# Priors for natural image statistics inform confidence in perceptual decisions

Rebecca K West[1]*, Emily J A-Izzeddin[2], William J Harrison[1,2]

[1] School of Psychology, University of Queensland, Australia

[2] Queensland Brain Institute, University of Queensland, Australia

* Corresponding author: rebecca.west1@uq.net.au

Competing Interests: The authors have no competing interests to declare.

Acknowledgements: We thank Cheyanne Gu for assistance with data collection. This work was funded by an ARC DECRA to WJH (DE190100136).

## Abstract

Decision confidence plays a critical role in humans' ability to make adaptive decisions in a noisy perceptual world. Despite its importance, there is currently little consensus about the computations underlying confidence judgements in perceptual decisions. One leading theory suggests that confidence is computed following the rules of Bayesian inference. Accordingly, the goal of the current study was to investigate a fundamental assumption of Bayesian models: the use of prior knowledge in subjective confidence. Rather than requiring participants to internalise the parameters of an arbitrary prior distribution, we capitalised on the existing probability distributions of features in natural scenes, which are known to play a critical role in guiding perception. Participants reported the subjective upright of naturalistic image target patches, and then reported their confidence in their orientation responses. We used computational modelling to relate the statistics of the targets to participants' responses, confirming that participants used the prior probability distribution of features in natural scenes to judge subjective upright. Critically, our results reveal that participants also used natural image priors to inform their confidence judgements. Our findings provide important evidence supporting a Bayesian characterisation of confidence and highlight the influence of environmental priors on confidence.

## Introduction

Humans make hundreds of decisions about their perceptual world every day and these decisions are often accompanied by a sense of confidence. Confidence can be defined as the belief that a choice or proposition is correct based on the available evidence (Pouget et al., 2016). In a perceptual world where the 'available evidence' is often corrupted with noise, confidence judgements play a crucial role in peoples' ability to make adaptive decisions in the absence of explicit feedback. For example, confidence has been shown to inform subsequent decisions (van den Berg, Zylberberg, et al., 2016) and to guide further information seeking before committing to a choice (Desender et al., 2018). Despite being a salient property of human decision-making, there is currently little consensus about how humans compute their confidence. A prominent theory suggests that confidence is computed following the rules of Bayesian inference (Adler & Ma, 2018; Aitchison et al., 2015; Bertana et al., 2021; Hangya et al., 2016; Li & Ma, 2020; Lisi et al., 2021; Locke et al., 2022; Navajas et al., 2017; Sanders et al., 2016). The goal of the current study, therefore, was to investigate a fundamental assumption of Bayesian models of confidence: the use of prior knowledge in the computation of confidence.

According to Bayesian decision theory, observers combine knowledge about the statistical structure of the world (the "prior") with the present sensory input (the "likelihood") to compute a posterior probability distribution over possible states of a stimulus (Drugowitsch et al., 2014b; Hangya et al., 2016; Kepecs and Mainen, 2012; Meyniel et al., 2015; Pouget et al., 2016). The posterior probability distribution, thus, represents every possible observation and its inferred probability. To form a single representation or choice, an ideal Bayesian observer chooses the most likely observation from the posterior probability distribution. According to Bayesian models, confidence can also be derived from this posterior probability distribution and can be defined as the degree of certainty (or probability) associated with the representation that the observer has chosen (or intends to choose; Ma & Jazayeri, 2014).

In contrast to Bayesian models, several alternative models of confidence have been proposed. These models posit that confidence reports are better explained by suboptimal computations, suggesting that confidence is based on learned associations with certain *heuristic* cues in the stimulus. These models have received some empirical support with a range of studies linking confidence to heuristic cues such as external noise (Bertana et al., 2021; Boldt et al., 2017; Spence et al., 2016), response time (Faivre et al., 2018; Patel et al., 2012; van den Berg, Anandalingam, et al., 2016) and task-difficulty variables (Mole et al., 2018).

Despite substantial research interest and myriad explanatory models, the computations underlying the generation of decision confidence remain unclear. Empirical studies investigating the plausibility of Bayesian and non-Bayesian models of confidence have shown mixed results, with some studies finding support for Bayesian models (Aitchison et al., 2015; Li & Ma, 2020; Navajas et al., 2017; Sanders et al., 2016) and others for non-Bayesian models (Adler & Ma, 2018; Aitchison et al., 2015; Bertana et al., 2021; Denison et al., 2018; Lisi et al., 2021; Locke et al., 2022; West et al., 2022). Furthermore, one of the major limitations of existing research is that many previous studies have exposed participants to stimuli drawn from arbitrarily pre-specified prior distributions, and compared their behaviour to that of a Bayesian optimal observer with full knowledge of that distribution (Denison et al., 2018; Li & Ma, 2020; Locke et al., 2022; Qamar et al., 2013; West et al., 2022). Such an approach is limited by the ability of humans to internalise the statistics of the prior distribution within the time frame of the experiment (Girshick et al., 2011). If participants are unable to internalise the statistics of the prior distribution within the time-limited context of the experiment, it is unlikely that their confidence judgements will match those of a Bayesian observer, even if priors are otherwise used to inform confidence.

To better understand the computations underlying confidence judgements and distinguish among candidate models, in this study we seek to address the extent to which confidence is informed by a prior distribution *without requiring participants to learn the parameters of that distribution over the short-term*. Instead, following A-Izzeddin, Mattingley and Harrison (2022), we took advantage of the distributions of low-level features present in naturalistic stimuli to define a prior distribution, the statistics of which have previously been shown to bias humans' perceptual decisions (Appelle, 1972; Berkley et al., 1975; Dakin, 2001; de Gardelle et al., 2010; Girshick et al., 2011).

A recent study introduced a novel psychophysical paradigm which shows a clear link between environmental statistics and perceptual inferences. A-Izzeddin et al. (2022) presented participants with randomly oriented target images of outdoor scenes and participants were required to infer their subjective “upright” orientation. The targets were designed so that they were windowed within a small aperture of the original image, meaning that the targets themselves contained very little high-level structure. Participants, therefore, had to rely on a strategy which depended on the low-level image features in the targets to judge “upright”. The authors sought to determine if participants had an internal representation of the average distribution of low-level features in the environment, a *prior*, which they used to guide their judgements. For example, as shown in **Figure 1**, horizontal orientation features, and, to a lesser extent, vertical orientation features are over-represented in natural images. A-Izzeddin and colleagues found that participants’ distribution of orientation responses was well approximated by the frequency distributions of orientations in natural images. The authors then used a model observer to show that participants’ responses were explained by a process in which they oriented the targets so that the low-level features in the targets, such as orientation and phase, best approximated the *prior* distribution for these features in the environment. This study, therefore, reveals that *humans have an existing, internal model of a prior probability distribution of low-level image statistics* which they use to inform their perceptual inferences about the world (A-Izzeddin et al., 2022; see also Girshick et al., 2011). Hence, in the present study, we used the same paradigm to investigate confidence because it allowed us to determine the extent to which participants use an internal prior distribution to compute their confidence without requiring them to explicitly learn the parameters of that distribution.

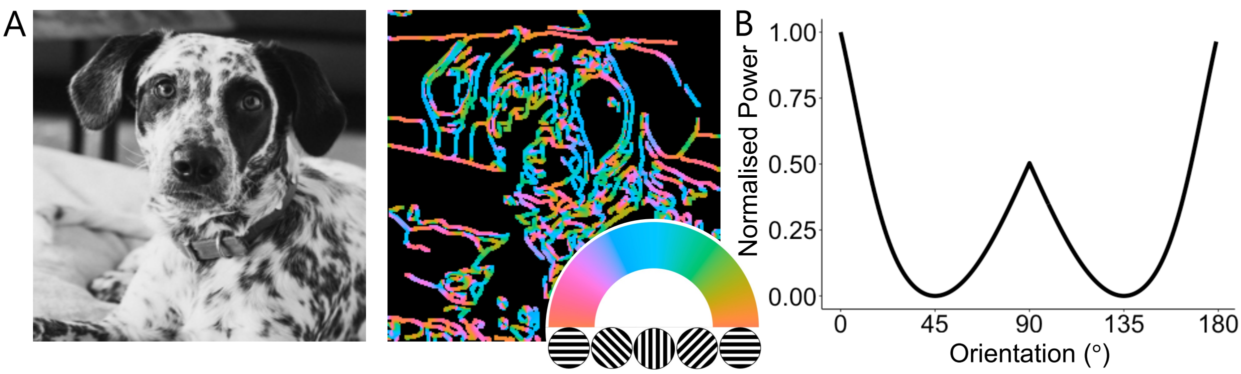


**Figure 1. Natural image statistics. (A)** Orientations of edges in digital photos have systematic biases. Original photo taken by Rafael Forseck and used under the Unsplash Licence. **(B)** Idealised distribution of orientations across many natural images (Hansen & Essock, 2004; Harrison et al., 2023).

In the present study, participants performed a perceptual task in which they made an orientation judgement about a target image, the *first-order* decision, and then reported their level of confidence in that orientation response, the *second-order* decision (see **Figure 2A**). To investigate whether participants used a prior for natural image statistics when making perceptual and confidence judgements, we used image processing techniques in combination with computational modelling. To preview our results, we confirmed that participants adopted a strategy that aligned the distribution of oriented features and phase with a prior probability model of the distribution of those features in the natural world (A-Izzeddin et al., 2022). Importantly, we also found that participants used a similar prior probability distribution to inform their confidence judgements, even without explicit instruction about the prior. Overall, our findings support a Bayesian characterisation of decision confidence in which its computation depends on multiple features of the incoming sensory information and their consistency with prior probability representations.

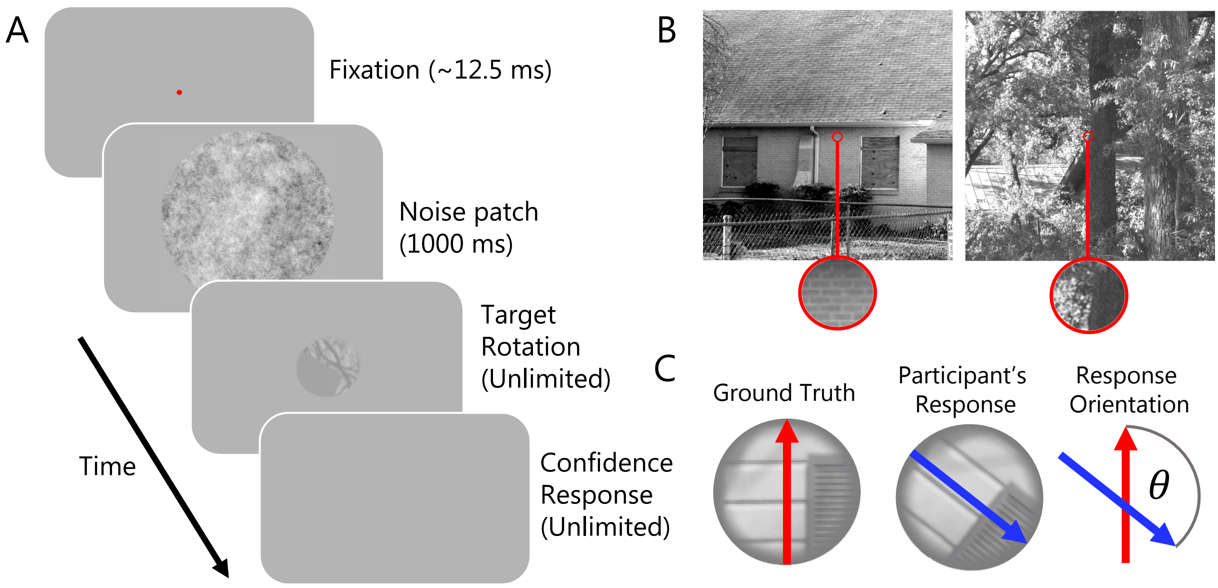


**Figure 2. Experimental paradigm. (A)** Schematic showing trial structure. Participants saw a fixation point, followed by a noise patch and then a target was presented at a random orientation. Participants used the mouse to rotate the target to appear upright and then made a confidence judgement (either "low confidence" or "high confidence") in their chosen orientation response. **(B)** Targets, such as the examples outlined in red, were extracted from high quality photos of natural scenes (Burge & Geisler, 2011). **(C)** To quantify perceptual performance, we computed the angular difference between the objective upright orientation of the target and the participants' chosen orientation.

## Results

To better understand the computations underlying decision confidence, we investigated whether confidence judgements are informed by a prior probability distribution. Rather than requiring participants to learn the statistics of an arbitrary prior distribution, we used a distribution that has been previously shown to affect participants' perceptual inferences: the distribution of orientation energy and phase in natural scenes. To validate this approach, in our first analysis, we confirmed a strong influence of the prior on participants' perceptual judgements as recently reported (A-Izzeddin et al., 2022).

### Perceptual Judgements

To quantify perceptual performance, we computed the angular difference between the objective upright orientation of the target and the participants' chosen orientation, referred to as the *response orientation* (see **Figure 2C**). Response orientation was measured in degrees ranging from -180° to 180°. The black line in **Figure 3A** shows the frequency distribution of response orientations. If participants ($N$ = 21) were not able to infer the objective upright orientation of the image using the low-level features in the target, the distribution of response orientations would be distributed uniformly across the range -180° to 180°. By contrast, **Figure 3A** shows that the most frequent response orientation was 0°, suggesting the modal response was highly accurate, and there are clear peaks at the cardinal orientations (±90° and ±180°), where targets were either inverted relative to their true orientation or offset by 90°. The presence of peaks aligned to cardinal orientations in **Figure 3A** is generally consistent with observers aligning edges within the target to the most common orientations found in nature. Note, however, that observers made ±180° inversions less frequently than 0° reports, which requires more than simply aligning the target energy to an orientation prior (in which case there would be an equal proportion of 0° and ±180° responses). To further investigate how participants made their responses, we used a model observer, described below.

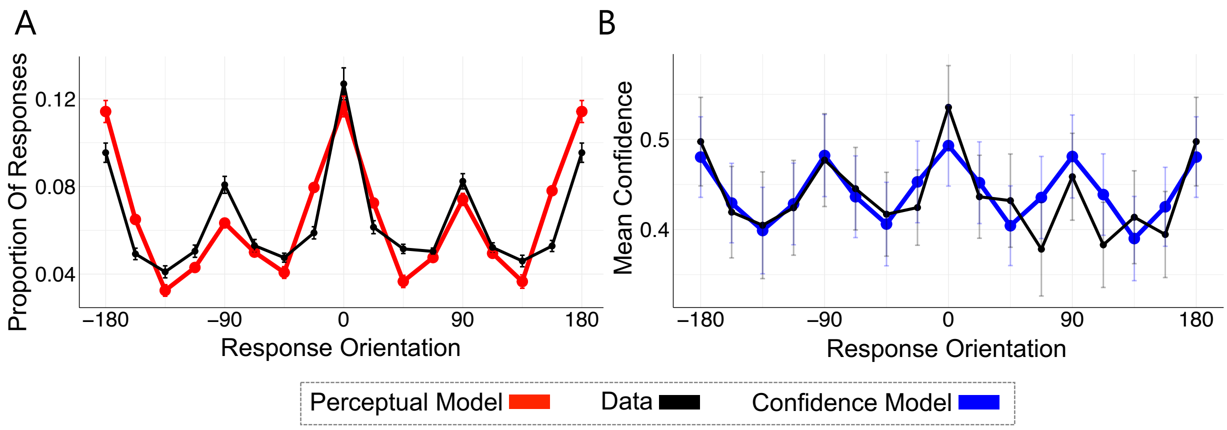


**Figure 3. Perceptual and Confidence Data and Models (A)** The distribution of observers' binned response orientations is shown in black. The output of the perceptual model is shown in red, and provides a good fit to the empirical data. **(B)** Mean confidence across binned response orientations shown in black. The output of the confidence model is shown in blue, and provides a good fit to the empirical data. Bin size = 22.5°. Error bars show ± 1 *SEM*.

### ***Perceptual Model***

We modelled participants' orientation judgements with a "pretty good observer model" developed by A-Izzeddin and colleagues (2022; see **Methods** for detailed model description). Whereas an *ideal* observer model exploits all possible sources of information to make a decision, the pretty good observer model was designed to use only a subset of image statistics to make a decision. The model included two stages. In the first stage, the model rotated the target to best match the distribution of orientation energy in the target to the prior distribution (see **Figure 4**). In the second stage, the model used a broadscale filter to estimate the phase of the horizon within the target, determining if the target needed to be a rotated a further 180°. As described in more detail by A-Izzeddin et al., the motivation of this second stage was to approximate a light-from-above prior (Adams, 2007; Brewster, 1826; Metzger, 1936; Murray, 2013; Ramachandran, 1988). This procedure produced a pattern of modelled response orientations for the 500 targets shown to each participant. To account for deviations from the model observer across participants, we fit a single noise parameter for each participant (see **Methods – Perceptual Model**).

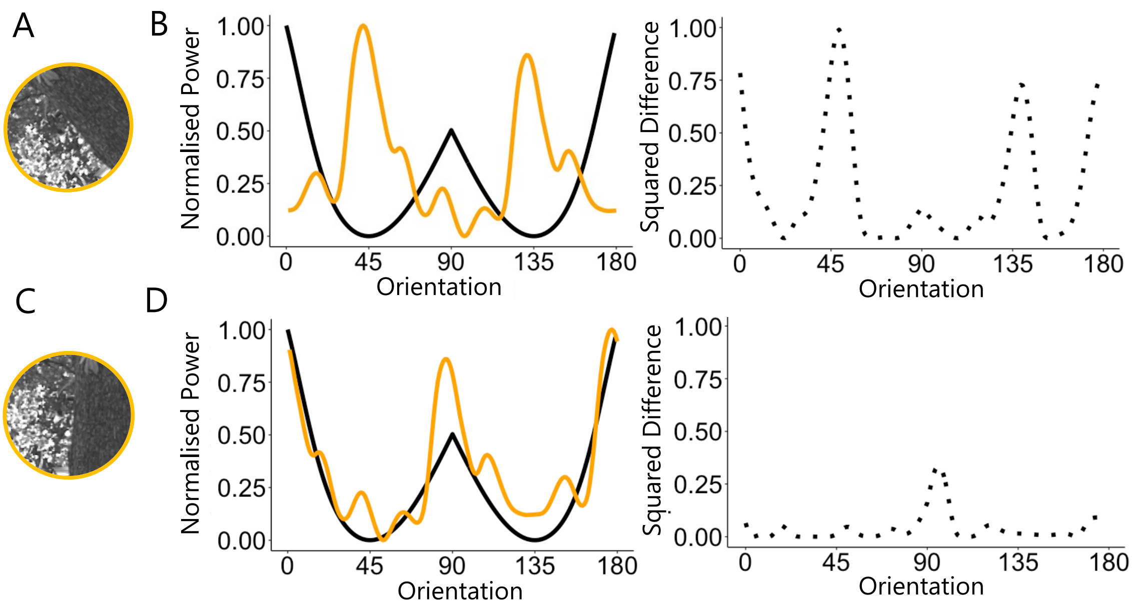


**Figure 4. Comparison of target distribution of orientation energy and prior distribution. (A)** An example target displayed at an arbitrary orientation. **(B)** Orange distribution is orientation energy for the target shown in **(A)** compared with the prior distribution of orientation energy (black). (**C**) The same target as in (**A**),but rotated to its objective upright. **(D)** Orange distribution is orientation energy for target shown in **(C)**, compared with the prior distribution of orientation energy (black)**.** Right panels of **(B)** and **(D)** show squared difference in normalised power between the target and the prior for each orientation bin. The perceptual model rotates the target to minimise the sum of the squared differences across all orientation bins.

As shown in **Figure 3A**, the perceptual model provided a very good approximation of participants' responses. The close correspondence between participants' responses and model responses suggests that participants judge the orientation of the target by matching a relatively simple set of low-level image statistics, namely orientation energy and phase, to an internal representation of the distribution of these statistics in the natural world. Thus, we confirmed that participants use prior knowledge of natural image statistics to make novel perceptual inferences about ambiguous stimuli (A-Izzeddin et al., 2022).

**Confidence Judgements**

If participants' confidence was informed by the same prior used to make their perceptual orientation inferences, we hypothesised that their confidence judgements should characterise how well they had matched the low-level statistics in the target with the prior. In contrast, if participants did not use the prior to inform their confidence, their confidence judgements would depend on other heuristic cues, not related to the prior. We tested these predictions with a second model, the confidence model, described below. We first provide an overview of the average distribution of confidence responses across response orientations and then outline the assumptions and performance of the confidence model.

The black data in **Figure 3B** shows mean confidence as a function of response orientations. Participants were most confident in their most accurate orientation responses (responses in the 0° bin), suggesting they had some metacognitive awareness of the match of their perceptual judgements to the upright image. There are also clear peaks in confidence at the cardinals (±90° and ±180°), demonstrating that participants' confidence responses show the same cardinal biases as the perceptual judgements. Our perceptual model revealed that participants use priors for the distribution of orientation energy and phase in natural scenes to inform their orientation judgements. We therefore wanted to develop a similar statistical model to determine if participants' confidence judgements were informed by the same prior.

***Confidence Model***

For the confidence model, we developed several measures that summarised the degree of overlap in orientation energy and phase in the target and the prior. In addition to these prior-related measures, we also considered the possibility that confidence depends on other stimulus cues not related to the prior, such as the contrast of the target patch or response time of the first-order decision (Bertana et al., 2021; Boldt et al., 2017; Faivre et al., 2018; Mole et al., 2018; Patel et al., 2012; Spence et al., 2016; van den Berg, Anandalingam, et al., 2016). To relate these measures to confidence and investigate their relative importance, we used a generalised linear mixed model framework in the form of a modified logistic regression. This allowed us to use a set of weighted stimulus variables to predict binary confidence responses. We describe the assumptions of each predictor in the model and the fixed effect of that predictor below. See **Methods – Confidence Model** for further model description. See **Supplementary Table 1 and Supplementary Figure 2** for model parameters and predictions. See **Supplementary Figure 5** for the correlation between fixed effects.

**Orientation Energy.** We tested two different assumptions about how participants may quantify the degree of overlap between the distribution of orientation energy in the target and the prior. These assumptions were quantified in what we refer to as the *prior mismatch predictor* and the cardinal and oblique orientation *energy predictors*. For the prior mismatch predictor, we assumed that participants directly compared a continuous representation of the distribution of orientation features in the stimulus and veridical knowledge of the prior distribution of orientation energy (see **Figure 5A**). The prior mismatch predictor, therefore, was a single statistic that summarised the difference between the entire distribution of orientation energy in the target at the rotational offset chosen by the observer and the prior. As shown in **Figure 6B**, however, the effect of the prior mismatch predictor on confidence was not significant ($e^{\beta} = 0.95, 95\%\ CI\ [0.88, 1.04], p = 0.261$), suggesting that participants do not appear to explicitly compare the distribution of orientation energy in the target and the prior veridically.

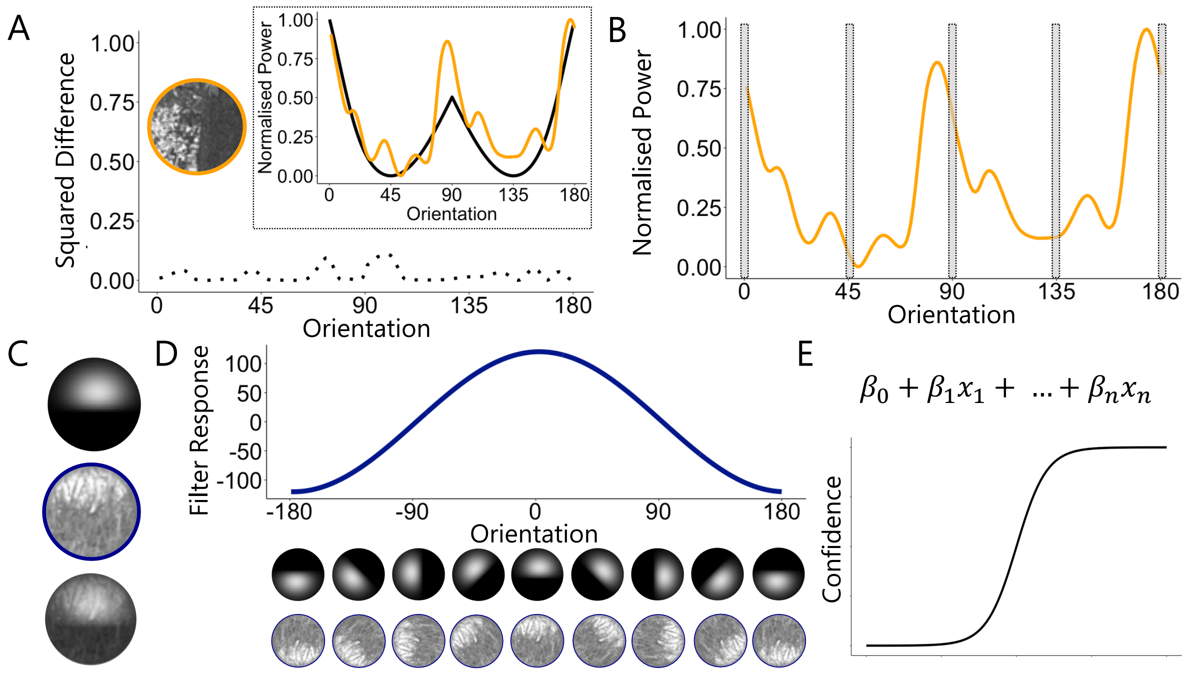


**Figure 5. Prior Predictors for Confidence Model. (A)** Example difference in orientation energy for each bin between the distribution for the target (orange distribution in inset) and the prior (black distribution in inset). The sum of these differences was used as the *prior mismatch predictor* in the confidence model. **(B)** Example distribution of orientation energy with horizontal (0° or 180°), vertical (90°) and oblique (45° or 135°) energy shaded. These values were calculated for each target at the rotational offset chosen by the observer and used to predict confidence. **(C)** Example of a broadscale filter with a positive polarity (top circle) and an example target (middle circle). The filter is applied to the target (bottom circle) to determine the direction and strength of lighting in the target. **(D)** Example response of a broadscale filter (blue distribution in plot). The broadscale filter is positioned at different rotational offsets (top row) based on the orientation of the target (bottom row). The response of the filter at the offset chosen by the observer was used as a predictor in the confidence model. **(E)** Linear predictors were passed through an inverse logistic link function to predict confidence responses.

As an alternative to computing confidence by comparing the full prior probability distribution of orientation energy with that measured from the target, we postulated that observers might instead be most confident in targets that contain strong vertical or horizontal features, based on their internal model of the over-representation of these features in natural scenes. We therefore included a second set of predictors in the confidence model which quantified the amount of vertical (where $\theta = 90°$), horizontal energy (where $\theta = 0° \; or \; 180°$) and oblique energy (where $\theta = 45° \; or \; 135°$) in the target at the rotational offset chosen by the participant (see **Figure 5B**). We found that vertical orientation energy ($e^{\beta} = 2.13, \; 95\% \; CI \; [1.47, 3.09], \; p < .001$), horizontal orientation energy ($e^{\beta} = 2.11,\ 95\%\ CI\ [1.45, 3.06],\ p < .001$) and oblique orientation energy ($e^{\beta} = 1.42,\ 95\%\ CI\ [1.12, 1.78],\ p = .003$) were all positively predictive of confidence. The effect of vertical and horizontal energy was almost twice that of oblique orientation energy (see **Figure 6B**), suggesting that participants most heavily weight the over-representation of vertical and horizontal features in natural scenes to inform their confidence.

**Phase.** Orientation energy is phase invariant, and so a participant cannot use orientation energy alone to distinguish a 0° and ±180° target. We therefore used a set of predictors in the confidence model that estimated phase (see **Methods**). We used the response of a broadscale filter positioned in the centre of the target at the rotational offset chosen by the observer as a phase, which we take to be indicative of the strength of a directional lighting signal. We used both the absolute value and a log transformed value of absolute phase in the confidence model to allow for non-linear relationships between target phase and confidence. The effects of absolute phase ($e^{\beta} = 0.89,\ 95\%\ CI\ [0.81, 0.98],\ p = .020$) and log transformed absolute phase on confidence were significant ($e^{\beta} = 1.32,\ 95\%\ CI\ [1.18, 1.49],\ p < .001$). The effect of the log transformed absolute phase on confidence suggests that the relationship between target phase and confidence was non-linear. This relationship is also positive such that the more evidence that there was a strong source of lighting in the target, the more confident the participant was in their response. Overall, these findings suggest that participants use low-level features such as the phase of the target, which can approximate the strength of a light source in the image, to inform their confidence.

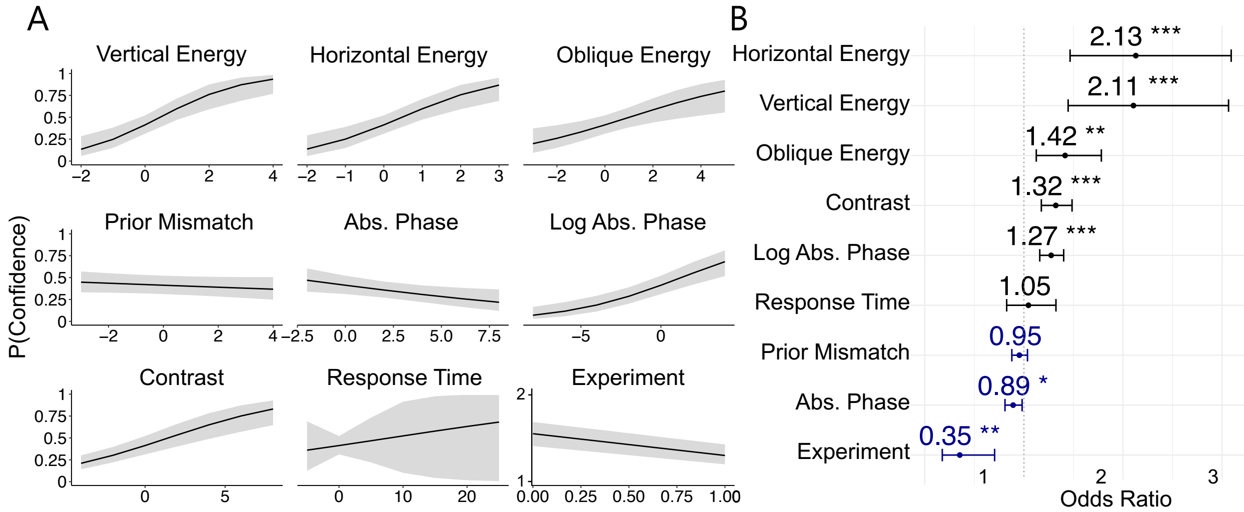


**Figure 6. Confidence Model. (A)** Marginal fixed effects of each predictor on confidence. Shaded regions show ± 1 standard error of the predictions. Note that predictors vertical energy, horizontal energy, oblique energy, prior mismatch, absolute phase, log absolute phase, contrast and response time were standardised. **(B)** Odds ratio for fixed effects. Dotted vertical line indicates odds ratio of 1. Error bars show 95% confidence intervals. Blue values indicate negative coefficients. ***$p < .001$,**$p < .01$, *$p < .05$.

**Other predictors**. If, contrary to the above effects of orientation energy and phase, participants did not use the prior to inform their confidence, we hypothesised that they might use other heuristic-like cues. We therefore included two predictors in the confidence model that were not directly related to the prior: the overall contrast of the target patch (RMS contrast) and the response time of the first-order decision.

The effect of response time on confidence was not significant ($e^{\beta} = 1.05$, $95\%\ CI\ [0.83, 1.32],\ p = .721$), suggesting that participants do not use their first-order response time as a heuristic cue for confidence. The effect of contrast on confidence, however, was significant ($e^{\beta} = 1.27,\ 95\%\ CI\ [1.16, 1.40],\ p < .001$), where increasing target contrast was predictive of increasing confidence. Considering the strong effect of oriented contrast energy on confidence reports, it is not particularly surprising that a measure of isotropic orientation was also a significant predictor. Indeed, RMS contrast was moderately correlated with vertical and horizontal energy (see **Supplementary Figure 5**). However, this finding could nonetheless also suggest that, over and above the effect of priors, participants use other stimulus features to compute their confidence.

To visualise how well these measures accounted for the data more generally, we generated predictions for a model using the statistically significant predictors only ($p < 0.05$ in **Figure 6**; horizontal orientation energy, vertical orientation energy, oblique orientation energy, absolute phase, log absolute phase, contrast and experiment; see **Supplementary Table 2**). The predicted confidence from this model is shown in **Figure 3B** (blue distribution). The model describes the empirical confidence data very well (black distribution; **Figure 3B**), capturing the major peaks in confidence at the cardinal orientations. To provide a reference point for model performance, we fit an intercept only model with an intercept term for each participant (see **Supplementary Figure 3** and **Supplementary Table 3**). We found substantially poorer performance for the intercept model ($AIC_{intercept}$ = 12657.19, $BIC_{intercept}$ = 12678.97, $AIC_{conf}$ = 12099.96, $BIC_{conf}$ = 12361.29; see **Supplementary Figure 4**). This finding suggests that the confidence model, with the predictors that quantify the low-level features in the targets and their degree of overlap with the prior, was more useful for explaining variations in confidence than a constant value for confidence for each participant.

Despite capturing many aspects of participants' confidence reports, the model is incomplete. As shown in **Figure 3B**, the confidence model underpredicts confidence for correctly oriented targets (those with response orientations of 0°), suggesting that observers may have had access to other features in the targets, not captured by the model, which led them to confidently infer the correct upright orientation of the target. The model also appears to overpredict confidence for response orientations between ~65-165°. We expect that this asymmetry in the confidence data, where the pattern of responding is different for response orientations of ~65-165° and ~-65-165°, is a result of non-stimulus-specific noise that is not captured by the model. Despite these minor limitations, a small set of environmental statistics provides a reasonable basis from which to understand confidence computations.

## Discussion

We investigated the influence of priors on decision confidence. Observers performed a task in which they rotated a naturalistic image patch to its upright orientation, the *first order decision*, and then made confidence judgements in their orientation responses, the *second order decision*. We found that participants use internal priors about the statistical regularities of low-level features in natural scenes to make their perceptual judgements, replicating a recent investigation (A-Izzeddin et al., 2022). Importantly, we also found that participants use the same priors to inform their confidence responses, even without explicit instruction about the prior. We discuss the implication of these findings for our understanding of decision confidence below.

## Priors Affect Confidence

We found that participants use prior knowledge about the statistical regularities of orientation energy and phase in natural scenes to inform both their perceptual inferences and confidence judgements. In the following sections, we discuss the role of this information in models of confidence.

### *Orientation Energy*

Our confidence modelling results showed that the amount of vertical and horizontal orientation energy in the target was an important predictor of confidence. In fact, the effect of vertical and horizontal energy on confidence was almost twice that of oblique orientation energy. Participants were not given any instruction about which image features to use to make their judgements and were not given any explicit instruction about a prior distribution. This finding, therefore, suggests that participants appear to use implicit knowledge about the over-representation of horizontal and vertical orientation features in natural scenes to inform their confidence. This finding broadens our understanding of the influence of the statistical regularities of orientation features in natural scenes on perceptual judgements (Appelle, 1972; Berkley et al., 1975; Campbell et al., 1966; Dakin & Watt, 1997; Girshick et al., 2011; Hansen et al., 2003, 2008; Pratte et al., 2016) and demonstrates, for the first time, that these low-level perceptual priors also effect observers' confidence.

Although participants appear to use knowledge about the prior probability of certain oriented features in natural scenes to inform their confidence, we did not find evidence that participants directly compare a full prior probability distribution of orientation energy in natural scenes and the distribution of orientation energy in the target directly (the *prior mismatch* predictor). Instead, our results suggest that participants use only a subset of orientations to inform their confidence, or, alternatively, use orientations within only some spatial frequencies. This finding suggests that perhaps observers may not use the entire prior probability distribution but instead, confidence is informed by only the most predictive features of the prior. As shown in **Figure 7**, perceptually apparent cardinal structures need not be defined by peaks in energy as defined in our prior. When quantifying energy, we aggregated over all spatial frequencies, whereas human vision is known to be bandpass (Campbell & Robson, 1968). Future models could therefore weight an image's orientation energy according to the human contrast sensitivity. Furthermore, phase alignment at different spatial scales is critical to perceptually relevant edge features that can guide the sorts of perceptual and confidence decisions in our study (Rideaux et al., 2022a). We discuss the role of phase in confidence computations in the next section.

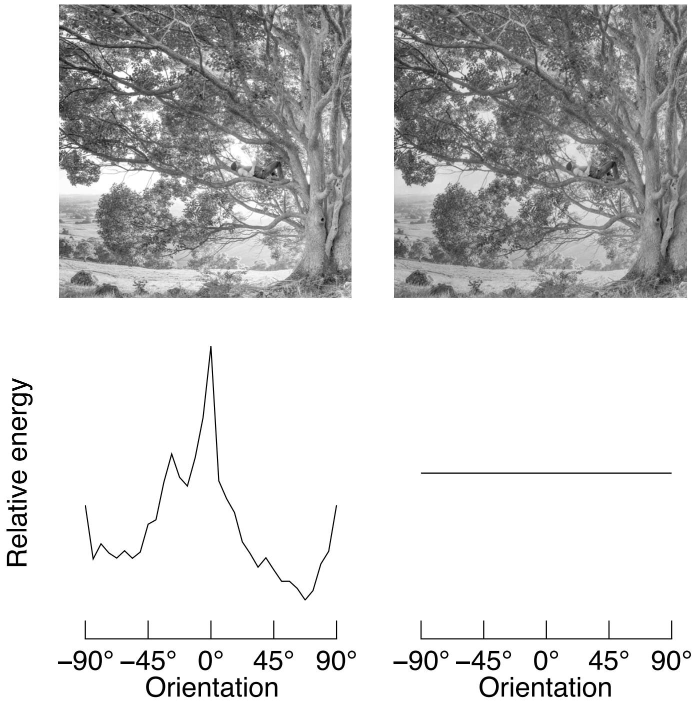


**Figure 7. Informativeness of Full Prior Distribution.** Comparing the full distribution of orientation energy to a prior may not always be functional. The image on the left shows a greyscale natural image. Its orientation energy is shown below the image, and closely matches the prior distribution shown in **Figure 1**. The image on the right is the same image, but we have "whitened" its amplitude spectrum with respect to orientation: as shown below the image, the modified image has equal energy at all orientations. Despite the large change in orientation energy, the images are perceptually similar. Because natural images are dominated by low-spatial frequency structures, energy within low spatial frequency bands were over-represented in our measure of orientation energy relative to the bandpass properties of human vision. However, this example nonetheless demonstrates that clearly oriented cardinal structures need not be defined by peaks in energy as defined in our prior per se. Instead, phase alignment at different spatial scales, and orientation energy within different spatial frequency sub-bands, are critical to perceptually relevant edge features that can guide the sorts of perceptual and confidence decisions in our study (Rideaux et al., 2022b).

***Phase***

Because orientation energy is phase-invariant, we included a set of predictors in the confidence model that estimated the phase of the horizon in the target in order to capture a proportion of responses that were inverted by 180°. We found a non-linear effect of the phase predictor on confidence such that the more evidence that there was a clear horizon in the target, the more confident participants were in their response.

The horizon in the target could be used to approximate both the direction and strength of lighting in the target. Lighting in the natural world is known to follow certain statistical regularities, mostly originating from above the horizon due to predominant sources of illumination being the sun and overhead lights. The demonstrated effect of phase on confidence is thus consistent with well documented perceptual effects that humans have priors for the direction of lighting in natural scenes (Adams, 2007; Brewster, 1826; Metzger, 1936; Murray, 2013; Ramachandran, 1988) and provides converging evidence that participants use priors based on experience with the statistical regularities of the natural world to inform their confidence.

**The Computational Basis of Confidence**

Our finding that priors are demonstratively used to inform confidence helps to resolve some of the theoretical debate about the computational basis of confidence. One of the leading theories of confidence suggests that it is computed according to the rules of Bayesian inference where humans combine a *prior* with a *likelihood* to compute a *posterior probability distribution*. Currently, however, the evidence for Bayesian accounts of confidence have been mixed with some studies finding evidence for Bayesian models (Aitchison et al., 2015; Li & Ma, 2020; Navajas et al., 2017; Sanders et al., 2016) and others for non-Bayesian models (Adler & Ma, 2018; Aitchison et al., 2015; Bertana et al., 2021; Denison et al., 2018; Lisi et al., 2021; Locke et al., 2022). One of the major limitations of existing research, however, is that many previous studies have assumed participants' effective internalisation of arbitrary prior distributions and compared their behaviour to that of a Bayesian optimal observer with full knowledge of that distribution (Denison et al., 2018; Li & Ma, 2020; Locke et al., 2022; Qamar et al., 2013; West et al., 2022). This assumption is rarely interrogated, despite the fact that if observers are unable to internalise the statistics of the prior distribution within the time-limited context of the experiment, it is unlikely their confidence judgements will match that of a Bayesian observer. This would make it difficult to find evidence of the use of priors in confidence judgments even if priors are used to inform confidence beyond the lab.

In this study, we used a task design in which we did not need to explicate a prior distribution and could instead rely on participants' existing priors. Such a task allowed us to confirm that priors do indeed influence confidence responses. Our findings are consistent with a Bayesian characterisation of confidence in which confidence depends on prior information, although further development of our models is required in order to compute a full posterior distribution as per other Bayesian formulations (Aitchison et al., 2015; Li & Ma, 2020; Navajas et al., 2017; Sanders et al., 2016; Denison et al., 2018; Locke et al., 2022; West et al., 2022). Our demonstration of the importance of priors suggests that previous research showing evidence *against* Bayesian models of confidence should be interpreted with caution. It is feasible that the use of arbitrary prior distributions in previous research meant that participants were unable to internalise the statistics of those distributions veridically and thus showed systematic deviations from prior-informed optimal behaviour. Our findings imply that future studies evaluating the computational basis of confidence could rely more on naturalistic task designs or use computational models that allow for inaccurate knowledge about the prior distribution.

**Other Stimulus Features as Cues for Confidence**

In addition to the effect of priors on confidence, we investigated if participants used other cues to inform their confidence. Inconsistent with some previous research, we did not find evidence that response time in the first order decision influenced confidence (Faivre et al., 2018; Patel et al., 2012; van den Berg, Anandalingam, et al., 2016). We did, however, find that the contrast of the target had a significant effect on confidence, where higher target contrast was associated with greater confidence. We postulate that stimulus contrast in our experimental paradigm may provide an important cue about sensory uncertainty. Lower contrast targets provide less clear and consistent visual information (Bex & Makous, 2002), and, therefore, contrast can provide a meaningful cue about the perceptual uncertainty of the source of information on which the decision is based. This is consistent with studies showing that confidence is influenced by perceived sensory uncertainty (Adler & Ma, 2018; Denison et al., 2018; Michael et al., 2015; West et al., 2022). Furthermore, it is generally consistent with theoretical work which links confidence with the hypothesis that perceptual uncertainty is encoded as the variance in firing rates across neural populations, with increased uncertainty leading to down-weighting of that evidence source in perceptual integration (Beck et al., 2008; Ma et al., 2006). Further research is required to confirm these hypotheses and understand the neural mechanisms by which certain stimulus features, like image contrast, are involved in the computation of confidence.

**Conclusions**

Our results support the idea that observers combine multiple features of incoming sensory information with prior probability representations to compute their confidence. Our experimental paradigm capitalises on statistical regularities in the structure of natural scenes such that participants relied on an existing, internal representation of prior probabilities to make their judgements. This meant that, unlike previous research, incomplete knowledge about the statistics of the prior distribution would not bias our interpretations about the computational basis of confidence. Our results provide important evidence supporting a Bayesian characterisation of confidence, highlighting the joint influence of priors and other perceptual features, such as sensory uncertainty, on confidence. Overall, our study demonstrates that prior knowledge plays an important role in both the perceptual and metacognitive decisions that humans make about the noisy, ambiguous sensory information they encounter every day.

## Methods

**Overview**. On each trial, a participant viewed a randomly oriented *target*, positioned in the centre of the display (see **Figure 2A**). The participant was informed that the target was a circular patch that had been cropped from the centre of a larger *source image,* where source images were randomly selected from a database of images of outdoor scenes (see **Figure 2B**). The participant was instructed to rotate the target to the "upright" orientation, with no additional contextual information given about the source image. For each target, the participant made their orientation judgement, the *first order decision*, and then made a confidence judgement in their chosen orientation response, the *second order decision*, reporting either "high confidence" or "low confidence".

**Participants**

In total, 21 participants ($M_{age}$ = 23.95, $SD_{age}$ = 4.65; no exclusions) completed the experiment. 10 participants completed Study 1 and 11 participants completed Study 2. All methods were the same for Study 1 and Study 2, unless indicated otherwise. As there did not appear to be any clear differences in the results of Study 1 and Study 2, we combined datasets across studies. We, therefore, report results for all 21 participants. See **Supplementary Material: Experiment 1 and Experiment 2 Comparison** for additional commentary. All participants were naïve to the purpose of the study and were reimbursed for their time ($20 per hour in cash). Ethics approval was granted by the University of Queensland Medicine, Low & Negligible Risk Ethics Sub-Committee.

**Stimuli & Apparatus**

All participants saw the same 500 digital natural images, selected randomly from a database of high-resolution photos of outdoor scenes and cropped to 1080x1080 pixel regions (Geisler & Perry, 2011; A-Izzeddin et al., 2022). Targets were circular patches cropped from the centre of the 1080x1080 images, subtending 2° of visual angle in diameter (see **Figure 2B**). All stimuli were converted to greyscale using MATLAB's rgb2gray() function. During practice, targets were selected randomly from a different set of images from the same database.

Stimuli were presented using the Psychophysics Toolbox (3.0.12; Brainard, 1997; Pelli, 1997) and a gamma correction was applied to the display, assuming gamma was 2. In Study 1, stimuli were presented on a 24-inch ASUS VG428QE monitor, 1920 x 1080-pixel resolution and a refresh rate of 100 Hz. Participants were seated in a dark room with their head positioned on a chin rest fixed at a viewing distance of 57cm. In Study 2, stimuli were presented on a 24-inch DELL P2414H monitor, 1920 x 1080-pixel resolution and a refresh rate of 60 Hz.

**Procedure**

As shown in **Figure 2A**, at the start of each trial participants were presented with a central fixation point for ~125 ms (fixation time sampled from a uniform distribution between 0 and 250 ms) followed by a black and white pink noise patch for 1000 ms (27° of visual angle in diameter). The target was then presented centrally at a random orientation and participants used the mouse to re-orient the patch to be "upright". Participants clicked the mouse to confirm their response and then used the arrow keys (left arrow = low confidence and right arrow = high confidence) to indicate their confidence in their chosen orientation. Half of the participants were instructed to use the high confidence and low confidence ratings approximately equally often (Study

1) and the other half of participants received no additional instructions about using the confidence ratings (Study 2). All participants completed 500 trials with the same targets, the order of which was randomised for each participant. Trials were split into 5 blocks of 100 trials with self-paced breaks in between blocks.

Prior to completing the experiment, participants did 40 practice trials to familiarise themselves with the task. In the first 20 practice trials, participants only rotated the targets to their “upright” orientation and, in the remaining 20 trials, participants rotated the targets and then made confidence ratings. Participants saw different target images during testing and training.

**Data Analyses**

We used statistical models and digital signal processing techniques (e.g. Harrison, 2022) to investigate whether participants use a prior for natural image statistics when making perceptual inferences. We then developed a novel statistical model to determine if participants use the same prior to inform their confidence judgments. In the following sections, we first describe the image processing methods used to derive the distribution of orientation energy and phase for the prior and each target patch. We then describe the perceptual and confidence models.

The targets were designed so that they were windowed within a small aperture of the larger source image. This meant that the targets themselves contained very little high-level structure that participants could use to unambiguously infer the upright orientation of the target (see **Supplementary Material: High-Level Image Features and “Informativeness” Data** for further discussion on this issue). Participants, therefore, had to rely on a strategy that depended on the low-level image features in the targets only (see **Supplementary Figure 8**). Based on A-Izzeddin and colleagues (2022), we expected that participants would adopt a strategy in which they chose rotational offsets for the targets which best matched the distribution of low-level features in the targets to the average distribution of these features in the environment, referred to as the *prior*. Consistent with A-Izzeddin and colleagues (2022), we focused on the use of two specific low-level features, the distribution of orientation energy and phase, and defined the prior as the average distribution of these features across thousands of images of natural scenes (see **Figure 1**). Importantly, for the experimental task, we did not provide any guidance on what features participants should use to make their decisions or give them any explicit instructions about the prior. Instead, we relied on participants' *internal* representation of the prior to do the task.

### *Orientation Energy*

**Prior.** We defined the prior distribution of orientation energy based on studies of natural images (e.g. Wei & Stocker, 2015):

$$p(\theta) \propto 2 - |sin\theta| \tag{1}$$

where $p(\theta)$ is the probability of observing orientation energy with an orientation of $\theta$ in radians. Equation **Error! Reference source not found.** assumes equal prevalence of horizontal and vertical orientations. Other studies, however, have shown that horizontal features are over-represented relative to vertical features (Hansen et al., 2003; Hansen & Essock, 2004; Harrison, 2022). We therefore modified Equation **Error! Reference source not found.** by increasing the proportion of horizontal energy according to a von Mises function which was normalised to have a peak of one:

$$p(\theta) \propto 2 - |sin\theta| + C\ exp\ (\kappa\ (\cos\theta - 1)\ ) \tag{2}$$

where *C* is the strength of the horizonal bias, and $\kappa$ is the width of the von Mises function which we set to 2.5. $p(\theta)$ was then normalised within the range 0 – 1. Small changes in $\kappa$ did not change the results. **Figure 1** shows the prior distribution of orientation energy (black distribution).

**Target Patch.** We calculated the distribution of orientation energy in the target to compare against the prior. For a given target patch, we computed orientation energy in 180 equally space orientation bands, each of which covered all spatial frequencies. These operations were performed in the frequency domain; energy was the absolute of the Fourier-transformed target patches. Orientation filters were also constructed in the frequency domain as raised cosine filters with a bandwidth of 45° (full width half height). Energy was summed within each orientation band, giving a distribution of energy across orientations. Because absolute energy fluctuates from one image to the next, the distribution of energy was normalised within the range 0 – 1, by subtracting the minimum value, and then dividing by the maximum value. **Figure 4B/D** shows example distributions of orientation energy (orange distribution) for an example target patch at two different rotational offsets in **Figure 4A/C**.

***Phase***

As described in A-Izzeddin et al. (2022), contrast energy alone is circular around ±90°, whereas observers' reports are circular around ±180°. To estimate fully circular responses, therefore, a second image statistic is required to determine whether an image needs to be inverted. We modelled this process as an estimate of the phase of the horizon.

**Prior.** Consistent with previous research, we assumed that participants had a light-from-above prior (Adams, 2007; Brewster, 1826; Metzger, 1936; Murray, 2013; Ramachandran, 1988). In other words, we assumed that the response of a broadscale filter positioned in the centre of the target should on average, be positive, consistent with light being above the horizon in naturalistic scenes.

**Target Patch.** To measure the phase of the target on a given trial, the broadscale filter was positioned in the centre of the target and at the orientation of the rotational offset of the target. The polarity of the response of the filter was used to determine lighting direction. That is, if the response of the filter was positive, light would appear to be above the horizon in the target and if the response of the filter was negative, light would appear to be below the horizon in the target (see **Figure 5C**).

**Perceptual Model**

We used the model observer developed by A-Izzeddin and colleagues (2022) to investigate whether participants use an internal prior model for low-level image features to inform their orientation judgements. This model is defined as a "pretty good observer" model because it uses a subset of image statistics to make a decision, as opposed to an "ideal observer" model which exploits all possible sources of information.

The model included two stages. In the first stage, we rotated the target to best match the distribution of orientation energy in the target with the prior. We used MATLAB's fminsearch() function to find the rotational offset that minimised the sum of squared differences in each orientation band between the target and the prior (see **Figure 4**). To avoid local minima, we fit the model with varying starting parameters and used the rotational offset at the global minimum from all fits. Because orientation contrast is phase invariant, in the second stage of the model after finding the best rotational offset, we estimated lighting direction in the patch using a broadscale filter and inverted the target so that it was consistent with a light-from-above prior, as described above. In other words, where the response of the filter was positive, the model observer would leave the target patch at the current orientation. If, on the other hand, the phase was negative, the model observer would rotate the target patch by 180°. The stronger the absolute value of the phase response, the more evidence the model observer had about the horizon.

The fitting of the perceptual model as described thus far was independent from participants' responses – we fit the targets' statistics to the prior. This meant that the model was deterministic. Because all participants saw the same targets, the model predicted the same pattern of responses for all participants. To account for deviations from the model observer across participants, we fit a single noise parameter for each participant: we added a random amount of noise, $\varepsilon$, to the model's predicted response orientation for each target. The amount of added noise was sampled from a normal distribution with a mean of 0 and a standard deviation, $\sigma$, that was estimated separately for each participant, *j*, such that:

$$N(\varepsilon; 0, \sigma_j) \tag{3}$$

where $N$ denotes the normal density function and $\varepsilon$ is the amount of noise added to the perceptual model's prediction. We estimated $\sigma$ by minimising the difference between each participant's observed distribution of response orientations and the mean predicted distribution of response orientations for that participant using ordinary least squares according to:

$$\sum_b \left(\delta_{j,b} - \rho_{j,b}\right)^2 \tag{4}$$

Where $\delta$ is the observed proportion of responses and $\rho$ is the predicted proportion of responses in bin, *b*, for participant, *j*. To calculate the predicted proportion of responses, we simulated 1000 noisy responses by drawing random samples of $\varepsilon$ according to Equation 3 for each trial and each participant and added these random samples to the model's predicted orientation response for that target. We then calculated the average proportion of responses in each orientation bin for each participant. See **Supplementary Table 4** for $\sigma$ parameter estimates.

## Confidence Model

For the associated confidence judgements, we sought to determine if participants use the same prior for the distribution of low-level features in natural scenes to inform their confidence. If participants use the same perceptual prior to inform their confidence, we hypothesised that their confidence judgements would reflect how closely the distribution of low-level features in their chosen orientation response matched the prior distribution. To test this prediction, we developed several measures that summarised the degree of overlap in orientation energy and phase in the target, as oriented by the participant, and the prior (described above). In contrast to these prior-related measures, we also considered the possibility that confidence depends on other stimulus cues not related to the prior, such as the contrast of the target patch or response time of the first-order decision. To evaluate these predictions, we used a generalised linear mixed model framework to investigate the relationship between certain features of the target and participants' confidence judgements. Importantly, we computed the measures (described in detail below) using the targets at the rotational offset chosen by the participant on each trial, reflecting the fact that confidence judgements are a second-order reflection on the first-order decision.

In the sections below, we first describe the general modelling framework (the generative model) and then describe each of the measures used to predict confidence.

### *Generative Model*

We used a generalised linear mixed model (GLMM) in the form of a modified logistic regression to relate a set of weighted predictor variables to binary confidence responses (low or high). In this framework, the sum of weighted predictor variables is passed through a logistic link function which transforms the unbounded weighted sum into the range of [0, 1]. The linear function is given by:

$$\gamma_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_U x_{iU} \tag{5}$$

where $x_{i1}$ ... $x_{iU}$ are the values for the *U* predictors on trial *i,* $\beta_1$ ... $\beta_U$ are the weights for the respective predictors and $\beta_0$ is the intercept term. The inverse logistic link function that the linear function is passed through is given by:

$$c_i = \frac{\exp(\gamma)}{1 + \exp(\gamma)} \tag{6}$$

We used a multilevel extension of this model, such that we estimate the predictor weights, $\beta$s, for each subject. These weights are assumed to come from the same population such that the individual-subject $\beta$s are estimated concurrently with the mean and variance of the weights at the population level (Wallis et al., 2015). For the confidence model, we used a model with 9 predictor variables, where the linear function was described by:

$$\begin{aligned} h_i = \ & \beta_{j,0} + \beta_{j,1} prior\ mismatch_i + \beta_{j,2} vertical\ energy_i \\ & + \beta_{j,3} horizontal\ energy_i + \beta_{j,4} oblique\ energy_i \\ & + \beta_{j,5}|phase_i| + \beta_{j,6}\log(|phase_i|) + \beta_{j,7} contrast_i \\ & + \beta_{j,8} rt_i + \beta_{j,9} experiment_i \end{aligned} \tag{7}$$

such that for subject *j*, $\beta_{j,1}$ ... $\beta_{j,5}$ were the weights for the respective predictor variables and $\beta_{j,0}$ was the intercept. $\beta_{j,9}$ was the weight for an experiment indicator variable which allowed for differences in mean confidence across experiments (see **Supplementary Figure 1**). All predictors were standardised, by subtracting the mean and dividing by the standard deviation, prior to model fitting (see **Supplementary Figure 7**). The computation of each predictor variable is described below.

***Orientation Energy***

The goal of the confidence model was to determine if participants' confidence judgements were informed by the same prior used for their perceptual judgements. We reasoned that if participants used the same prior to inform their confidence, there would be a statistical association between confidence and the amount of correspondence in the distribution of low-level features in the target and the prior. To summarise the degree of overlap between the distribution of orientation features in the target and the prior, we used two sets of predictors: *prior mismatch* and *cardinal and oblique orientation energy*. For both sets of predictors, we assumed that participants use prior knowledge about the statistics of the distribution of orientation features in natural scenes to inform their confidence. For the prior mismatch predictor, we assumed that observers directly compare a veridical representation of the distribution of orientation energy in the stimulus and veridical knowledge of the prior distribution of orientation energy to compute their confidence. This means that observers use a statistic that summarises the mismatch between the *entire* distribution of orientation energy in the target and the prior to compute their confidence. For the cardinal and oblique orientation energy predictors, in contrast, we assume that observers rely on only a *subset* of features in the stimulus and the prior model to compute their confidence. Specifically, observers use a set of salient orientation cues, namely vertical, horizontal, and oblique features, for confidence, based on prior knowledge about the probability of these orientation features in the natural environment.

**Prior Mismatch.** For the prior mismatch predictor, we calculated the average difference between the distribution of orientation energy in the target at the rotational offset chosen by the participant and the prior distribution of orientation energy. We did this by calculating the difference in orientation energy between the target and the prior in each orientation bin, squaring this difference and then summing across bins.

**Cardinal and Oblique Orientation Energy.** For the cardinal and oblique orientation energy predictors, we assumed that participants use a subset of orientation features to estimate their confidence. We calculated orientation energy across all orientations for each target positioned at the rotational offset chosen by the participant. We then normalised energy for each orientation band so that each band expressed a proportion of orientation energy in that bin relative to total orientation energy in the target, a form of divisive normalisation (Carandini & Heeger, 2012; see **Supplementary Figure 6**). To predict confidence, we use 3 predictors: *vertical energy* (where $\theta = 90 \cdot \frac{\pi}{180}$), *horizontal energy* (where $\theta = 0 \; or \; 180 \cdot \frac{\pi}{180}$), and *oblique energy* (where $\theta = 45 \cdot \frac{\pi}{180} \; or \; 135 \cdot \frac{\pi}{180}$).

***Phase***

To summarise the degree of overlap between the phase in the target and the prior, we used 2 *phase* predictors. For these predictors, we computed the response of a broadscale filter positioned at the centre of the target with an orientation that matched the rotational offset of the target. The response of the filter approximated the strength of the horizon in the target and could be used to approximate the strength of lighting in the scene (see **Figure 5C**). We used both the absolute value of the filter response and a log transformation of the absolute value of the filter response as predictors of confidence to allow for non-linear relationships between phase and confidence.

***Response Time***

As described above, we also considered the possibility that confidence depended on other heuristic cues not related to the prior. Based on previous research (Faivre et al., 2018; Patel et al., 2012; van den Berg, Anandalingam, et al., 2016), we therefore included a predictor in the confidence model for *response time*. For the response time predictor, we used the response time of the orientation response for each trial, measured in milliseconds.

***Contrast***

We also postulated that participants may use the contrast of the target as a heuristic cue for confidence. For the *contrast* predictor, we used the root-mean-square (RMS) contrast of each target (Bex & Makous, 2002; Harrison, 2022). RMS contrast is the standard deviation of the luminance (i.e., pixel) values:

$$C_{rms} = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(i_k - L)^2} \tag{8}$$

Where *k* is a pixel index, *N* is the total number of pixels, and *L* is the mean luminance.

***Summary of Modelling Approach***

To summarise our expectations about the confidence modelling results, we reasoned that if participants used the perceptual prior to inform their confidence, we would find that the measures which summarise the degree of overlap in orientation energy between the target and the prior (prior mismatch and/or cardinal/oblique orientation energy) and phase (absolute phase and/or log absolute phase) predicted confidence. If, however, participants used other non-prior related measures to inform their confidence, we would expect the other heuristic cues, such as the first order response time or contrast, to be the only predictors of confidence.

## References


Adams, W. J. (2007). A common light-prior for visual search, shape, and reflectance judgments. *Journal of Vision*, *7*(11), 11. https://doi.org/10.1167/7.11.11

Adler, W. T., & Ma, W. J. (2018). Comparing Bayesian and non-Bayesian accounts of human confidence reports. *PLOS Computational Biology*, *14*(11), e1006572. https://doi.org/10.1371/journal.pcbi.1006572

Aitchison, L., Bang, D., Bahrami, B., & Latham, P. E. (2015). Doubly Bayesian Analysis of Confidence in Perceptual Decision-Making. *PLOS Computational Biology*, *11*(10), e1004519. https://doi.org/10.1371/journal.pcbi.1004519

A-Izzeddin, E. J., Mattingley, J. B., & Harrison, W. J. (2022). *Contextual influences of visual perceptual inferences* [Preprint]. Neuroscience. https://doi.org/10.1101/2022.09.15.507892

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin*, *78*(4), 266–278. https://doi.org/10.1037/h0033117

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., & Pouget, A. (2008). Probabilistic Population Codes for Bayesian Decision Making. *Neuron*, *60*(6), 1142–1152. https://doi.org/10.1016/j.neuron.2008.09.021

Berkley, M. A., Kitterle, F., & Watkins, D. W. (1975). Grating visibility as a function of orientation and retinal eccentricity. *Vision Research*, *15*(2), 239–244. https://doi.org/10.1016/0042-6989(75)90213-8

Bertana, A., Chetverikov, A., van Bergen, R. S., Ling, S., & Jehee, J. F. M. (2021). Dual strategies in human confidence judgments. *Journal of Vision*, *21*(5), 21. https://doi.org/10.1167/jov.21.5.21

Bex, P. J., & Makous, W. (2002). Spatial frequency, phase, and the contrast of natural images. *Journal of the Optical Society of America A*, *19*(6), 1096. https://doi.org/10.1364/JOSAA.19.001096

Boldt, A., de Gardelle, V., & Yeung, N. (2017). The impact of evidence reliability on sensitivity and bias in decision confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(8), 1520–1531. https://doi.org/10.1037/xhp0000404

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brewster, D. (1826). On the optical illusion of the conversion of cameos into intaglios, and of intaglios into cameos with an account of other analogous phenomena. *Edinburgh Journal of Science*, *4*(7), 99–108.

Burge, J., & Geisler, W. S. (2011). Optimal defocus estimation in individual natural images. *Proceedings of the National Academy of Sciences*, *108*(40), 16849–16854. https://doi.org/10.1073/pnas.1108491108

Campbell, F. W., Kulikowski, J. J., & Levinson, J. (1966). The effect of orientation on the visual resolution of gratings. *The Journal of Physiology*, *187*(2), 427–436. https://doi.org/10.1113/jphysiol.1966.sp008100

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*, *197*(3), 551–566.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62. https://doi.org/10.1038/nrn3136

Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A*, *18*(5), 1016. https://doi.org/10.1364/JOSAA.18.001016

Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, *37*(22), 3181–3192. https://doi.org/10.1016/S0042-6989(97)00133-8

de Gardelle, V., Kouider, S., & Sackur, J. (2010). An oblique illusion modulated by visibility: Non-monotonic sensory integration in orientation processing. *Journal of Vision*, *10*(10), 6–6. https://doi.org/10.1167/10.10.6

Denison, R. N., Adler, W. T., Carrasco, M., & Ma, W. J. (2018). Humans incorporate attention-dependent uncertainty into perceptual decisions and confidence. *Proceedings of the National Academy of Sciences*, *115*(43), 11090–11095. https://doi.org/10.1073/pnas.1717720115

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective Confidence Predicts Information Seeking in Decision Making. *Psychological Science*, *29*(5), 761–778. https://doi.org/10.1177/0956797617744771

Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioral, Modeling, and Electrophysiological Evidence for Supramodality in Human Metacognition. *Journal of Neuroscience*, *38*(2), 263–277. https://doi.org/10.1523/JNEUROSCI.0322-17.2017

Geisler, W. S., & Perry, J. S. (2011). Statistics for optimal point prediction in natural images. *Journal of Vision*, *11*(12), 14–14. https://doi.org/10.1167/11.12.14

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, *14*(7), 926–932. https://doi.org/10.1038/nn.2831

Hangya, B., Sanders, J. I., & Kepecs, A. (2016). A Mathematical Framework for Statistical Decision Confidence. *Neural Computation*, *28*(9), 1840–1858. https://doi.org/10.1162/NECO_a_00864

Hansen, B. C., & Essock, E. A. (2004). A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes. *Journal of Vision*, *4*(12), 5. https://doi.org/10.1167/4.12.5

Hansen, B. C., Essock, E. A., Zheng, Y., & Deford, J. K. (2003). Perceptual anisotropies in visual processing and their relation to natural image statistics. *Network: Computation in Neural Systems*, *14*(3), 501–526. https://doi.org/10.1088/0954-898X_14_3_307

Hansen, B. C., Haun, A. M., & Essock, E. A. (2008). The “Horizontal Effect”: A perceptual ansiotropy in visual processing of naturalistic broadband stimuli. In T. A. Portocello & R. B. Velloti (Eds.), *Visual cortex: New research* (2nd ed.). Nova Science Publishers.

Harrison, W. J. (2022). Luminance and Contrast of Images in the THINGS Database. *Perception*, *51*(4), 244–262. https://doi.org/10.1177/03010066221083397

Harrison, W. J., Bays, P. M., & Rideaux, R. (2023). *Neural tuning instantiates prior expectations in the human visual system* (p. 2023.01.26.525790). bioRxiv. https://doi.org/10.1101/2023.01.26.525790

Li, H.-H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. *Nature Communications*, *11*(1), 2004. https://doi.org/10.1038/s41467-020-15581-6

Lisi, M., Mongillo, G., Milne, G., Dekker, T., & Gorea, A. (2021). Discrete confidence levels revealed by sequential decisions. *Nature Human Behaviour*, *5*(2), 273–280. https://doi.org/10.1038/s41562-020-00953-1

Locke, S. M., Landy, M. S., & Mamassian, P. (2022). Suprathreshold perceptual decisions constrain models of confidence. *PLOS Computational Biology*, *18*(7), e1010318. https://doi.org/10.1371/journal.pcbi.1010318

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*(11), 1432–1438. https://doi.org/10.1038/nn1790

Ma, W. J., & Jazayeri, M. (2014). Neural Coding of Uncertainty and Probability. *Annual Review of Neuroscience*, *37*(1), 205–220. https://doi.org/10.1146/annurev-neuro-071013-014017

Metzger, W. (1936). *Gesetze des Sehens. [Laws of vision.]* (pp. xvi, 172). Kramer.

Michael, E., de Gardelle, V., Nevado-Holgado, A., & Summerfield, C. (2015). Unreliable Evidence: 2 Sources of Uncertainty During Perceptual Choice. *Cerebral Cortex*, *25*(4), 937–947. https://doi.org/10.1093/cercor/bht287

Mole, C. D., Jersakova, R., Kountouriotis, G. K., Moulin, C. J., & Wilkie, R. M. (2018). Metacognitive judgements of perceptual-motor steering performance. *Quarterly Journal of Experimental Psychology*, *71*(10), 2223–2234. https://doi.org/10.1177/1747021817737496

Murray, R. F. (2013). The Statistics of Shape, Reflectance, and Lighting in Real-World Scenes. In S. J. Dickinson & Z. Pizlo (Eds.), *Shape Perception in Human and Computer Vision: An Interdisciplinary Perspective* (pp. 225–235). Springer London. https://doi.org/10.1007/978-1-4471-5195-1_16

Navajas, J., Hindocha, C., Foda, H., Keramati, M., Latham, P. E., & Bahrami, B. (2017). The idiosyncratic nature of confidence. *Nature Human Behaviour*, *1*(11), 810–818. https://doi.org/10.1038/s41562-017-0215-1

Patel, D., Fleming, S. M., & Kilner, J. M. (2012). Inferring subjective states through the observation of actions. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1748), 4853–4860. https://doi.org/10.1098/rspb.2012.1847

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374. https://doi.org/10.1038/nn.4240

Pratte, M. S., Park, Y. E., Rademaker, R. L., & Tong, F. (2016). Accounting for stimulus-specific variation in precision reveals a discrete capacity limit in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(1), 6–17. https://doi.org/10.1037/xhp0000302

Qamar, A. T., Cotton, R. J., George, R. G., Beck, J. M., Prezhdo, E., Laudano, A., Tolias, A. S., & Ma, W. J. (2013). Trial-to-trial, uncertainty-based adjustment of decision boundaries in visual categorization. *Proceedings of the National Academy of Sciences*, *110*(50), 20332–20337. https://doi.org/10.1073/pnas.1219756110

Ramachandran, V. S. (1988). *Perception of shape from shading*.

Rideaux, R., West, R. K., Wallis, T. S. A., Bex, P. J., Mattingley, J. B., & Harrison, W. J. (2022a). Spatial structure, phase, and the contrast of natural images. *Journal of Vision*, *22*(1), 1–19. https://doi.org/10.1167/jov.22.1.4

Rideaux, R., West, R. K., Wallis, T. S. A., Bex, P. J., Mattingley, J. B., & Harrison, W. J. (2022b). Spatial structure, phase, and the contrast of natural images. *Journal of Vision*, *22*(1), 4. https://doi.org/10.1167/jov.22.1.4

Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron*, *90*(3), 499–506. https://doi.org/10.1016/j.neuron.2016.03.025

Spence, M. L., Dux, P. E., & Arnold, D. H. (2016). Computations underlying confidence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(5), 671–682. https://doi.org/10.1037/xhp0000179

van den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). A common mechanism underlies changes of mind about decisions and confidence. *ELife*, *5*, e12192. https://doi.org/10.7554/eLife.12192

van den Berg, R., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). Confidence Is the Bridge between Multi-stage Decisions. *Current Biology*, *26*(23), 3157–3168. https://doi.org/10.1016/j.cub.2016.10.021

Wallis, T. S. A., Dorr, M., & Bex, P. J. (2015). Sensitivity to gaze-contingent contrast increments in naturalistic movies: An exploratory report and model comparison. *Journal of Vision*, *15*(8), 3. https://doi.org/10.1167/15.8.3

Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. *Nature Neuroscience*, *18*(10), 1509–1517. https://doi.org/10.1038/nn.4105

West, R. K., Harrison, W. J., Matthews, N., Mattingley, J. B., & Sewell, D. K. (2022). *Modality Independent or Modality Specific? Common Computations Underlie Confidence Judgements in Visual and Auditory Decisions* [Preprint]. Neuroscience. https://doi.org/10.1101/2022.10.31.514447