distributions representing cases where the signal came from S_1 or S_2 following standard SDT conventions (Green and Swets, 1966; Macmillan and Creelman, 2005) (Figure 1b). The discrimination criterion, which is hardcoded at the optimum in this simple system (halfway between the distributions to maximize performance, i.e. by setting the criterion where a sample is equally likely to have been generated by either stimulus class, which under equal priors corresponds to the location where the distributions intersect), is shown through a dashed vertical line. When the internal response variable falls above this threshold, the observer answers "S_2" to a discrimination question, and otherwise answers "S_1". The internal response criterion in the models below is specified in posterior probability ratio space, making this observer a Bayesian observer, to facilitate the confidence readout as the posterior probability of the choice that the observer made. See Methods for specific details.

To this simple Type 1 system we add an additional hierarchical or metacognitive (Type 2) inference layer to explore several hypotheses about how the noise in the system (both internal noise and external noise (Lu and Dosher, 2008)) governs the Type 1 decision space. Drawing inspiration from (1) Bayesian ideal observer analysis demonstrating that natural scene statistics govern Type 1 perception (Adams et al., 2004; Girshick et al., 2011; Peters et al., 2015; Stocker and Simoncelli, 2006), and (2) hierarchical models in which a Bayesian ideal observer's inference about latent variables also governs the ultimate percept (Knill and Richards, 1996; Knill and Saunders, 2003; Körding et al., 2007; Körding and Tenenbaum, 2007a; Odegaard et al., 2015; Peters et al., 2018, 2016; Samad et al., 2015; Yuille and Bülthoff, 1996), we developed and compared a series of 'flat' and 'hierarchical' inference models with varying 'knowledge' or reliance of natural scene statistics of noise in central versus peripheral vision to evaluate their predictions for peripheral inflation.

B.1.ii Environmental statistics of noise in central versus peripheral vision

A critical factor in an observer's ability to make accurate Type 1 decisions is the noisiness (variability) of the system itself: noise can cause the observer to respond "S_1" when a signal actually came from the S2 distribution, or vice versa, with varying levels of precision or error (Figure 1d). The noise in the system depends on the location in the visual field, central or peripheral, with central vision being more reliable (less noisy) than peripheral vision (Provis et al., 2013); (Gloriani and Schütz, 2019). It is therefore reasonable to assume that central and peripheral vision have different hypothesized underlying (parent) distributions governing the distributions (Lu and Dosher, 1998) of noise typically experienced in these two regions of the visual field (Figure 1b & c). Specifically, the variance (noise) corresponding to peripheral vision is assumed to be higher than for central vision.

A Bayesian observer that is sensitive to environmental statistics about a given variable will represent expectations about such variables as prior distributions -- for example, about light source location, motion speed, or contour orientation (Adams et al., 2004; Girshick et al., 2011; Stocker and Simoncelli, 2006), as mentioned above. Here, we hypothesized that central and peripheral environmental distributions of *noise* experienced by the visual system also lead the visual system to form prior expectations about variability as a latent variable, following previous convention in hierarchical Bayesian inference in vision and multisensory integration (Beierholm et al., 2009; Knill and Richards, 1996; Knill and Saunders, 2003; Körding et al., 2007; Körding and Tenenbaum, 2007a; Landy et al., 2011; Odegaard et al., 2015; Peters et al., 2018, 2016, 2015; Samad et al., 2015; Shams et al., 2000; Wozny et al., 2008; Yuille and Bülthoff, 1996). That is, the visual system learns to expect that central vision typically involves less noisy signals, while peripheral vision typically involves noisier signals.

Our hypothesis is further extended by the observation that variance is a property that cannot be mathematically negative, such that any distribution of variance must by definition exist in the domain $\sigma > 0$. When the distribution of variance is narrow and symmetric, a truncated Gaussian distribution (with domain of $\sigma > 0$) might be modestly appropriate to represent the empirical distribution for central noise (Figure 1a); that is, we hypothesized that the empirical distribution of noise is appropriately represented by a *symmetric* distribution around a typical (mode/most common) level of noise. However, in the periphery it is possible that a broader distribution of noise is experienced, especially under varied lighting conditions. Although it would be possible for such a distribution to be symmetric as in central vision, here we hypothesized a positively skewed empirical distribution for noise in the periphery (Figure 1b); that is, we hypothesized that when the variance does fluctuate it does so *asymmetrically* around a typical (mode/most common) level of noise. We note that although such a distribution is mathematically justified by properties of noise (variance), it has yet to be empirically demonstrated (see Discussion). Thus, the purpose of this project was to serve as a precursor to measuring such statistics by evaluating the impact of such natural statistics on decisions and confidence judgments, including how an observer's differential 'knowledge' of this true empirical distribution might explain puzzling behaviors such as peripheral subjective inflation.

[Figure 1 about here]

B.1.iii Intuitive introduction to models

Typically, the visual system is assumed to be optimal (Bejjanki et al., 2016; Landy et al., 2007) in that it uses an accurate estimate of its own noise because it has 'knowledge' of its own internal statistics (King and Dehaene, 2014; Lau, 2008); this optimality is assumed to propagate to the visual metacognitive system (Drugowitsch, 2016; Drugowitsch et al., 2019; Fleming and Daw, 2017), perhaps corrupted by some additional metacognitive noise (Maniscalco and Lau, 2016). However, a number of heuristic models have previously been shown to capture metacognitive behavior better than optimal models, including variance misperception (Peters et al., 2017a; Zylberberg et al., 2014) (this is also seen in cognitive decision-making; (Herce Castañón et al., 2019)), fixed Type 2 (metacognitive) criterion setting under varying noise (Li et al., 2018; Rahnev et al., 2012b, 2011; Solovey et al., 2014), a bias indicating overweighting evidence favoring a decision (Maniscalco et al., 2016; Peters et al., 2017b; Zylberberg et al., 2012), or suboptimal Type 2 criterion setting (Adler and Ma, 2018), among others. In particular, overconfidence relative to true Type 1 sensitivity is more often observed in the visual periphery than the center (Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014).

Combining these observations, we hypothesized that the visual system in both peripheral and central vision may not be able to accurately perceive the noisiness of the stimulus. To explore how various heuristic strategies could lead to overconfidence specific to the periphery, we defined and compared five candidate computational models, falling into two classes: *flat* models, which include no explicit metacognitive estimation of noise at all; and *hierarchical* models, which specify various algorithms for the metacognitive system to estimate the noise. Simple descriptions of each model are provided in Table 1; we provide full details of these models in Methods.

[Table 1 about here]

For each of these models, we estimated the behavior of the Type 1 and Type 2 decision-making system under different variances, i.e. different levels of noise in the signal, under central versus peripheral vision conditions. The critical factor to evaluate the models' performance is how each model makes metacognitive decisions relative to Type 1 performance capacity (which is dictated by the true noise regardless of the metacognitive noise estimation process). Model details are presented in Methods.

<u>B.3 Attentional manipulations</u>

Subjective inflation of confidence relative to performance has been shown previously to be stronger under cases of inattention (Li et al., 2018; Rahnev et al., 2011), such that -- somewhat counterintuitively -- increased sensory precision due to attention leads to a reduction in peripheral subjective inflation in particular. We next investigated the behavior of all five model observers under simulated conditions of increasing attention by assuming that attention may modify two factors: (1) the precision in the sensory estimate, and (2) the precision in the metacognitive estimate of noise. Details are presented in Methods.

B.4 Central versus peripheral confidence relative to performance (no attentional manipulation)

We first examined the behavior of the five simulated observers under 'default' conditions, i.e. using the parameters presented in Table 1 without any attentional manipulations ($a_1 = a_2 = 1$). The key measure is that of confidence inflation, measured here using the effect size measure of Cohen's *d*: mean difference between confidence (Equation 5) and the percent correct answers (performance governed by Equation 4), scaled by their pooled standard deviation, such that positive values of *d* indicate confidence > performance, while and negative values indicate confidence < performance. See Methods for details.

B.4.i Model F1: Flat Bayesian ideal observer

The flat Bayesian ideal observer (F1) followed optimal expected behavior in both central and peripheral vision (Figure 2a, first column). Under higher noise conditions performance (p(correct)) falls and this drop is mirrored in confidence estimates for both visual field locations. Thus, this observer displayed neither under- nor overconfidence on average (Figure 2b). The average effect size for F1 in the center was Cohen's *d* = -0.0104, and in the periphery was Cohen's *d* = -0.0054 -- that is, there is almost no overconfidence on average for either visual field.

B.4.ii Model F2: Flat fixed criterion heuristic observer

The flat fixed criterion heuristic observer (F2) also behaved as reported in previous literature (Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014) (Figure 2a, second column). Performance dropped as expected for increasingly noisy conditions in both central and peripheral vision, but due to the fixed Type 2 criterion (set here according to an average of the noise in central and peripheral visual fields) F2 displayed a general underconfidence bias under less noisy conditions and a general overconfidence bias under noisier conditions. Importantly, because peripheral vision is on average much noisier than central vision, the fixed confidence criterion led to biases towards underconfidence relative to performance in central vision, accompanied by significant peripheral inflation (overconfidence relative to performance) (Figure 2b). The average effect size for F2 in the center was Cohen's *d* = -0.8510 (i.e., strong underconfidence), and in the periphery was Cohen's *d* = 1.0273 (i.e., strong overconfidence). If the fixed criterion were instead set only according to central vision noise, this would translate to optimal confidence in central vision and extreme overconfidence in the periphery; however, this is not the focus of the present project, so we leave such explorations to future studies.

B.4.iii Model H1: Hierarchical Bayesian ideal observer

The hierarchical Bayesian ideal observer (H1) estimates the noise in its system optimally using both prior experience (optimal knowledge of the empirical prior governing noise in central versus peripheral vision) and an estimate of noise at the given moment. Thus, it also displayed similar behavior to model F1, albeit with slight bias towards underconfidence under extremely not-noisy conditions and towards overconfidence under extremely noisy conditions due to reliance on knowledge of the empirical distributions of noise (i.e., use of the prior distributions of noise) (Figure 2a, third column). Despite these slight biases, H1 displayed neither under- nor overconfidence on average in either central or peripheral vision (Figure 2b). The average effect size for H1 in the center was Cohen's d = -0.2832, and in the periphery was Cohen's d = -0.2212 -- that is, both central and peripheral vision displayed modest underconfidence, but in similar magnitude.

B.4.iv Model H2: Hierarchical 'mode prior' heuristic observer

The hierarchical 'mode prior' heuristic observer (H2) uses the most probable level of noise in each of the central and peripheral visual fields, based on its previous experience, rather than represent and use the entire empirical prior. Due to the fact that this empirical distribution is (here hypothesized to be) symmetrical around a typical noise level in central vision, this strategy leads H2 to exhibit on average no over- or underconfidence in the center of the visual field (Figure 1a, fourth column, top row; Figure 2b). In contrast, like F1 the H2 observer displayed overconfidence in the visual periphery -- but this time, it is due to the fact that the peripheral noise distribution is (here hypothesized to be) positively skewed. Thus, reliance on the mode of the empirical prior distribution led to overconfidence more often than underconfidence, and overconfidence on average in the visual periphery (Figure 1a, fourth column, bottom row; Figure 2b). The average effect size for H2 in the center was Cohen's d = -0.1522 (i.e., modest underconfidence) and in the periphery was Cohen's d = 0.7287 (i.e., strong overconfidence).

B.4.v Model H3: Hierarchical 'Gaussian assumption' heuristic observer

Finally, we examined the predicted behavior of the hierarchical 'Gaussian assumption' heuristic observer (H3). H3 is nearly identical to H1 (the hierarchical Bayesian ideal observer) with the important difference that H3 uses a Gaussian estimate for the empirical distribution of noise in the visual periphery. Thus, while H3 displayed identical behavior to H1 in the center (Figure 2a, fifth column, top row; Figure 2b), this Gaussian assumption for peripheral noise led H3 to more often under- than over-estimate noise in the visual periphery, leading once again to peripheral overconfidence on average (Figure 2a, fifth column, bottom row; Figure 2b). Thus, even though H3 possesses knowledge of the scale *and* location of the empirical distribution of noise in the visual periphery, its assumption that this distribution is symmetrical led to an average under-estimation of noise and therefore overconfidence relative to performance. The average effect size for H3 in the center was Cohen's d = -0.2857 (i.e., modest underconfidence because it is identical to H1), but in the periphery was Cohen's d = 0.4846 (i.e., medium-strong overconfidence).

[Figure 2 about here]

B.5 Attentional manipulations

Given that several of the model observers we tested displayed overconfidence in the visual periphery, we next turned to attentional manipulations as a possible avenue for arbitrating models' fit to previously-reported effects. In particular, it has previously been reported that attentional manipulations may affect peripheral inflation in a somewhat counterintuitive manner, with high-attention conditions leading to less strong peripheral overconfidence and inattention causing increases in subjective inflation in the periphery (Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014). Thus, we simulated anticipated effects of increasing attention as decreases in sensory noise and increases in metacognitive precision (see Methods). We quantified the effect size of inflation, i.e. the difference between performance and confidence, as Cohen's d, with positive values indicating confidence > performance and negative values indicating confidence < performance.

As expected, increasing attentional allocation increased sensory precision and thus increased performance in both central and peripheral vision (Figure 3a,b; Table 3). However, this was generally accompanied by a relatively flat or if anything slightly increasing effect size (Cohen's d) for subjective inflation (confidence overestimating performance) for both central and peripheral vision (Figure 3c,d; Table 3), with the exceptions of model F2 in central vision (Figure 3c; Table 3) and models H2 and H3 in peripheral vision (Figure 3d; Table 3).

In particular, while model F2 (flat fixed criterion heuristic observer) showed increasingly strong *under*confidence in central vision due to inflation (despite increasing performance), it displayed flat or if anything increasing overconfidence in peripheral vision. If a correction were to be applied such that central Cohen's d for inflation were to remain flat across increasing attentional allocation, this would translate to *increasing* inflation in the visual periphery -- the opposite of empirical reports in the literature. On the other hand, models H2 and H3 showed flat inflation in central vision due to increasing attentional allocation, but if anything slightly *decreasing* overconfidence in the visual periphery. Qualitatively, this behavior matches reports from the literature (Li et al., 2018; Rahnev et al., 2011).

[Figure 3 about here]

[Table 2 about here]

C. DISCUSSION

Here, we asked how knowledge and use of noise statistics across the visual field might inform metacognitive judgments of perceptual decisions. We specifically focused on the phenomenon of visual subjective inflation in the visual periphery -- the phenomenon wherein confidence judgments overestimate decisional accuracy -- including how this phenomenon interacts with attentional manipulations. We hypothesized that the metacognitive system makes use of knowledge of noise statistics in central vision and peripheral vision differently, and that errors in estimation of noise or in use of this prior knowledge might explain why peripheral inflation occurs, and why it becomes stronger under conditions of inattention. Specifically, we proposed that peripheral noise statistics might exhibit a positively-skewed empirical distribution while central noise statistics show a symmetric noise distribution, and that metacognitive heuristics might lead the visual system to systematically overestimate confidence relative to performance in the visual periphery.

We compared five possible model observers which estimate noise and make decisions and confidence in various ways according to Bayesian computations: two *flat* observers that possessed no knowledge of natural statistics of noise across the visual field, and three *hierarchical* observers that used such prior knowledge in varying ways. Using Monte Carlo simulations and simple parameter choices as a proof of concept, we showed that Bayesian ideal observer models that either have optimal access to true noise in a sample (model F1: flat Bayesian ideal observer) or optimally estimate noise using accurate empirical noise priors in the center and periphery (i.e., 'knew' the peripheral empirical noise distribution was positively skewed; model H1: hierarchical Bayesian ideal observer) predict no over- or underconfidence at all, instead showing confidence behavior that tracks decisional accuracy essentially perfectly. In contrast, heuristic observers that either used a fixed Type 2 criterion for both central and peripheral judgments (model F2: flat fixed criterion heuristic observer), used the most likely noise level in central or peripheral vision (model H2: hierarchical 'mode prior' heuristic observer), or used a Gaussian estimate for the peripheral noise prior (model H3: hierarchical 'Gaussian assumption heuristic observer) predicted clear patterns of underconfidence under less noisy conditions and overconfidence under noisier conditions; because of the hypothesized positive skew of empirical noise in the visual periphery, this led these observers to display peripheral inflation. However, only model F2 predicted underconfidence in central vision and overconfidence in the periphery, while the two hierarchical heuristic observers (H2 and H3) predicted near-perfect confidence in the center and overconfidence in the periphery. And crucially, only H2 and H3 also correctly predicted that increasing sensory and metacognitive precision due to attentional allocation ought to decrease peripheral inflation.

Our results may explain findings reported in previous empirical work showing overconfidence in the visual periphery (Li et al., 2018; Odegaard et al., 2018; Rahnev et al., 2011; Solovey et al., 2014). They also further extend previous signal detection and Bayesian models of peripheral inflation and errors in variance estimation (Fleming and Daw, 2017; Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014; Zylberberg et al., 2014) to include explicit formulation and exploration of computations that may underlie metacognitive estimates of noise, especially how such mechanisms may rely on natural statistics and previous experience. This formulation thus places the metacognitive system in the same conceptual space as other hierarchical models of perceptual inference which use prior knowledge of natural scene statistics (Adams et al., 2004; Girshick et al., 2011; Landy et al., 2011; Stocker and Simoncelli, 2006; Yuille and Bülthoff, 1996) and which estimate distributions and values for latent variables. Previous work in this area has posited that such hierarchical Bayesian inference might underlie visual inferences about shape by first estimating luminance or specularity (Kersten et al., 2004; Yuille and Bülthoff, 1996), visual inferences about planar slant by first estimating texture isotropy (Knill and Saunders, 2003), multisensory inferences about object heaviness by first estimating relative density (Peters et al., 2018, 2016), or multisensory inferences about numerosity (Shams et al., 2000; Wozny et al., 2008), spatial location (Odegaard et al., 2015; Wozny et al., 2010; Wozny and Shams, 2011), body ownership (Samad et al., 2015), or sensorimotor signal processing (Körding and Tenenbaum, 2007b; Körding and Wolpert, 2003; Wei and Körding, 2011) by first estimating causal relationships among multimodal signals (Körding et al., 2008, 2007; Shams and Beierholm, 2010), among others.

Other previous studies have demonstrated that seemingly suboptimal estimation of noise leading to metacognitive errors may stem from a heuristic estimate of noise rather than a fixed estimate (as in the fixed Type 2 criterion models (Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014), simple variance misperception (Zylberberg et al., 2014)), or belief errors about noise (Fleming and Daw, 2017). It has also been suggested that a metacognitive update rule which takes into account changing reliability due to sensory noise (Adler and Ma, 2018) or attentional manipulations (Denison et al., 2020) in a linear or quadratic manner rather than fully Bayes-optimal fashion can explain decisional confidence estimates in simple perceptual tasks. These latter studies found that while the fixed criterion type models fail to capture behavior, models that assume the metacognitive system is suboptimally sensitive to changing noise in this particular heuristic fashion can better mimic human behavior. However, these models have not been applied to the differential confidence errors found between central versus peripheral vision, nor do they explain how specifically the shape of the empirical noise priors in central versus peripheral vision (and its mismatch with the observers *beliefs* about such priors) would impact such metacognitive errors. Instead, here we show that Bayesian inference -- especially for model H3, which assumes that the perceptual system possesses an erroneous prior expectation about the shape of the distribution of noise itself -- produces empirically-observed errors in metacognitive estimation as well as the critical pattern of increasing peripheral inflation under increasing inattention. Future work should formally compare the models we explored here with the linear, quadratic, and other heuristic models proposed by other authors in an empirical dataset.

Our results suggest that an efficient coding scheme for expectations about noise conditioned on visual field location, based on empirical priors, may contribute to peripheral inflation of confidence. In particular, that hierarchical models H2 and H3 showed peripheral inflation suggests that the system's prior about noise in the visual periphery may not reflect the true noise experienced, but instead a simplified representation of the most likely noise level to be experienced by the observer in the visual periphery (the mode of the empirical prior) in addition to an impoverished metacognitive ability to estimate such noise in the periphery as compared to the center ($\varsigma_{center} < \varsigma_{periphery}$) . This simplification of prior expectations for complex, skewed, or otherwise non-Gaussian empirical priors into (mixtures of) Gaussians (sometimes formulated as competing priors) has been noted previously (Knill, 2007, 2003; Knill and Saunders, 2003; Yuille and Bülthoff, 1996), and has been suggested to underlie other perceptual illusions (Peters et al., 2018, 2016). Incorrect priors in general, regardless of simplification, can lead to illusions in many areas of perception (Flanagan et al., 2008; Geisler and Kersten, 2002; Teufel et al., 2013; Valton et al., 2019; Weiss et al., 2002), and even if skewed priors are initially learned appropriately it seems the system has difficult using them optimally in cue integration (Acerbi et al., n.d.). Therefore, perhaps due to biological constraints restricting information coding (Heng et al., 2020), we've shown here how it is possible that in metacognition, too, efficient, simplified coding of noise expectations in the periphery may lead to overconfidence relative to performance capacity.

Our findings here also suggest other fruitful avenues for future study. In particular, they suggest that the natural statistics of noise in the visual periphery may be highly positively skewed. Validating this assumption would likely require comprehensive measurement of experienced noise statistics in central versus peripheral vision across a large range of stimulus types. A strong challenge to this endeavor would be that the natural statistic to be measured is *experienced* noise in the visual system rather than natural statistics of noise in the environment. Thus, one promising future approach might be to use established methods for measuring additive and multiplicative internal noise, e.g. through the triple-threshold-versus-contrast function (triple-TVC) approach (Dosher and Lu, 2017, 2000; Lu and Dosher, 1998, 2008) and double-pass procedures (Awwad Shiekh Hasan et al., 2012; Gold et al., 2004; Levi and Klein, 2003; Ratcliff et al., 2018; Vilidaite and Baker, 2017) to quantify internal noise. It would be necessary to conduct these procedures under a range of attentional manipulations, in both the center and visual periphery, and across a large range of tasks *in the same observer* (i.e., a within-subjects design) in order to accurately measure the shape of the internal noise distributions. Although this approach is daunting, smaller studies may make headway by comparing a few tasks at a time. Unfortunately, such an endeavor is beyond the scope of the current project, which aims to provide an exploratory proof of concept for how noise priors might be used by the metacognitive system to result in metacognitive errors in noise estimation; we therefore leave such studies for future research.

We admit the present study is limited in its scope due to its nature as proof-of-concept simulations only, without comprehensive parameter fitting or behavioral data; this of course limits its utility, although it provides a principled jumping off point for future empirical studies, as has been done previously (see e.g. (Fleming and Daw, 2017)). These models also do not take into account more nuanced differences between central versus peripheral vision in terms of other factors, such as crowding or visual search, differential impact by attentional manipulation, and so on (Rosenholtz, 2016; Rosenholtz et al., 2012a, 2012b). However, despite the models' simplicity, the results shown here pave the way for integrating metacognitive noise or uncertainty estimation with a long and established history of hierarchical Bayesian models in perception and cue combination by explicating the specific hypothesis that the metacognitive system builds prior distributions of expected noise that are sensitive not only to experienced ‘environmental’ (within itself) noise statistics but are also sensitive to attentional manipulations, visual field location, and other contextual modulations. We therefore believe that our results provide an important step in realizing the power of such modeling frameworks and empirical approaches for fully explaining how metacognitive computations are performed, and how they may be implemented in neural architecture.

D. METHODS

D.2 Formal computational models

D.2.i Type 1 decisions

Two Gaussian distributions represent the internal response distributions for signal and noise conditions, respectively (Figure 1b). To make a decision, the conditional probability of the source being present given an internal response variable on that trial is calculated through Bayes rule:

$$p(S|x;\ \mu,\ \sigma) = \frac{p(x|S)p(S)}{p(x)} \tag{1}$$

where *S* is source (S_1 or S_2), and *x* is the internal response, with

$$\begin{aligned} x_{S_1} &\sim N(\mu_{S_1}, \sigma) \\ x_{S_2} &\sim N(\mu_{S_2}, \sigma) \end{aligned} \tag{2}$$

where μ_{S_2} is the mean of the internal response to a signal from distribution S2 with a given strength, $\mu_{S_1} = -\mu_{S_2}$ following previous convention , and σ represents their (assumed to be equal) standard deviations under the simplest implementation. Prior probabilities for sources S1 and S2 are defined by

$$p(S_1) = p(S_2) = 0.5 \tag{3}$$

The decision made (D) is determined by the following

$$\begin{aligned} D = S_1 &\text{ if } p(S_1|x) > 0.5 \\ S_2 &\text{ if } p(S_2|x) \geq 0.5 \end{aligned} \tag{4}$$

where 0.5 serves as the optimal criterion in an SDT framework (set to maximize percent correct choices). That is, the criterion is set so the observer selects S_i according to which is most likely to be correct.

The standard convention is that confidence is computed via

$$C = p(S_{chosen}|x) \tag{5}$$

The models below selectively alter elements of this standard process, with focus on Equations 1 and 2.

D.2.i Empirical prior distributions governing noise in central versus peripheral vision

The standard deviations of both the x_{S_1} and x_{S_2} internal response distributions are set to the same σ (Equation 2) within a given visual field location, governed by the noise distributions at each location: central vision's empirical distribution of standard deviations (i.e., the natural statistics of noise in central vision) is symmetric (Gaussian), while peripheral vision's is positively skewed:

$$\sigma_{center} \sim N(M_{center}, \Sigma_{center}) \tag{6}$$
$$\sigma_{periphery} \sim logN(M_{periphery}, \Sigma_{periphery}) \tag{7}$$

where M_{center} and $M_{periphery}$ represent the mean (or log mean) of true standard deviations experienced by central and peripheral visual fields across a wide range of situations, respectively, Σ_{center} and $\Sigma_{periphery}$ represent the respective variabilities in these empirical distributions, and σ is always constrained to be in domain $\sigma > 0$ (i.e., Equation 6 is a truncated normal distribution). Thus, central vision experiences a symmetric distribution of noise around a central mean (Figure 1a), while peripheral vision experiences on average higher noise and with positive skew (Figure 1b).

D.2.ii Definitions of Flat and Hierarchical models of metacognition

F1: Flat Bayesian ideal observer
This observer 'knows' the true noise in its system, separately for central and peripheral vision, without any explicit metacognitive noise estimation process or knowledge of empirical noise distributions. This is the standard model employed in simple signal detection and Bayesian decision systems, with true knowledge of σ. This model calculates the Type 1 decision via Equations 1, 2 and 4, and confidence via Equation 5 as written above.

F2: Flat fixed criterion heuristic observer

This observer sets a single estimate for the noise in the environment across both central and peripheral vision based on optimal 'knowledge' of both, without any explicit metacognitive noise estimation process or knowledge of empirical noise distributions. Thus, it is conceptually akin to the 'fixed Type 2 criterion' model described in previous literature (Li et al., 2018; Rahnev et al., 2011; Solovey et al., 2014). This model alters Equations 1 & 2 such that all decisions and confidence judgments for both central and peripheral vision are based on a single heuristic estimate of σ, $\hat{\sigma}_{F2}$:

$$\hat{\sigma}_{F2} = \frac{\sigma_{center} + \sigma_{periphery}}{2} \quad (8)$$

Thus, Equation 1 is rewritten as:

$$p(S|x;\ \mu,\ \hat{\sigma}_{F2}) = \frac{p(x|S)p(S)}{p(x)} \quad (9)$$

with Equation 2 rewritten as the observer *assuming*

$$\begin{aligned} x_{S_1} &\sim N(\mu_{S_1}, \hat{\sigma}_{F2}) \\ x_{S_2} &\sim N(\mu_{S_2}, \hat{\sigma}_{F2}) \end{aligned} \quad (10)$$

Decisions and confidence judgments are made according to Equations 4 and 5 as above.

H1: Hierarchical Bayesian ideal observer

This observer has optimal knowledge of the empirical distributions of noise in central versus peripheral vision, including the positive skew of peripheral noise. It uses Bayes' rule to optimally estimate the actual noise experienced at a given moment based on these prior expectations and the likelihood experienced at a given moment, separately in central and peripheral vision. That is, this model alters Equations 1 & 2 such that all decisions and confidence judgments for both central and peripheral vision are based on a optimal estimates of σ, $\hat{\sigma}_{H1}$, with one estimate for each of center and periphery:

$$p(\sigma|\sigma_{H1}) = \frac{p(\sigma_{H1}|\sigma)p(\sigma)}{p(\sigma_{H1})} \quad (11)$$

with assumed Gaussian noise in estimating self-noise, i.e. $p(\sigma_{H1}|\sigma)$ follows a normal distribution with mean σ (the sample of noise drawn on this particular trial) and standard deviation of this likelihood distribution defined as ς -- again with separate definitions ς for each of the (separable) center and periphery. $p(\sigma)$ refers to the empirical prior distribution of noise for a given visual field location (center or periphery). That is, there is an amount of noise that happens on this trial, but the observer does not have 'perfect' access to this noise; instead, it has a noisy representation of this noise (akin to how one would have a noisy representation of some other aspect of a physical stimulus [length, location, size, speed, etc.] which is combined with prior expectations for that noise (Alais and Burr, 2004; Beierholm et al., 2009; Burge and Girshick, 2010; Girshick et al., 2011; Knill, 2007, 2003; Knill and Richards, 1996; Knill and Saunders, 2003; Körding et al., 2007; Landy et al., 2011; Odegaard et al., 2015; Peters et al., 2018; Weiss et al., 2002; Wozny et al., 2010, 2008; Yuille and Bülthoff, 1996)). The estimate for $\hat{\sigma}_{H1}$ is chosen as the maximum a posteriori estimate, i.e.

$$\hat{\sigma}_{H1} = argmax_{\hat{\sigma}_{H1}} p(\sigma_{H1}|\sigma) \tag{12}$$

Subsequently, decisions and confidence are made as above by altering Equations 1 & 2 to read for each of the central and peripheral judgments:

$$p(S|x;\ \mu,\ \hat{\sigma}_{H1}) = \frac{p(x|S)p(S)}{p(x)} \tag{13}$$

$$\begin{aligned} x_{S_1} &\sim N(\mu_{S_1}, \hat{\sigma}_{H1}) \\ x_{S_2} &\sim N(\mu_{S_2}, \hat{\sigma}_{H1}) \end{aligned} \tag{14}$$

and then following Equations 4 and 5 as above to complete the readout.

H2: Hierarchical 'mode prior' heuristic observer

This observer uses an estimate of the most likely noise experienced at a given moment, separately for central and peripheral vision. Rather than optimally estimating the noise as the H1 observer does, it instead uses the mode of the empirical prior distribution as a heuristic estimate -- it picks the most likely amount of noise experienced at this location in the visual field. That is, rather than using σ, this observer sets $\hat{\sigma}_{H2}$ via

$$\hat{\sigma}_{H2} = argmax_{\sigma}\ p(\sigma) \tag{15}$$

again, separately for center and periphery as before. Then, just as H1, we redefine decisions and confidence by rewriting Equations 1 & 2 to reflect these assumptions for each of the central and peripheral judgments:

$$p(S|x;\ \mu,\ \hat{\sigma}_{H2}) = \frac{p(x|S)p(S)}{p(x)} \tag{16}$$

$$\begin{aligned} x_{S_1} &\sim N(\mu_{S_1}, \hat{\sigma}_{H2}) \\ x_{S_2} &\sim N(\mu_{S_2}, \hat{\sigma}_{H2}) \end{aligned} \tag{17}$$

and then following Equations 4 and 5 as above to complete the readout.

H3: Hierarchical 'Gaussian assumption' heuristic observer
This observer uses Bayes' rule to estimate the noise experienced at a given moment (similar to H1), but assumes that the empirical distribution of noise in peripheral vision is Gaussian rather than using the true positively-skewed empirical prior. (It also uses a Gaussian assumption for central vision, but since this is accurate, the central vision H3 model is identical to the central vision H1 model.) We assume that the observer understands the general location of this peripheral noise distribution, but possesses poorer knowledge of its variability or level of skewness.

Thus, this observer first estimates the location (mode) of the empirical prior governing peripheral noise as:

$$\hat{M}_{periphery} = argmax_{\sigma_{periphery}} p(\sigma_{periphery}) = e^{M_{periphery} - \Sigma^2_{periphery}} \tag{18}$$

which follows from the fact that the peripheral noise distribution is lognormal (Equation 7). That is, the location (mean and mode) of the "estimated", symmetrical prior for σ in the visual periphery is computed so as to match the most probable noise level (mode) of the "true," asymmetrical prior.

Variability in the estimated peripheral noise empirical prior distribution is then set to some multiple of the variability in the central noise prior:

$$\hat{\Sigma}_{periphery} = k\Sigma_{center} \tag{19}$$

The strength of this assumption (i.e., the magnitude of *k*, the variability assumed in the peripheral noise priori) controls the strength of peripheral inflation but does not qualitatively affect results. Because the purpose of the present project is to provide a proof of concept, we do not fit *k* here, and set *k* = 1 for simplicity. Future work ought to fit *k* to empirical data.

Then, we redefine the *estimated* empirical prior for noise in the periphery (Equation 7) for use by this observer as:

$$\sigma^*_{periphery} \sim N(\hat{M}_{periphery}, \hat{\Sigma}_{periphery}) \tag{20}$$

(The corresponding distribution for σ^*_{center} is equivalent to σ_{center}, since $\hat{M}_{center} = M_{center}$ and $\hat{\Sigma}_{center} = \Sigma_{center}$ due to the correct expectation Gaussian noise distributions in the visual center.) Then, just as with H1, all decisions and confidence judgments for both central and peripheral vision are based on a optimal estimates of σ (now based on the incorrect prior expectations for the distribution of σ^*), $\hat{\sigma}_{H3}$, again for each of center and periphery:

$$p(\sigma^*|\sigma_{H3}) = \frac{p(\sigma_{H3}|\sigma^*)p(\sigma^*)}{p(\sigma_{H3})} \tag{21}$$

with assumed Gaussian noise in estimating self-noise, i.e. $p(\sigma_{H3}|\sigma^*)$ follows a normal distribution with mean σ^* and standard deviation ς, again for each of the center and periphery. That is, $p(\sigma^*)$ now refers to the *incorrect* expected prior distribution of noise for a given visual field location (center or periphery). As with H1, the estimate for $\hat{\sigma}_{H3}$ is chosen as the maximum a posteriori estimate, i.e.

$$\hat{\sigma}_{H3} = argmax_{\sigma^*}\, p(\sigma_{H3}|\sigma^*) \tag{22}$$

Subsequently, as with H1, decisions and confidence are made by altering Equations 1 & 2 to reflect these assumptions, reading for each of the central and peripheral judgments:

$$p(S|x;\ \mu,\ \hat{\sigma}_{H3}) = \frac{p(x|S)p(S)}{p(x)} \tag{23}$$

$$\begin{aligned} x_{H3,signal} &\sim N(\mu_{signal}, \hat{\sigma}_{H3}) \\ x_{H3,noise} &\sim N(\mu_{noise}, \hat{\sigma}_{H3}) \end{aligned} \tag{24}$$

and then following Equations 4 and 5 as above to complete the readout.

D.3 Attentional manipulations

As described in Results, we investigated model predictions under simulated attentional manipulations assuming that attention may modify two factors: (1) the precision in the sensory estimate (i.e., the means of the empirical distributions of noise, M), and (2) the precision in the metacognitive estimate of noise (i.e., the standard deviations of the likelihood distributions of noise, ς). For this simple proof of concept, we introduced two additional scalar multiplicative factors that could modify these two parameters for both Type 1 and Type 2 (metacognitive) judgments: a_1 and a_2 for each noise factor, respectively. That is, M is actually defined as $a_1 M$ and ς as $a_2 \varsigma$, with $a_1 = a_2 = 1$ under default conditions. Thus, under increasing attention, a_1 and a_2 are both assumed to decrease, increasing the precision of the Type 1 samples themselves as well as the Type 2 (metacognitive) estimates about noise. Increasing attention is referred to as ‘attention+’ and ‘attention++’ with reference to ‘neutral’ (no attentional manipulation) in Results.

D.4 Simulation details

We used Monte Carlo simulations to evaluate the performance and confidence behavior of all five models described above, assuming the same signal strength in all scenarios. Parameters for all simulations are presented in Table 3; although these parameter values are arbitrary for the purposes of demonstration, the models' qualitative behavior is similar across a range of parameter values (data not shown).

[Table 3 about here]

For each of the above-described model observers (F1, F2, H1, H2, H3), we sampled 1000 standard deviations σ from each of the central and peripheral distributions (samples with $\sigma < 0$ were set to $\sigma = eps =$ 2.2204e-16); at each sampled standard deviation we calculated the performance (% correct) and average confidence (mean *C*) for each model across 50,000 simulated trials (25,000 signal and 25,000 noise). As described in Results, we quantified over- and underconfidence using Cohen's *d* as a measure of effect size, with positive values indicating confidence > performance and negative values indicating confidence < performance:

$$Cohen's\ d = \frac{\mu_C - \mu_{\%\ correct}}{\sigma_{pooled}} \tag{25}$$

with μ_C the mean of confidence judgments (Equation 5), $\mu_{\%\ correct}$ the proportion of correct responses, and σ_{pooled} their pooled standard deviation (for paired samples) according to standard definitions. All simulations were carried out through custom scripts written in Matlab R2019b.

E. ACKNOWLEDGEMENTS

This work was supported by the Canadian Institute for Advanced Research Azrieli Global Scholars Program in Brain, Mind, & Consciousness (to MAKP).

F. REFERENCES

Acerbi L, Marius't Hart B, Behbahani FMP, Peters MAK. n.d. Optimality under fire: Dissociating learning from Bayesian integration.

Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. 2013. The computational anatomy of psychosis. *Front Psychiatry* **4**:47.

Adams WJ, Graf EW, Ernst MO. 2004. Experience can change the'light-from-above'prior. *Nat Neurosci* **7**:1057–1058.

Adler WT, Ma WJ. 2018. Comparing Bayesian and non-Bayesian accounts of human confidence reports. *PLoS Comput Biol* **14**:e1006572.

Alais D, Burr D. 2004. The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* **14**:257–262.

Awwad Shiekh Hasan B, Joosten E, Neri P. 2012. Estimation of internal noise using double passes: does it matter how the second pass is delivered? *Vision Res* **69**:1–9.

Beierholm U, Quartz S, Shams L. 2009. Bayesian priors are encoded independently from likelihoods in human multisensory perception. *J Vis* **9**:1–9.

Bejjanki VR, Knill DC, Aslin RN. 2016. Learning and inference using complex generative models in a spatial localization task. *J Vis* **16**:9.

Burge J, Girshick A. 2010. Visual–Haptic Adaptation Is Determined by Relative Reliability. *J Neurosci*.

Denison RN, Adler WT, Carrasco M, Ma WJ. 2018. Humans incorporate attention-dependent uncertainty into perceptual decisions and confidence. *Proc Natl Acad Sci U S A* **115**:11090–11095.

Denison RN, Block N, Samaha J. 2020. What do models of visual perception tell us about visual phenomenology? doi:10.31234/osf.io/7p8jg

Dosher BA, Lu ZL. 2017. Visual Perceptual Learning and Models. *Annu Rev Vis Sci* **3**:343–363.

Dosher BA, Lu ZL. 2000. Noise exclusion in spatial attention. *Psychol Sci* **11**:139–146.

Drugowitsch J. 2016. Becoming Confident in the Statistical Nature of Human Confidence Judgments. *Neuron* **90**:425–427.

Drugowitsch J, Mendonça AG, Mainen ZF, Pouget A. 2019. Learning optimal decisions with confidence. *PNAS*. doi:10.1073/pnas.1906787116

Drugowitsch J, Moreno-Bote R, Pouget A. 2014. Relation between belief and performance in perceptual decision making. *PLoS One* **9**:e96511.

Ehinger BV, Häusser K, Ossandón JP, König P. 2017. Humans treat unreliable filled-in percepts as more real than veridical ones. *Elife* **6**. doi:10.7554/eLife.21761

Fetsch CR, Kiani R, Shadlen MN. 2014. Predicting the Accuracy of a Decision: A Neural Mechanism of Confidence. *Cold Spring Harb Symp Quant Biol* **79**:185–197.

Flanagan JR, Bittner J, Johansson RS. 2008. Experience can change distinct size-weight priors engaged in lifting objects and judging their weights. *Curr Biol* **18**:1742–1747.

Fleming SM, Daw ND. 2017. Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychol Rev* **124**:91–114.

Geisler WS, Kersten D. 2002. Illusions, perception and Bayes. *Nat Neurosci*.

Girshick AR, Landy MS, Simoncelli EP. 2011. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat Neurosci* **14**:926–932.

Gloriani AH, Schütz AC. 2019. Humans Trust Central Vision More Than Peripheral Vision Even in the Dark. *Curr Biol* **29**:1206–1210.e4.

Gold JM, Sekuler AB, Bennett PJ. 2004. Characterizing perceptual learning with external noise. *Cogn Sci* **28**:167–207.

Gorea A, Sagi D. 2001. Disentangling signal from noise in visual contrast discrimination. *Nat Neurosci* **4**:1146–1146.
Green DM, Swets JA. 1966. Signal Detection Theory and Psychophysics. New York: John Wiley & Sons, Inc.
Heng JA, Woodford M, Polania R. 2020. Efficient sampling and noisy decisions. *Elife* **9**. doi:10.7554/eLife.54962
Herce Castañón S, Moran R, Ding J, Egner T, Bang D, Summerfield C. 2019. Human noise blindness drives suboptimal cognitive inference. *Nat Commun* **10**:1719.
Hess RF, Baker DH, May KA, Wang J. 2008. On the decline of 1st and 2nd order sensitivity with eccentricity. *J Vis* **8**:19.1–12.
Kersten D, Mamassian P, Yuille A. 2004. Object perception as Bayesian inference. *Annu Rev Psychol* **55**:271–304.
King J-R, Dehaene S. 2014. A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philos Trans R Soc Lond B Biol Sci* **369**:20130204.
Knill DC. 2007. Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *J Vis* **7**:5:1–24.
Knill DC. 2003. Mixture models and the probabilistic structure of depth cues. *Vision Res* **43**:831–854.
Knill DC, Richards W. 1996. Perception as Bayesian inference. Cambridge University Press.
Knill DC, Saunders JA. 2003. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res* **43**:2539–2558.
Knotts JD, Michel M, Odegaard B. 2020. Defending subjective inflation: An inference to the best explanation. *PsyArxiv*. doi:10.31234/osf.io/fhywz
Koizumi A, Maniscalco B, Lau H. 2015. Does perceptual confidence facilitate cognitive control? *Atten Percept Psychophys* **77**:1295–1306.
Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L. 2007. Causal Inference in Multisensory Perception. *PLoS One* **2**:e943–e943.
Körding KP, Shams L, Ma WJ. 2008. Comparing Bayesian models for multisensory cue combination without mandatory integration. *Adv Neural Inf Process Syst*.
Körding KP, Tenenbaum JB. 2007a. Causal inference in sensorimotor integration. *NIPS*.
Körding KP, Tenenbaum JB. 2007b. Causal inference in sensorimotor integration In: Schölkopf B, Platt JC, Hoffman T, editors. Advances in Neural Information Processing Systems 19. MIT Press. pp. 737–744.
Körding KP, Wolpert D. 2003. Probabilistic inference in human sensorimotor processing. *Adv Neural Inf Process Syst* **16**.
Landy MS, Banks MS, Knill DC. 2011. Ideal-Observer Models of Cue Integration. *Sensory Cue Integration*. doi:10.1093/acprof:oso/9780195387247.003.0001
Landy MS, Goutcher R, Trommershäuser J, Mamassian P. 2007. Visual estimation under risk. *J Vis* **7**:4.
Lau HC. 2008. A higher order Bayesian decision theory of consciousness. *Prog Brain Res* **168**:35–48.
Levi DM, Klein SA. 2003. Noise provides some new signals about the spatial vision of amblyopes. *J Neurosci* **23**:2522–2526.
Li MK, Lau H, Odegaard B. 2018. An investigation of detection biases in the unattended periphery during simulated driving. *Atten Percept Psychophys* **80**:1325–1332.
Lu ZL, Dosher B. 1998. External noise distinguishes attention mechanisms. *Vision Res*.
Lu ZL, Dosher BA. 2008. Characterizing observers using external noise and observer models:

assessing internal representations with external noise. *Psychol Rev* **115**:44–82.
Macmillan NA, Creelman CD. 2005. Detection theory: A user's guide, 2nd ed **2**:492.
Maniscalco B, Castaneda OG, Odegaard B, Morales J, Rajananda S, Peters MAK. 2020. The metaperceptual function: Exploring dissociations between confidence and task performance with type 2 psychometric curves. doi:10.31234/osf.io/5qrjn
Maniscalco B, Lau H. 2016. The signal processing architecture underlying subjective reports of sensory awareness. *Neuroscience of Consciousness* 1–41.
Maniscalco B, Odegaard B, Grimaldi P, Cho SH, Basso MA, Lau H, Peters MAK. 2019. Tuned normalization in perceptual decision-making circuits can explain seemingly suboptimal confidence behavior. *bioRxiv*. doi:10.1101/558858
Maniscalco B, Peters MAK, Lau H. 2016. Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. *Atten Percept Psychophys*. doi:10.3758/s13414-016-1059-x
Morales J, Odegaard B, Maniscalco B. 2020. The Neural Substrates of Conscious Perception without Performance Confounds. *philpapers.org*.
Odegaard B, Chang MY, Lau H, Cheung S-H. 2018. Inflation versus filling-in: why we feel we see more than we actually do in peripheral vision. *Philos Trans R Soc Lond B Biol Sci* **373**. doi:10.1098/rstb.2017.0345
Odegaard B, Wozny DR, Shams L. 2015. Biases in Visual, Auditory, and Audiovisual Perception of Space. *PLoS Comput Biol* **11**:e1004649.
Peters MAK, Balzer J, Shams L. 2015. Smaller= denser, and the brain knows it: natural statistics of object density shape weight expectations. *PLoS One*.
Peters MAK, Fesi J, Amendi N, Knotts JD, Lau H, Ro T. 2017a. Transcranial magnetic stimulation to visual cortex induces suboptimal introspection. *Cortex* **93**:119–132.
Peters MAK, Ma WJ, Shams L. 2016. The Size-Weight Illusion is not anti-Bayesian after all: a unifying Bayesian account. *PeerJ* **4**:e2124–e2124.
Peters MAK, Thesen T, Ko YD, Maniscalco B, Carlson C, Davidson M, Doyle W, Kuzniecky R, Devinsky O, Halgren E, Lau H. 2017b. Perceptual confidence neglects decision-incongruent evidence in the brain. *Nature Human Behaviour*.
Peters MAK, Zhang L-Q, Shams L. 2018. The material-weight illusion is a Bayes-optimal percept under competing density priors. *PeerJ* **6**:e5760.
Pouget A, Drugowitsch J, Kepecs A. 2016. Confidence and certainty: distinct probabilistic quantities for different goals. *Nat Neurosci* **19**:366–374.
Provis JM, Dubis AM, Maddess T, Carroll J. 2013. Adaptation of the central retina for high acuity vision: cones, the fovea and the avascular zone. *Prog Retin Eye Res* **35**:63–81.
Rahnev D, Bahdo L, de Lange FP, Lau H. 2012a. Prestimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception. *J Neurophysiol* **108**:1529–1536.
Rahnev D, Koizumi A, McCurdy LY, D'Esposito M, Lau H. 2015. Confidence Leak in Perceptual Decision Making. *Psychol Sci* **26**:1664–1680.
Rahnev D, Maniscalco B, Graves T, Huang E, de Lange FP, Lau H. 2011. Attention induces conservative subjective biases in visual perception. *Nat Neurosci* **14**:1513–1515.
Rahnev D, Maniscalco B, Luber B, Lau H, Lisanby SH. 2012b. Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. *J Neurophysiol* **107**:1556–1563.
Ratcliff R, Voskuilen C, McKoon G. 2018. Internal and external sources of variability in perceptual decision-making. *Psychol Rev* **125**:33–46.
Rosenholtz R. 2016. Capabilities and Limitations of Peripheral Vision. *Annu Rev Vis Sci*

2:437–457.

Rosenholtz R, Huang J, Ehinger KA. 2012a. Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. *Front Psychol* **3**:13.

Rosenholtz R, Huang J, Raj A, Balas BJ, Ilie L. 2012b. A summary statistic representation in peripheral vision explains visual search. *J Vis* **12**. doi:10.1167/12.4.14

Rounis E, Maniscalco B, Rothwell JC, Passingham RE, Lau H. 2010. Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn Neurosci* **1**:165–175.

Samad M, Chung AJ, Shams L. 2015. Perception of Body Ownership Is Driven by Bayesian Sensory Inference. *PLoS One* **10**:e0117178–e0117178.

Samaha J, Barrett JJ, Sheldon AD, LaRocque JJ, Postle BR. 2016. Dissociating Perceptual Confidence from Discrimination Accuracy Reveals No Influence of Metacognitive Awareness on Working Memory. *Front Psychol* **7**:851.

Samaha J, Iemi L, Postle BR. 2017. Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Conscious Cogn*. doi:10.1016/j.concog.2017.02.005

Seriès P, Seitz AR. 2013. Learning what to expect (in visual perception). *Front Hum Neurosci* **7**:1–14.

Shams L, Beierholm U. 2010. Causal inference in perception. *Trends Cogn Sci* **14**:425–432.

Shams L, Kamitani Y, Shimojo S. 2000. Illusions: What you see is what you hear. *Nature* **408**:788–788.

Solovey G, Graney GG, Lau H. 2014. A decisional account of subjective inflation of visual perception at the periphery. *Atten Percept Psychophys* **77**:258–271.

Stocker AA, Simoncelli EP. 2006. Noise characteristics and prior expectations in human visual speed perception. *Nat Neurosci* **9**:578–585.

Teufel C, Subramaniam N, Fletcher PC. 2013. The role of priors in Bayesian models of perception. *Front Comput Neurosci* **7**:25.

Valton V, Karvelis P, Richards KL, Seitz AR, Lawrie SM, Seriès P. 2019. Acquisition of visual priors and induced hallucinations in chronic schizophrenia. *Brain* **142**:2523–2537.

Vilidaite G, Baker DH. 2017. Individual differences in internal noise are consistent across two measurement techniques. *Vision Res* **141**:30–39.

Wei K, Körding KP. 2011. Causal Inference in Sensorimotor Learning. *Sensory Cue Integration* 30–30.

Weiss Y, Simoncelli EP, Adelson EH. 2002. Motion illusions as optimal percepts. *Nat Neurosci* **5**:598–604.

Wozny DR, Beierholm U, Shams L. 2010. Probability Matching as a Computational Strategy Used in Perception. *PLoS Comput Biol* **6**:e1000871–e1000871.

Wozny DR, Beierholm U, Shams L. 2008. Human trimodal perception follows optimal statistical inference. *J Vis* **8**:24:1–11.

Wozny DR, Shams L. 2011. Recalibration of Auditory Space following Milliseconds of Cross-Modal Discrepancy. *Journal of Neuroscience* **31**:4607–4612.

Yuille AL, Bülthoff HH. 1996. Bayesian decision theory and psychophysics In: Knill DC, Richards W, editors. New York: Cambridge University Press. pp. 123–161.

Zylberberg A, Barttfeld P, Sigman M. 2012. The construction of confidence in a perceptual decision. *Front Integr Neurosci* **6**:79–79.

Zylberberg A, Fetsch CR, Shadlen MN. 2016. The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *Elife* **5**. doi:10.7554/eLife.17688

Zylberberg A, Roelfsema PR, Sigman M. 2014. Variance misperception explains illusions of confidence in simple perceptual decisions. *Conscious Cogn* **27C**:246–253.

G. FIGURES & FIGURE LEGENDS

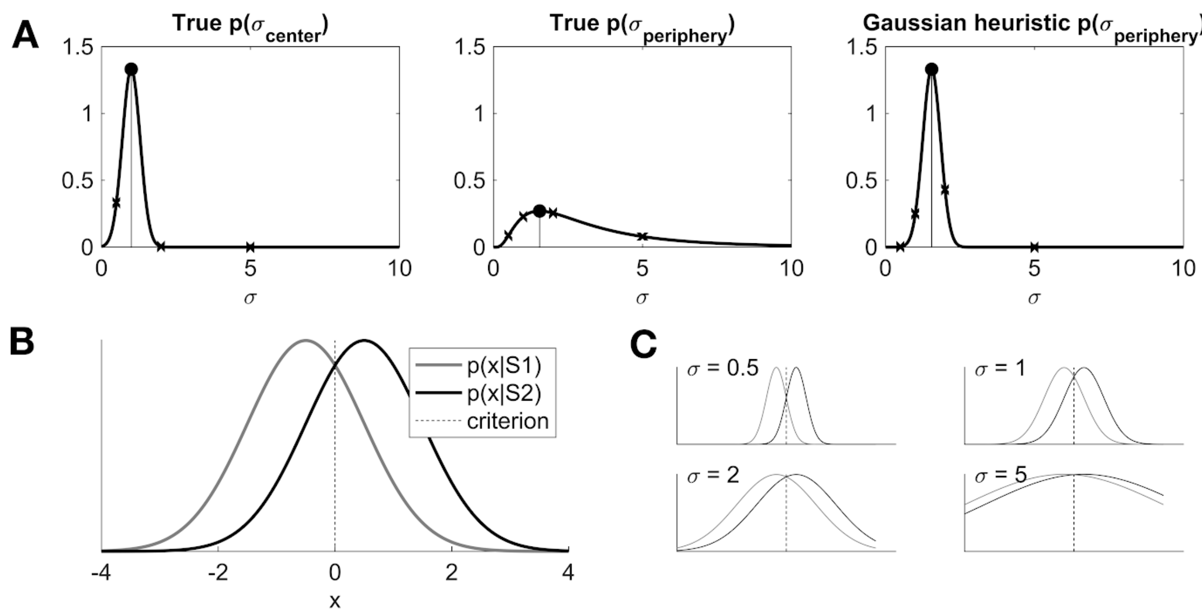

Figure 1. Hypothesized empirical prior distributions of noise in central and peripheral vision, and how they govern Type 1 decision-making in a Bayesian decision theory framework. (A) The empirical distribution of noise in central vision (left panel) is assumed to be Gaussian with relatively small mean and variance. The empirical distribution of noise in the visual periphery (middle panel) is assumed to be positively skewed. A heuristic metacognitive observer (e.g. model H3: hierarchical 'Gaussian assumption' heuristic observer; see Methods) may correctly represent the location and scale of the peripheral noise distribution, but misrepresent its skewed shape as symmetric (right panel). Small x's refer to possible noise levels sampled from these distributions, with consequences on the decision framework noted in (C).) (B) The noise in the central or peripheral visual field dictates the behavior of the Type 1 decision system, here represented in a signal detection or Bayesian decision theoretic framework. (C) As noise changes due to environmental conditions or location in the visual field (small x's in the panels in (A)), the Type 1 decision system will become more or less precise, i.e. more or less sensitive to the signal in the environment.

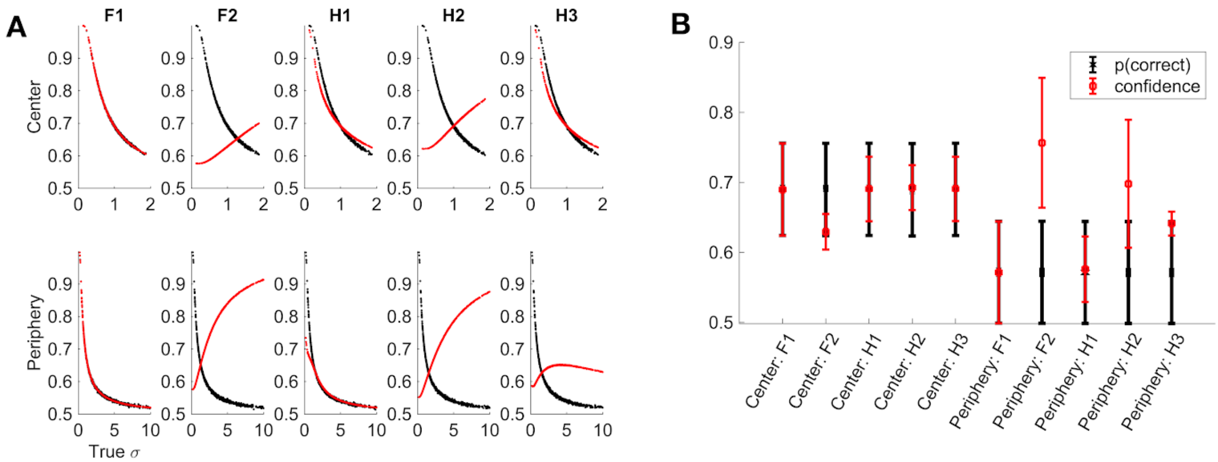

Figure 2. Predicted performance and confidence behavior from simulations of all five model observers. (A) Predicted performance (p(correct)) and confidence in central and peripheral vision for each model across a range of noise levels. (B) Median performance and confidence for each model observer in central and peripheral vision. Models F1 (flat Bayesian ideal observer) and H1 (hierarchical Bayesian ideal observer) display close match between performance (p(correct)) and confidence in both central and peripheral vision, leading to neither under- nor overconfidence on average. In contrast, models F2 (flat fixed criterion heuristic observer) and H2 (hierarchical 'mode prior' heuristic observer) display patterns of both under- and overconfidence in central and peripheral vision. Finally, while H3 (hierarchical 'Gaussian assumption' heuristic observer) correctly estimates confidence in central vision due to knowledge of the shape of the empirical prior distribution of noise, due to its assumption of symmetry in the visual periphery noise distribution, it systematically under-estimates noise leading to average over-confidence for the visual periphery only. See Table 1 and Methods for details of models.

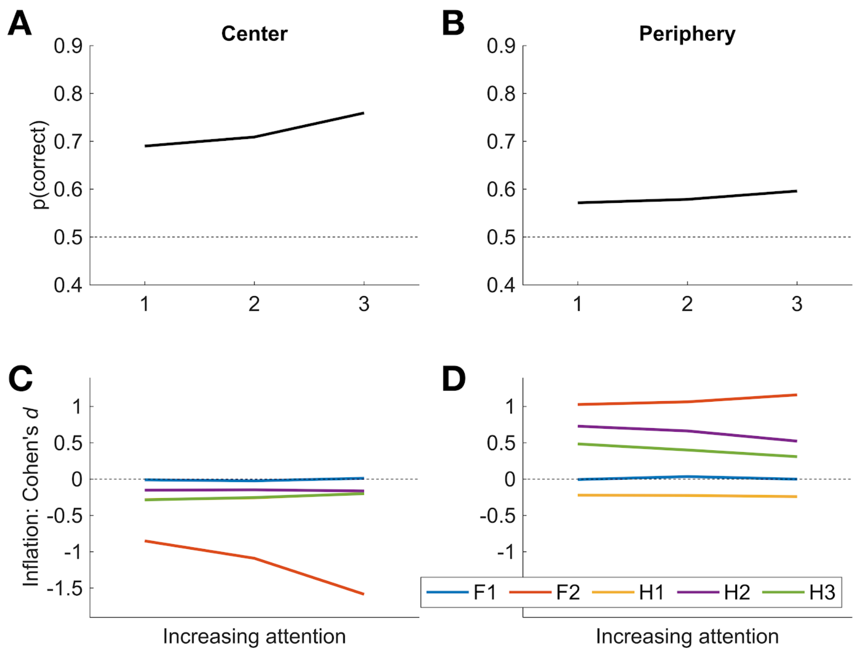

Figure 3. Results of simulated attentional manipulations on the effect size (Cohen's *d*) of subjective inflation (overconfidence). Increasing attention is modeled as increasing precision in sensory estimates and increasing metacognitive noise estimation precision (see Methods). In central vision, increasing attentional allocation leads as expected to improved Type 1 performance (A), but generally flat or slightly increasing effect size for inflation (C) with the exception of model F2, which shows stronger and stronger *under*confidence as attention increases. In contrast, in the periphery, despite attentional allocation increasing performance as expected (B), most models display flat or slightly increasing effect size for peripheral inflation (D) with the exception of models H2 and H3, the two heuristic hierarchical models.

H. TABLES

Model name	**Brief description**
F1: Flat Bayesian ideal observer	Sets a single estimate for the noise in its own system across both central and peripheral vision based on optimal 'knowledge' of both, without any explicit metacognitive noise estimation process.
F2: Flat fixed criterion heuristic observer	Has optimal knowledge of the empirical distributions of noise in central versus peripheral vision, including the positive skew of peripheral noise. Uses Bayes' rule to optimally estimate the actual noise experienced at a given moment based on these prior expectations and the likelihood experienced at a given moment, separately in central and peripheral vision.
H1: Hierarchical Bayesian ideal observer	Has optimal knowledge of the empirical distributions of noise in central versus peripheral vision, including the positive skew of peripheral noise. Uses Bayes' rule to optimally estimate the actual noise experienced at a given moment based on these prior expectations and the likelihood experienced at a given moment, separately in central and peripheral vision.
H2: Hierarchical 'mode prior' heuristic observer	Uses an estimate of the most likely noise experienced at a given moment (i.e., the mode of the empirical prior distribution), separately for central and peripheral vision.
H3: Hierarchical 'Gaussian assumption' heuristic observer	Uses Bayes' rule to estimate the noise experienced at a given moment (similar to H1), but relies on a Gaussian estimate of the empirical distribution of noise in peripheral vision rather than the true positively-skewed empirical prior -- i.e., a distribution of noise at the correlect location, but lacking skew.

Table 1. Introduction to the *Flat* and *Hierarchical* models tested.

		neutral	attention+	attention++
a_1		1	0.9	0.8
a_2		1	0.8	0.6
Model		**Cohen's *d***		
F1	**center**	-0.0104	-0.0236	0.0120
	periphery	-0.0054	0.0351	-0.0005
F2	**center**	-0.8510	-1.0903	-1.5854
	periphery	1.0273	1.0637	1.1605
H1	**center**	-0.2832	-0.2572	-0.1974
	periphery	-0.2212	-0.2248	-0.2412
H2	**center**	-0.1522	-0.1474	-0.1632
	periphery	0.7287	0.6634	0.5230
H3	**center**	-0.2857	-0.2536	-0.2001
	periphery	0.4846	0.4007	0.3090

Table 2. Cohen's *d* measures of effect size for inflation, i.e. the difference between confidence ($p(choice|x)$) and performance (p(correct)), as a function of increasing attentional allocation (neutral, attention+, and attention++; see Methods). Positive Cohen's *d* indicates inflation (confidence overestimates performance), while negative Cohen's *d* indicates deflation (confidence underestimates performance). Empirical evidence shows that peripheral inflation is stronger under inattention, so increasing attention ought to reduce inflation strength. Only H2 and H3 correctly predicted that peripheral inflation ought to decrease with increasing attention to the periphery.

Parameter	Description	Used by	Value(s)
μ_{signal}	mean of signal distribution	all models	0.5
μ_{noise}	mean of noise distribution	all models	-0.5
M_{center}	mean of central vision noise prior distribution	H1, H2, H3	1
$M_{periphery}$	mean of peripheral vision noise prior distribution	H1, H2, H3	1
Σ_{center}	standard deviation of central vision noise prior distribution	H1, H2, H3	0.3
$\Sigma_{periphery}$	standard deviation of peripheral noise prior distribution	H1, H2, H3	0.75
ς_{center}	standard deviation of central noise likelihood distribution	H1, H3	0.15
$\varsigma_{periphery}$	standard deviation of peripheral noise likelihood distribution	H1, H3	0.5
a_1	scaling factor for the mean of the noise prior distributions	all models	[1,0.9,0.8]
a_2	scaling factor for the standard deviation of the noise likelihood distributions	H1, H3	[1,0.8,0.6]

Table 3. Parameters used in all simulations.