[PREPRINT: DOI 10.17605/OSF.IO/MT2QK]

Guessing reveals internal models of perceptual precision

Caroline Myers*, Chaz Firestone, Justin Halberda

Department of Psychological and Brain Sciences, Johns Hopkins University
*Corresponding author: cmyers60@jhu.edu

Abstract

When observers lack sufficient information to support a confident response, they often guess. Guessing plays a pervasive role in visual cognition and working memory, yet the mechanisms that govern how observers generate guesses remain poorly understood. Standard models traditionally assume that responses produced in the absence of information are either uniformly distributed over feature space or are perhaps weighted towards prevailing environmental statistics. In contrast, here we consider an intriguing alternative: that guesses incorporate observers' knowledge of their own perceptual capacities. We empirically measured guessing by eliciting responses under extreme target uncertainty (Experiment 1) as well as a novel "0ms presentation" approach in which no stimulus appeared but subjects believed one had (Experiment 2). We evaluated three accounts of guesses under these conditions: unsystematic (lapse) responding, biases toward environmental statistics, and a self-representational account in which guesses reflect observers' knowledge of their own feature-dependent precision (e.g., preferring to guess feature values they believe they would be likely to miss). Guess responses were non-uniform and systematically biased toward feature values typically encoded with the least precision (e.g., oblique orientations) — a counterintuitive bias away from high-frequency, high-fidelity feature values (e.g., cardinal orientations). This complementary relationship between guessing and perceptual fidelity held within individuals and across paradigms, and was recoverable via an empirical-guess mixture model that replaced the standard uniform assumption with empirically measured guess distributions. Our findings challenge prevailing views that guesses reflect random noise, and suggest instead that guessing behavior reflects metacognitive knowledge of internal precision. Rather than defaulting to environmental priors, observers appear to model their own sensory limitations and leverage these representations to inform decisions in the absence of evidence. These results reframe guessing as a theoretically informative behavior that expresses observers' own beliefs about their perceptual capacities.

Significance

Guessing is commonly treated as random noise in models of perception and memory, assumed to reflect lapses or uninformed responses. Instead, we show that human guesses are systematically structured across feature space: observers preferentially guess values they typically encode with the least precision, revealing a consistent, strategic bias away from high-fidelity representations. By directly measuring guess behavior on stimulus-absent trials and integrating these empirical distributions into a mixture model, we find that guesses on stimulus-present trials can be systematically recovered, and that they too form the complement of perceptual precision. These findings challenge foundational psychophysical modeling assumptions and position guessing as a strategic, informative behavior that engages self-representation.

MAIN TEXT

When we don't know what we've seen, we often guess. This is familiar in daily life and ubiquitous in behavioral experiments that probe perception and memory. Despite its central role, guessing is typically treated as theoretically uninformative. Standard models assume that responses produced in the absence of perceptual information are random, approximated by a uniform distribution over feature space and captured by a lapse parameter. This assumption underlies widely used mixture models and shapes how guess rates are interpreted across behavioral, neural, and computational studies (Prins 2012; Klein, 2001; Zchaluk & Foster, 2009; Wichmann & Hill, 2001), but it has never been tested directly.

In principle, guessing need not be random. At least three distinct computations are possible. Guess responses could be unsystematic, reflecting arbitrary responding when information is absent. Alternatively, guessing could be systematic in a manner consistent with environmental priors, such that observers leverage knowledge that some feature values are more probable in the world than others (e.g., a preference for cardinal orientations, which are more prevalent in natural scenes; Girshick, Landy, & Simoncelli 2011; Harrison, Bays & Rideaux, 2023; Oliva 2001; Hansen & Essock 2004; Girshick, Simoncelli & Landy, 2011; Appelle, 1972). A third possibility, central here, is that guessing reflects self-representation. On this view, observers possess internal knowledge of how precisely different feature values are internally represented, and strategically condition their guesses on this knowledge when perceptual evidence is insufficient. Guessing would therefore be structured by feature-dependent precision, rather than by external stimulus statistics.

This self-representational account makes a distinctive and counterintuitive prediction. When forced to guess, observers should avoid feature values they normally represent with high precision and instead preferentially sample from low-precision regions of the feature space. In orientation, this predicts an increased tendency to guess oblique orientations and a systematic avoidance of cardinal orientations, despite the fact that cardinals are both perceptually advantaged and more prevalent in natural scenes. Such a pattern would contradict environmental-prior accounts and instead imply that guess behavior reflects metacognitive knowledge of one's own perceptual limitations.

Here, we tested these alternatives in orientation report tasks using two complementary approaches. In Experiment 1, we elicited guess-like responses under extreme target uncertainty using extremely high load and brief exposure trials, isolating responses that are effectively untuned to the target. In Experiment 2, we embed backward-masked 0-ms trials in which no stimulus is presented, providing an observer-specific empirical measure of pure guessing. These empirical guess distributions allow us to test whether guessing is uniform, prior-driven, or structured by internal precision. We further use these distributions to constrain a two-component mixture model, replacing the standard uniform guess term with each observer's measured guess density. This approach yields trial-wise posterior estimates that distinguish responses likely driven by target-centered internal representations from those best characterized as guesses, enabling a direct test of whether guessing and perceptual precision exhibit a systematic relationship.To preview our results, we found that guesses were reliably non-uniform and biased toward feature values observers typically represent with the least precision. Thus, guessing behavior provides a direct readout of internal models of representational fidelity, revealing that observers not only monitor their own perceptual limitations, but actively exploit this knowledge to guide behavior when evidence is absent.

RESULTS

Experiment 1

In Experiment 1, (n = 10) observers completed a continuous-report working memory orientation task with two interleaved two trial types designed to separate high-fidelity encoding from guess-like responding: low-load, long-duration “precision” trials (1 item, 1000 ms) and high-load, ultra-brief “guess” trials (36 items, 16 ms; see **SI Methods**).

Precision and guess trials engage distinct response regimes.
Responses on 1-item, 1000ms precision trials tightly tracked the target orientation (Pearson $r = .892$, 95% bootstrap CI [0.862, 0.922]., $p < .0001$; **Figure 2A**). By contrast, responses on 36-item, 16 ms guess trials were essentially untuned to the target (Pearson $r = -.011$, 95% bootstrap CI [−0.0582, 0.0336], $p > .5$; **Figure 2B**). The within-observer difference was reliable (mean $\Delta r = 0.908$, CI [0.863, 0.949]; exact sign test $p < 0.001$).

Guess responses are non-uniform across orientation space.
To characterize guess behavior, responses from 36-item/16-ms trials were collapsed across targets (to which they were unresponsive) and fit using a circular kernel density estimate over the 180° axial orientation space. If guessing were a pure random process, the distribution of responses on guess trials should be approximately uniform. Instead, guess response distributions deviated from uniformity in 9/10 observers (Kuiper permutation $p < 0.05$; **Figure 2C**).

Guessing forms the complement of precision.
We next tested whether guessing relates systematically to an observer’s own feature-dependent precision. Each observer’s precision landscape was estimated from 1-item/1000-ms trials as the inverse of the smoothed SD of signed error as a function of target orientation (**SI Methods**). If guess behavior were driven by an environmental prior favoring cardinals, the guess distribution should peak near those high-precision axes. In contrast, a self-representational account predicts that guess density should be highest where the observer’s precision is lowest– i.e., that guessing should align with the complement of the precision landscape.

Guess densities were compared to precision landscapes at their true physical alignment (relative to the cardinal axes), using an exact circular-shift alignment test that preserves smoothness while breaking physical correspondence (SI Methods). Guess density and precision were anticorrelated (mean $r = -0.336$, 95% CI [−0.519, −0.132]; sign-flip p(left)

= 0.007805; **Figure 2C**), consistent with guesses concentrating in orientations that the same observer represents least precisely.

Experiment 2

Pure guessing is structured and supports empirical-guess mixture modeling.
Experiment 1 used conditions intended to induce guess-like states, but weak stimulus traces could in principle still influence responding even under extreme load and brief exposure. Experiment 2 therefore embedded backward-masked 0-ms (stimulus-absent) trials among stimulus-present trials varying set size (1, 3, 5) and display duration (16–300 ms; **SI Methods**).

These stimulus-absent trials both (1) provide an observer-specific measure of pure guessing and (2) enable a stronger test of the complement account: if guessing reflects self-represented precision limitations, complementarity should be evident not only on stimulus-absent trials, but also in the subset of stimulus-present trials on which observers effectively guessed.

Stimulus-absent trials yield a non-uniform empirical guess distribution
Responses on 0-ms trials were systematically non-uniform in 10/10 observers (within-observer Kuiper permutation test $p < 0.05$; **Figure 3**). Moreover, each observer's 0-ms guess density was positively similar to that same observer's induced-guess density from Experiment 1 (mean within-observer $r = 0.444$, 95% CI [0.199, 0.635]; sign test $p = 0.0107$), consistent with a shared guessing strategy across paradigms.

Empirical-guess mixture model
Standard continuous-report mixture models treat guesses as uniformly distributed over the response space, and attempt to control for guessing via the inclusion of a parametric lapse parameter. The preceding results show that this assumption is violated in a reliable, structured way. More importantly, they raise a deeper inferential problem: when guessing is structured in absolute feature space, collapsing responses into target-centered error distributions obscures this structure and renders guessing statistically indistinguishable from extreme imprecision. Can guess behavior be empirically identified and separated from stimulus-driven representations without discarding the structure it exhibits across feature values?

We fit a two-component mixture model to target-present trials in Experiment 2 (**Figure 4A)**, with (1) an internal component defined by a normalized target-centered distribution whose SD varies with target orientation according to a two-parameter symmetric oblique-effect function and (2) a nonparametric guess component fit to the observer's measured 0-ms guess KDE **(SI Methods)**.

Trial-wise posteriors recover structured guessing on stimulus-present trials.
To test whether the same structured guessing measured on 0-ms trials also operates when a stimulus is physically present, we reconstructed a guess-weighted response density from stimulus-present trials by weighting each response by its posterior probability of guessing and then estimating a circular KDE on these weighted

responses. The recovered guess density from stimulus-present trials closely matched the independently measured 0-ms guess density (mean r = 0.869, 95% bootstrap CI [0.836, 0.902]; **Figure 4B**), indicating that the model's trial-wise separation recovers empirical guess structure from stimulus-present data.

We also quantified the expected recovered q_0 similarity under a null in which stimulus-present guesses are uniform across orientation. Specifically, for each observer we simulated stimulus-present responses preserving the fitted internal component and trial structure, but drawing guess responses from a uniform distribution; we then applied the identical posterior-weighted KDE recovery procedure. All observers' recovered–q_0 correlations exceeded this uniform-guess null (p < . 0.01). indicating that the recovered–q_0 match is not an artifact of using q_0 in the posterior.

Guessing on stimulus-present trials is biased toward low-precision orientations. We then tested the key theoretical prediction: if guessing reflects self-representational knowledge of perceptual limitations, guess-like responses on stimulus-present trials should be biased toward the least precise orientations and away from the most precise orientations. Using the recovered $(1-z)$-weighted guess density from stimulus-present trials, complementarity with the independently measured Experiment 1 precision landscape was reliably negative across observers (mean r = -0.495, 95% bootstrap CI [−0.669, −0.311]; one-sided sign-flip test p(left) = 0.00195). Thus, on the subset of stimulus-present trials where observers behave as if they are guessing, their responses are systematically biased toward the least precise regions of their internal feature space.

DISCUSSION

Across two experiments, we directly measured guess behavior under conditions of extreme uncertainty (Experiment 1) and on completely stimulus-absent trials (Experiment 2) and found that guess responses were reliably non-uniform. Guess density varied systematically with orientation in a manner that closely tracked the inverse of each observer's precision landscape. Observers preferentially guessed orientations they represented least precisely and avoided orientations they represented most precisely. This inverse relationship was evident within individual observers and was recovered independently from both stimulus-absent trials and guess-like responses on stimulus-present trials.

These findings bear directly on how continuous-report data are modeled and interpreted. In standard discrete-state working-memory mixture models, responses are decomposed into an internal-representation component and a guess component assumed to be uniform over feature space (Zhang & Luck, 2008; Rouder et al., 2008; Adam, Vogel & Awh, 2017). Variable precision models reject this dichotomy, treating all responses as noisy but meaningful, and positing a continuous distribution over representational fidelity (van den Berg et al., 2012; Schurgin, Wixted & Brady, 2020). In these models, large errors are assumed to arise from extremely low precision, rather than the absence of memory altogether.

Our results place new pressure on both classes of models. The defining assumption shared by both is that guesses are unstructured: either uniform (in the discrete case), or non-existent (in the continuous case). But when guess behavior is structured, this assumption fails in principled ways.

Our results show that when internal signal is absent, guessing is not uniform, and when signal is present but degraded, guess-like responses still follow orientation-dependent patterns that are unrelated to the stimulus shown. These patterns are consistent across stimulus-absent and stimulus-present conditions and are defined in orientation space, not error space. That is, the structure in guessing is organized around internal feature representation, not the stimulus target.

This introduces a problem of misattribution. For discrete models, the estimated guess weight no longer isolates trials without information, because responses produced under guessing are not exchangeable across feature values. As a result, orientation-dependent regularities in guessing may be misinterpreted as properties of the internal representation, such as reduced precision or increased variability. For continuous models, this poses a deeper problem: The assumption that even guess responses reflect degraded signal becomes difficult to maintain once systematic biases appear even in the absence of a stimulus.

In both cases, the presence of orientation-specific structure in guess responses introduces parameter ambiguity. Model fits may conflate true imprecision with feature-dependent bias in guessing, echoing prior critiques of how mixture model outputs can be distorted by misattributed structure in error distributions (Panichello, DePasquale & Buschman, 2019; Bays, 2014). This misattribution is compounded when responses in continuous-report tasks are collapsed into target-centered error distributions. This effectively removes absolute feature information and projects all responses into a space defined as relative to the target. A non-uniform distribution of responses that are truly independent from any target (e.g., guesses) will therefore appear as uniform noise in this error space, not because the responses lack structure, but because the structure is defined in absolute feature space rather than relative to the target. Models fit to such collapsed error distributions may therefore appear to support uniform guessing or extremely low-precision representations, not because those assumptions are warranted by the data, but because the preprocessing step itself renders alternative explanations impossible. In this sense, common modeling conclusions about guessing and precision reflect how the data are transformed rather than how observers actually respond. This also reflects a deeper modeling assumption: that responses without target correspondence lack meaningful structure. Our results directly challenge that assumption.

We resolve this tension by introducing an empirical-guess mixture model that replaces the uniform guess term with each observer's measured guess distribution. This model both (1) captures non-uniform, structured guessing, without introducing new parameters, and (2) provides posterior estimates that separate stimulus-present trials based on whether they reflect an internal representation of the stimulus or informed guessing. Indeed, replacing the uniform guess term with each observer's empirically measured 0-ms guess distribution produced a large and consistent improvement in model fit in all observers, indicating that structured guessing captures systematic variance missed by standard lapse assumptions.

This raises a natural question about the computational role of guessing: What does it mean to guess strategically? Our findings suggest that even in the absence of perceptual evidence, observers draw on internal knowledge of their own perceptual limitations to constrain responses to likely guesses. In orientation space, this manifests as a systematic bias toward oblique orientations where perceptual precision is poor.

This pattern runs counter to that predicted by most standard Bayesian observer models, which emphasize the role of environmental regularities in shaping perceptual estimates (Wei & Stocker, 2015; Stocker & Simoncelli, 2006). For instance, observers are typically biased toward cardinal orientations, a pattern long attributed to their statistical prevalence in visual environments (Girshick et al., 2011; Harrison, Bays & Rideaux, 2023; Coppola et al., 1998; Howe & Purves, 2005; Hansen & Essock, 2004). These environmental anisotropies are mirrored in perceptual performance, where thresholds for detecting or discriminating cardinal orientations are consistently lower than for oblique ones, a well-established phenomenon known as the oblique effect (Appelle, 1972; Essock, 1980; Heeley & Timney, 1988; Westheimer 2003).

Importantly, these behavioral asymmetries also align with anisotropies in neural tuning: cardinal preferences are reliably observed in early visual cortex, including in coarse-scale population coding in V1 (Roth, Kay & Merriam, 2022) and orientation-selective BOLD responses (Patten, Mannion & Clifford, 2017), supporting a model in which neural architecture and perceptual priors are jointly shaped by environmental regularities (Geisler, 2008).

However, our results show that when external structure is absent or unavailable, observers default not to what is likely in the world, but to what is likely given their own representational limitations. This is not a prior over the stimulus, but a bias conditioned on introspective uncertainty that cannot be captured by models that conflate encoding priors with perceptual likelihoods (even those that unify efficient coding with Bayesian decoding; Wei & Stocker, 2015). In those models, repulsive perceptual biases arise from asymmetric likelihoods shaped by natural stimulus statistics. But our data reveal repulsive biases in the absence of a stimulus altogether, suggesting that these patterns reflect internal models of sensory fidelity rather than environmental probability.

A guessing bias conditioned on introspective uncertainty may be more consistent with theories of resource-rational inference, in which decision strategies are shaped by internal constraints as much as by external statistics (Lieder & Griffiths, 2017; Prat-Carrabin et al., 2024; Van den Berg & Ma, 2018). Rather than optimize guess responses for environmental likelihood (e.g., guessing high-prevalence cardinal orientations), observers optimize for representational plausibility: when uncertain, they report what they expect themselves to miss. This form of introspective strategy is not captured by standard Bayesian accounts, which only assume access to environmental priors, nor by more general models that deny metacognitive access to internal noise.

Instead, it reflects a form of pragmatic metacognition, where responses are generated based on knowledge of one's own uncertainty. This interpretation is supported by emerging work suggesting that confidence and precision can be tracked even in the absence of external stimuli, and that metacognitive judgments may rely on internal monitoring of representational fidelity (Fleming & Daw, 2017; Boldt, de Gardelle & Yeung, 2017). Our results extend these findings by demonstrating that such monitoring influences primary task responses.

This complementarity links guessing to self-knowledge about perceptual representation. When perceptual evidence was insufficient to support an internal representation of the target, responses were biased toward orientations associated with lower precision. This supplies a behavioral route to metacognition in a domain where measurement is typically indirect. Standard approaches rely on confidence judgments or other secondary reports that are vulnerable to both response bias and the tight coupling between performance and awareness, making it difficult to isolate metacognitive knowledge from first-order task demands (Lau, 2008; Morales, Chiang, & Lau, 2015). Moreover, existing measures necessarily depend on the presentation of a stimulus, introducing additional variables that complicate interpretation (Morales, Odegaard, & Maniscalco, 2022). In our tasks, observers' beliefs about their own precision are expressed in the primary report itself. The structure of guessing therefore provides a quantitative readout of observers' models of their own perceptual limitations, without requiring explicit report or relying on stimulus-driven responses.

In sum, guessing is structured, feature-specific, and systematically related to perceptual precision. Measuring that structure reveals a complementary relationship between guessing and precision and provides a new tool for metacognitive inference.

More broadly, our results challenge the assumption that guess responses reflect failure. When observers lack evidence, they still respond in ways that reflect internal models of error and representational uncertainty. Recognizing this reframes guessing as a theoretically informative behavior and establishes a new approach for probing internal models of perceptual representation without relying on confidence judgments or explicit report.

SUPPLEMENTARY MATERIALS AND METHODS

Participants

Ten participants participated in both experiments in a fully within-subjects design. Participants were recruited via campus recruitment and the Johns Hopkins University SONA subject pool and received course credit for completion of one-hour in-house sessions. All participants reported normal or corrected-to-normal vision. Study procedures were approved by the Johns Hopkins University Institutional Review Board.

Apparatus and software

All stimuli were generated and presented using custom JavaScript in conjunction with the jsPsych framework (de Leeuw, 2015) and run locally on a high–refresh-rate LCD monitor (nominal refresh rate: 120 Hz) in a dark, sound-attenuated testing room.

Viewing distance was fixed at 57 cm using a chinrest. The display was calibrated and luminance-linearized using a vPixx i1Display Pro Spectra-Colorimeter photometer. All preprocessing, statistical analyses, and modeling were conducted using custom-written MATLAB code (R2017b).

Stimuli

Stimuli in both experiments were sinusoidal Gabor patches presented on a uniform gray background. Patches were windowed by a circular Gaussian envelope and rendered at full Michelson contrast. Orientation was defined over a 180° axial space, such that orientations separated by 180° were treated as equivalent. On trials with multiple items, orientations were sampled without replacement subject to a minimum angular separation of 5° to avoid identical or near-identical orientations.

Experiment 1: Guess-like responding under extreme uncertainty

Task overview

Participants completed a continuous-report orientation memory task designed to contrast high-fidelity encoding with guess-like responding under extreme uncertainty. Each participant completed a single one-hour session comprising 360 total trials. Two trial types were randomly interleaved throughout the session: precision trials (50%) and guess trials (50%).

Trial sequence

Each trial began with a centrally presented fixation cross. Participants initiated the trial by clicking the fixation cross with the mouse. A black circular response ring remained visible throughout the trial. Responses were made by adjusting the orientation of a centrally presented Gabor patch using the mouse and confirming the response with a mouse click. Accuracy was emphasized over response time.

Precision trials

On precision trials, a single oriented Gabor patch (180 px diameter) was presented for 1000 ms at one of nine equally spaced locations on an imaginary circle centered on fixation (radius: 320 px). Target orientation was sampled uniformly from the 180° orientation space. Following stimulus offset, the response display appeared immediately. The response display consisted of an adjustable central Gabor patch and a solid black outline marking the target location. Participants adjusted the central Gabor to match the remembered target orientation and confirmed their response with a mouse click. The starting orientation of the response Gabor was randomized on each trial to minimize motor and response biases. Accuracy was emphasized over speed.

Guess trials

Guess trials were designed to elicit responses under conditions of extreme perceptual uncertainty while preserving the general structure of precision trials. On each guess trial, thirty-six Gabor patches (one target and 35 distractors; each 45 px diameter) were presented simultaneously for 16 ms. Stimuli were positioned at locations sampled without replacement from a fixed set of 91 possible spatial positions distributed across four concentric rings centered on fixation (radii: 120, 200, 280, and 360 px). Immediately following stimulus presentation, all items were masked for 132 ms by spatially overlapping, randomly oriented Gabor patches drawn uniformly from the stimulus space (90 X 90 px). At response, the target location was indicated by a solid black outline, and distractor locations were indicated by dashed outlines to minimize spatial uncertainty. Participants reported the target orientation using the same continuous-report procedure as on precision trials.

Experiment 2: Stimulus-absent guessing and empirical-guess modeling

Task overview
Experiment 2 extended the task used in Experiment 1 to directly measure guessing and to support model-based separation of guessing from target-driven responses. Participants completed four one-hour in-house sessions comprising 3600 total trials. Apparatus, response method, and general task structure were identical to Experiment 1.

Trial sequence
Each trial began with a centrally presented fixation cross. Participants initiated the trial by clicking the fixation cross with the mouse. A black circular response ring remained visible throughout the trial. Responses were made by adjusting the orientation of a centrally presented Gabor patch using the mouse and confirming the response with a mouse click. Accuracy was emphasized over response speed.

Stimulus-present trials
On stimulus-present trials, oriented Gabor patches (90 px diameter) were presented at set sizes of 1, 3, or 5 items for one of four display durations (16, 66, 132, or 300 ms). Stimuli were positioned at locations sampled without replacement from a fixed set of 24 possible spatial positions distributed across two concentric rings centered on fixation (radii: 120 and 300 px).

Following stimulus offset, all trials were immediately followed by a 132-ms backward mask composed of randomly oriented Gabors drawn uniformly from the stimulus space. Following the masks, the response display appeared. The target location was indicated by a solid outline: dashed placeholders indicated locations of distractors to minimize location uncertainty at the time of response. Participants reported the target orientation using the same continuous-report procedure as in Experiment 1.

Stimulus-absent trials
On stimulus-absent (0-ms) trials, no stimulus was presented prior to the mask. These trials were otherwise identical to stimulus-present trials, including masking and response displays, and were randomly interleaved with stimulus-present trials. Participants were not informed of the presence of stimulus-absent trials. Participants were probed following successful completion of the experiment: none none spontaneously identified the purpose of the study none spontaneously identified the purpose of the study, the presence of the 0-ms trials, or reported any strategies consistent with deliberate or strategic guessing.

QUANTIFICATION AND STATISTICAL ANALYSIS

Experiment 1

Axial orientation transformation

All analyses were conducted in an axial 180° orientation space (i.e., θ and θ+180° are equivalent). We represent orientations on a 1° grid:

$$r \in \{1,2,\ldots,180\},$$

Continuous (non-integer) target and response angles are wrapped to this axial domain by:

$$\mathrm{wrap}_{180}(x) \;=\; ((x-1)\,mod\,180)+1,$$

which maps values into $[1,180]$ while preserving degrees.

The signed axial difference (response error) between response r and target t was defined as:

$$\Delta(r,t) \;=\; ((r-t+90)\,mod\,180)-90, \Delta(r,t) \in [-90°,90°).$$

We also use the corresponding axial distance for computations in which only error magnitude matters as:

$$\delta(r,t) = \mid \Delta(r,t) \mid \in [0°,90°],$$

Circular KDE for guess densities

Circular guess densities $q(r)$were estimated over the 180° axial orientation space using a circular Gaussian kernel density estimate (KDE) evaluated on the 1° grid $r \in \{1,\ldots,180\}$. Given a set of orientation responses $\{r_j\}_{j=1}^{N}$, we compute the unnormalized guess density:

$$\tilde{q}(r) \;=\; \sum_{j=1}^{N} \exp\left[-\frac{1}{2}\left(\frac{\delta(r,r_j)}{h}\right)^2\right],$$

with bandwidth $h = 3°$. The KDE is then normalized to unit mass on the grid:

$$q(r) \;=\; \frac{\tilde{q}(r)}{\sum_{r'=1}^{180}\tilde{q}\,(r')}, \sum_{r=1}^{180} q(r) = 1.$$

Non-uniformity

Within each observer, deviation from uniformity was tested with Kuiper's statistic on the empirical CDF of responses mapped to [0, 1). Significance values were computed by permutation against n_{perm} = 20,000 samples of matched size.

Precision landscape

Precision landscapes were estimated from precision trials as the inverse of the target-dependent standard deviation of signed axial error. For precision trial i, signed error was

$$e_i = \Delta(r_i, t_i) \in [-90°, 90°),$$

where $t_i$is the target orientation and r_i is the response orientation.

For each orientation bin $\theta \in \{1, \dots, 180\}$, we compute Gaussian weights over target orientation using axial distance:

$$d_i(\theta) = \min\left(|\, t_i - \theta \,|\,,\, 180 - |\, t_i - \theta \,|\right), \qquad w_i(\theta) = \exp\left[-\frac{1}{2}\left(\frac{d_i(\theta)}{\sigma}\right)^2\right],$$

with $\sigma = 10°$. The weighted mean and SD of error are

$$\mu(\theta) = \frac{\sum_i w_i\,(\theta) e_i}{\sum_i w_i\,(\theta)}, SD(\theta) = \sqrt{\frac{\sum_i w_i\,(\theta)\,(e_i - \mu(\theta))^2}{\sum_i w_i\,(\theta)}}.$$

Precision is defined as:

$$P(\theta) = \frac{1}{SD(\theta)}\,(\deg^{-1}).$$

Complementarity and exact circular-shift tests

Complementarity was assessed within each observer as the Pearson correlation between the observer's guess density and precision landscape at their true physical alignment in absolute orientation space.

Let $q(r)$denote the observer's guess KDE and $P(\theta)$ denote the observer's precision landscape (both defined on the same 1° grid). To match smoothness across the two curves for correlation and alignment testing, both curves were additionally smoothed with the same circular Gaussian kernel ($\sigma = 10°$; truncated at $\pm 4\sigma$ and normalized), producing $q_{\text{smooth}}$and P_{smooth}.

The observed complementarity statistic is the correlation at true alignment:

$$r_{\text{obs}} = \text{corr}(q_{\text{smooth}}, P_{\text{smooth}}).$$

To test whether this correlation depended on physical alignment (rather than arbitrary phase matching induced by smoothing), we constructed an exact circular-shift null by shifting one curve by all non-zero offsets $k \in \{1, \ldots, 179\}$and recomputing the correlation:

$$r_k = \mathrm{corr}\,(\mathrm{circshift}(q_{\mathrm{smooth}}, k), P_{\mathrm{smooth}}).$$

For a **left-tailed** complementarity test (negative correlation), the exact p-value is

$$p_{\mathrm{left}} = \frac{\#\{k : r_k \le r_{\mathrm{obs}}\} + 1}{180}.$$

Group-level confidence intervals over observers were computed by bootstrapping observers (5,000 resamples). Group-level p-values for a mean correlation were computed with a one-sided sign-flip test on observer-level correlations.

Experiment 1 target-response tuning (precision vs. guess trials)
Within each observer, target–response tuning was quantified as the Pearson correlation between target orientation and response orientation on precision trials and on guess trials. For group-level inference on the within-observer difference in tuning $(r_{\mathrm{precision}} - r_{\mathrm{guess}})$, we used an exact sign test on the number of observers showing stronger tuning on precision trials.

Experiment 2 and mixture model

Model definition
We fit, within each observer, a two-component mixture model to stimulus-present trials in Experiment 2. For trial i, let $t_i$be target orientation, r_i response orientation, and c_i index the condition (set size × display duration). The response likelihood is:

$$p(r_i \mid t_i, c_i) \;=\; w_{c_i}\, p_{\mathrm{int}}(r_i \mid t_i; a, b) \;+\; (1 - w_{c_i})\, q_0(r_i),$$

where:

(1) $q_0(r)$is the observer's fixed 0-ms KDE (unit mass on the 1° grid: $\sum_{r=1}^{180} q_0\,(r) = 1$); in likelihood computations, $q_0(r_i)$is evaluated at continuous $r_i$by circular linear **interpolation** on the 1° grid.
(2) $w_c \in [0,1]$is the condition-specific weight of the internal (target-centered) component.
(3) $p_{\mathrm{int}}$is the internal component, defined as a target-centered axial Gaussian whose width depends on target orientation via a two-parameter "bowtie" function.

The internal component is defined as:

$$p_{\text{int}}(r_i \mid t_i; a, b) = \frac{\exp\left[-\frac{1}{2}\left(\frac{\delta(r_i, t_i)}{SD(t_i; a, b)}\right)^2\right]}{\sum_{r=1}^{180} \exp\left[-\frac{1}{2}\left(\frac{\delta(r, t_i)}{SD(t_i; a, b)}\right)^2\right]},$$

where $\delta(\cdot,\cdot)$is axial distance and the orientation-dependent SD is:

$$SD(\theta; a, b) = \frac{a+b}{2} - \frac{a-b}{2}\cos(4\theta).$$

(Here, θ is in degrees; equivalently $\cos(4\theta)$can be written $\cos(4\theta\pi/180)$. This parameterization yields $SD = a$ at obliques and $SD = b$ at cardinals; during fitting we enforced $a \geq b$, consistent with known anisotropies in human precision across orientation space.

EM fitting and posteriors

Parameters were fit via an EM algorithm with a, b shared across conditions within each observer and w_c allowed to vary by condition. The trial-wise posterior probability that a response arose from the internal component is:

$$z_i = P(\text{int} \mid r_i, t_i, c_i) = \frac{w_{c_i}\, p_{\text{int}}(r_i \mid t_i; a, b)}{w_{c_i}\, p_{\text{int}}(r_i \mid t_i; a, b) + (1 - w_{c_i})\, q_0(r_i)}.$$

M-step updates

Mixture weights were updated in closed form:

$$w_c \leftarrow \frac{1}{n_c} \sum_{i:\, c_i = c} z_i,$$

where n_c is the number of stimulus-present trials in condition c. The bowtie parameters (a, b) were updated by minimizing the posterior-weighted negative log-likelihood of the internal component:

$$(a, b) \leftarrow \arg\min_{a,b} \left[-\sum_i z_i \log\, p_{\text{int}}(r_i \mid t_i; a, b)\right]$$,with constraint $a \geq b$.

Optimization was performed using fminsearch (MATLAB 2020B).

Initialization and stopping

Initial (a, b)values were seeded from the observer's best-fitting bowtie parameters estimated from Experiment 1 precision data when available; otherwise we used the sample median across observers. EM iterations were capped at 200 iterations and terminated early when the absolute change in log-likelihood satisfied:

$$|\ \Delta LL\ | < 10^{-6}\, n,$$

(where n is the number of stimulus-present trials used for fitting.)

Recovered guessing on stimulus-present trials

To reconstruct guessing structure from stimulus-present trials, we computed a posterior-weighted KDE over stimulus-present responses using the posterior probability of guessing $(1-z_i)$as a trial weight:

$$\tilde{q}_{\text{rec}}(r) = \sum_i (1 - z_i) \exp\left[-\frac{1}{2}\left(\frac{\delta(r, r_i)}{h}\right)^2\right], \qquad h = 3°,$$

and normalized it to unit mass on the 1° grid:

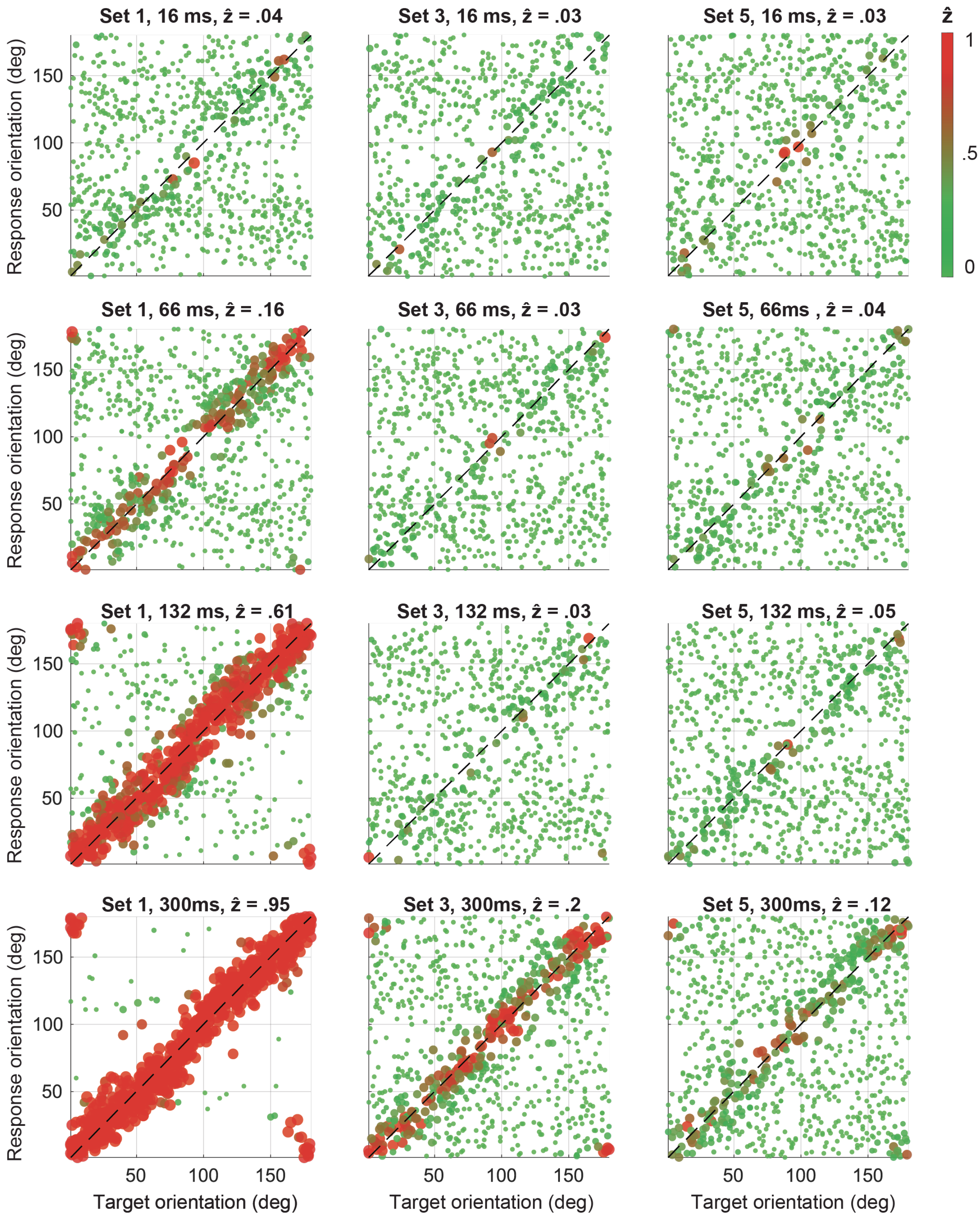
$$q_{\text{rec}}(r) = \frac{\tilde{q}_{\text{rec}}(r)}{\sum_{r'=1}^{180} \tilde{q}_{\text{rec}}(r')}$$

Comparison to a uniform-guess null (posterior recovery control)

Because $q_0(r)$ enters the posterior z_i, posterior-weighted recovery could have, in principle, biased $q_{\text{rec}}(r)$ toward $q_0(r)$. To quantify this bias under a null in which stimulus-present guesses are uniform, we simulated stimulus-present responses for each observer while preserving that observer's fitted internal component and trial/condition structure, but drew guess responses from a distribution $q_{\text{unif}}(r) = 1/180$. We then applied the identical posterior-recovery procedure to obtain $q_{\text{rec,null}}(r)$ and computed its correlation with $q_0(r)$. This simulation was repeated (300 iterations per observer) to obtain an observer-specific null distribution for the recovered-vs-q_0 similarity.

Recovered guess complementarity with independently measured precision

Complementarity between stimulus-present recovered guessing and independently measured precision was assessed within each observer as the Pearson correlation between the smoothed recovered guess density $q_{\text{rec,smooth}}$(from stimulus-present trials) and the smoothed precision landscape P_{smooth}(from Experiment 1 precision trials), using the same exact circular-shift alignment test described above (left-tailed). Group-level summaries used observer bootstrapping for confidence intervals and a one-sided sign-flip test over observer-level correlations.



Supplementary Figure 1. Trial-wise posterior classification (z_i) by condition in Experiment 2. Each panel shows target orientation (x-axis) plotted against response orientation (y-axis) for one set size (columns: 1, 3, 5) and display duration (rows: 16, 66, 132, 300 ms). Points are individual trials, colored by the empirical-guess mixture model's trial-wise posterior probability of the internal (target-driven) component, z_i (color bar). Warmer colors indicate higher z_i (responses more likely generated from the target-centered internal component); cooler colors indicate lower z_i (responses more likely generated from the empirical guess component derived from 0-ms trials). The dashed diagonal indicates perfect target–response correspondence. Panel titles report the mean posterior $\bar{z}$for that condition.

REFERENCES

Adam, K. C., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. Cognitive psychology, 97, 79-97.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the" oblique effect" in man and animals. Psychological bulletin, 78(4), 266.

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.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. Annu. Rev. Psychol., 59(1), 167-192.

Girshick, A. R., Landy, M. S. & Simoncelli, E. P. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. Nat. Neurosci. 14, 926–932 (2011).

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), 5-5.

Harrison, W. J., Bays, P. M., & Rideaux, R. (2023). Neural tuning instantiates prior expectations in the human visual system. bioRxiv, 2023-01.

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. Perception & psychophysics, 63(8), 1421-1455.

Lau, H. (2008). Are we studying consciousness yet. Frontiers of consciousness: Chichele lectures, 2008, 245.

Lieder, F., & Griffiths, T. L. (2017). Strategy selection as rational metareasoning. Psychological review, 124(6), 762.

Luck, S J., and Edward K. Vogel (1997). "The capacity of visual working memory for features and conjunctions." Nature 390.6657 : 279-281.

Morales, J., Chiang, J., & Lau, H. (2015). Controlling for performance capacity confounds in neuroimaging studies of conscious awareness. Neuroscience of Consciousness, 2015(1), niv008.

Morales, J., Odegaard, B., & Maniscalco, B. (2022). The Neural Substrates of Conscious Perception without Performance Confounds. Neuroscience and Philosophy.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. International journal of computer vision, 42, 145-175.

Panichello, M. F., DePasquale, B., Pillow, J. W., & Buschman, T. J. (2019). Error-correcting dynamics in visual working memory. Nature communications, 10(1), 3366.

Patten, M. L., Mannion, D. J., & Clifford, C. W. (2017). Correlates of perceptual orientation biases in human primary visual cortex. Journal of Neuroscience, 37(18), 4744-4750.

Prat-Carrabin, A., Meyniel, F., & da Silveira, R. A. (2024). Resource-rational account of sequential effects in human prediction. Elife, 13, e81256.

Roth, Z. N., Kay, K., & Merriam, E. P. (2022). Natural scene sampling reveals reliable coarse-scale orientation tuning in human V1. Nature communications, 13(1), 6469.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. Proceedings of the National Academy of Sciences, 105(16), 5975-5979.

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. Nature human behaviour, 4(11), 1156-1172.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. Nature neuroscience, 9(4), 578-585.

Van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. ELife, 7, e34963.

Van den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. Proceedings of the National Academy of Sciences, 109(22), 8780-8785.

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.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature, 453(7192), 233-235.

FIGURES

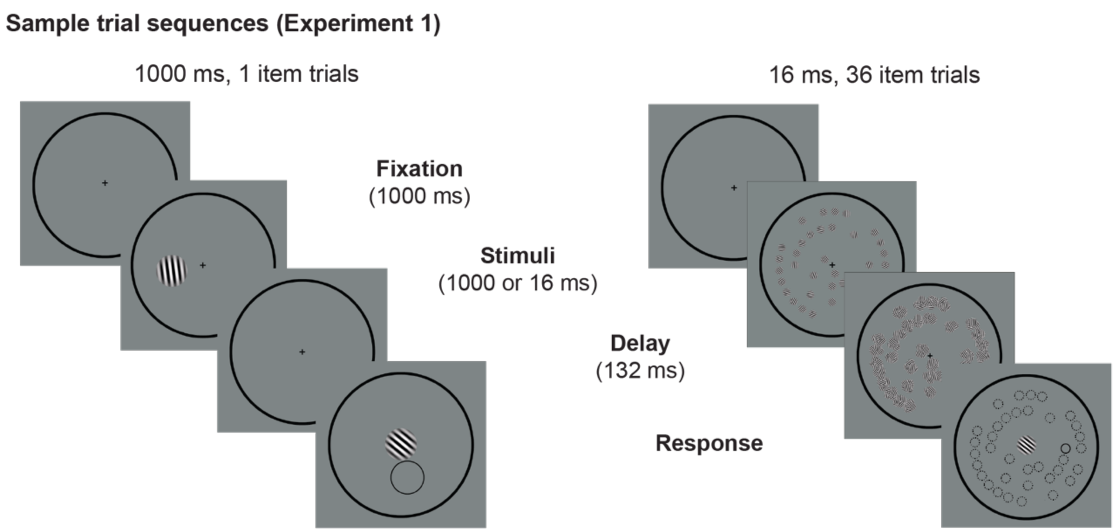


Figure 1. Sample trial sequences for the two trial types used in Experiment 1. Each trial began with a centrally presented fixation cross, which participants clicked to initiate the trial. A circular response ring remained visible throughout. On **precision trials** (left), a single oriented Gabor patch was presented for 1000 ms at one of nine possible peripheral locations. Following stimulus offset, participants adjusted the orientation of a centrally presented response Gabor to match the remembered target orientation. On **guess trials** (right), thirty-six Gabors (one target and 35 distractors) were presented simultaneously for 16 ms, followed by a 132-ms backward mask composed of randomly oriented Gabors. At response, the target location was indicated by a solid outline and distractor locations by dashed outlines. In both trial types, responses were made using a continuous-report procedure, and accuracy was emphasized over response speed.

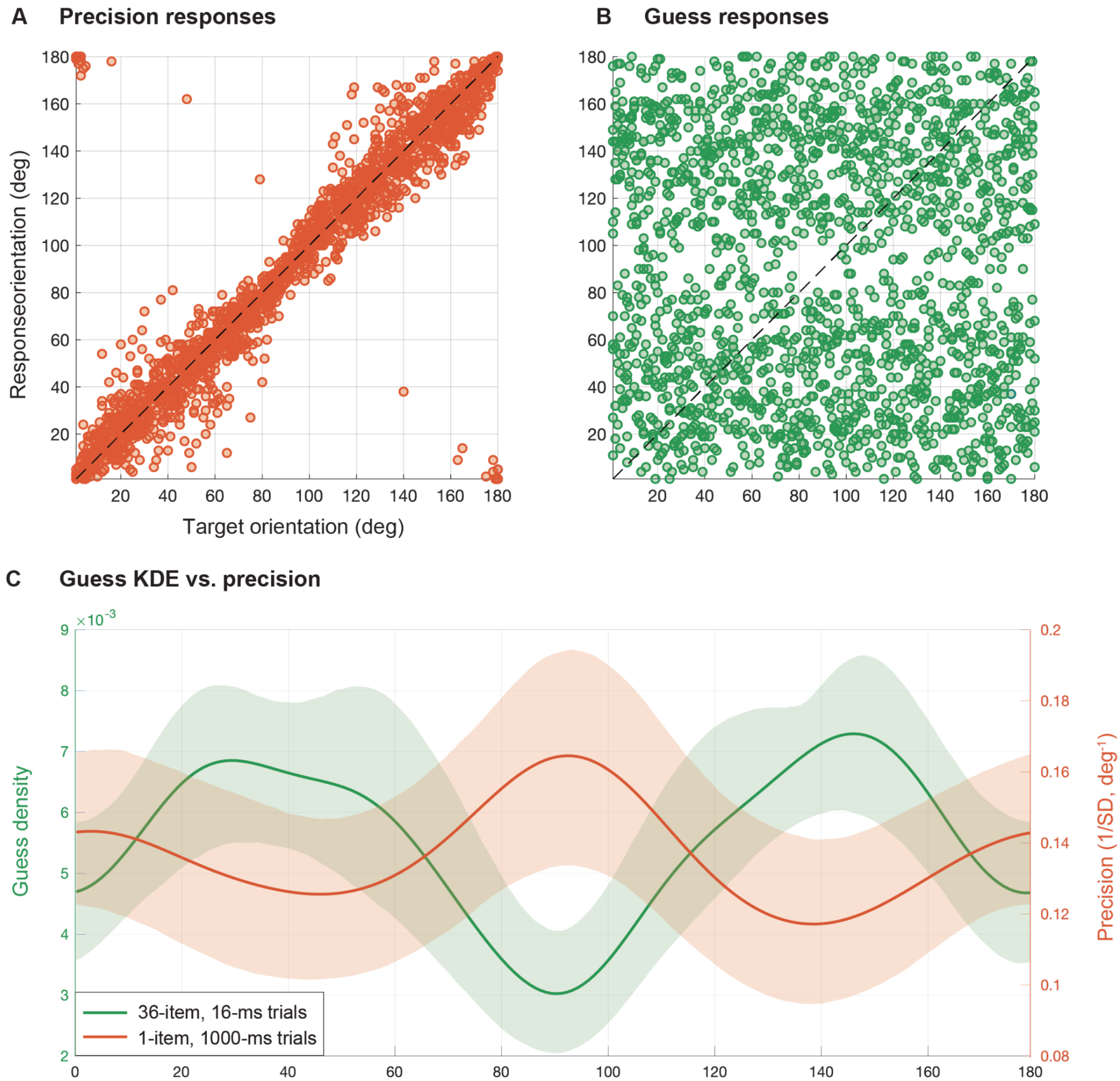


Figure 2. Guesses are non-uniform, forming the complement of precision. (A) Pooled target-response correlations on precision trials (set size 1, 1000ms), showing strong target tuning; dashed unity line indicates perfect performance. **(B)** Subject-wise target-response correlations on guess trials (set size 36, 16ms), showing near-zero tuning in the minimal-information condition. **(C)** Guess response probability (left axis) and precision landscape (right axis; precision = 1/SD of signed error, deg^{-1}) overlaid as a function of absolute orientation. Guess distributions were estimated per observer using circular KDE; precision landscapes were estimated from precision trials using Gaussian weighting over target orientation (see **SI Methods**). Across observers, guess density and precision were reliably anticorrelated (mean $r = -0.335857$; sign-flip p(left) = 0.007805). For visualization, both curves were additionally smoothed with the same circular Gaussian kernel (σ = 10 deg), matching smoothness across curves; shaded bands show 95% bootstrap CIs across observers.

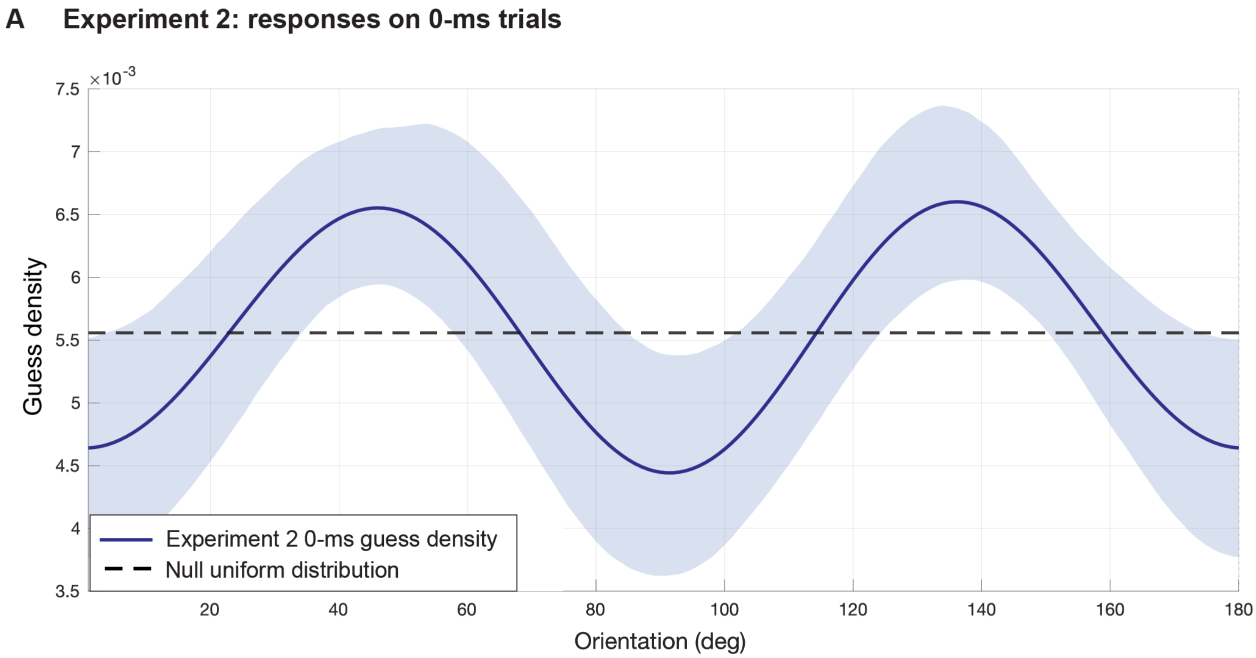
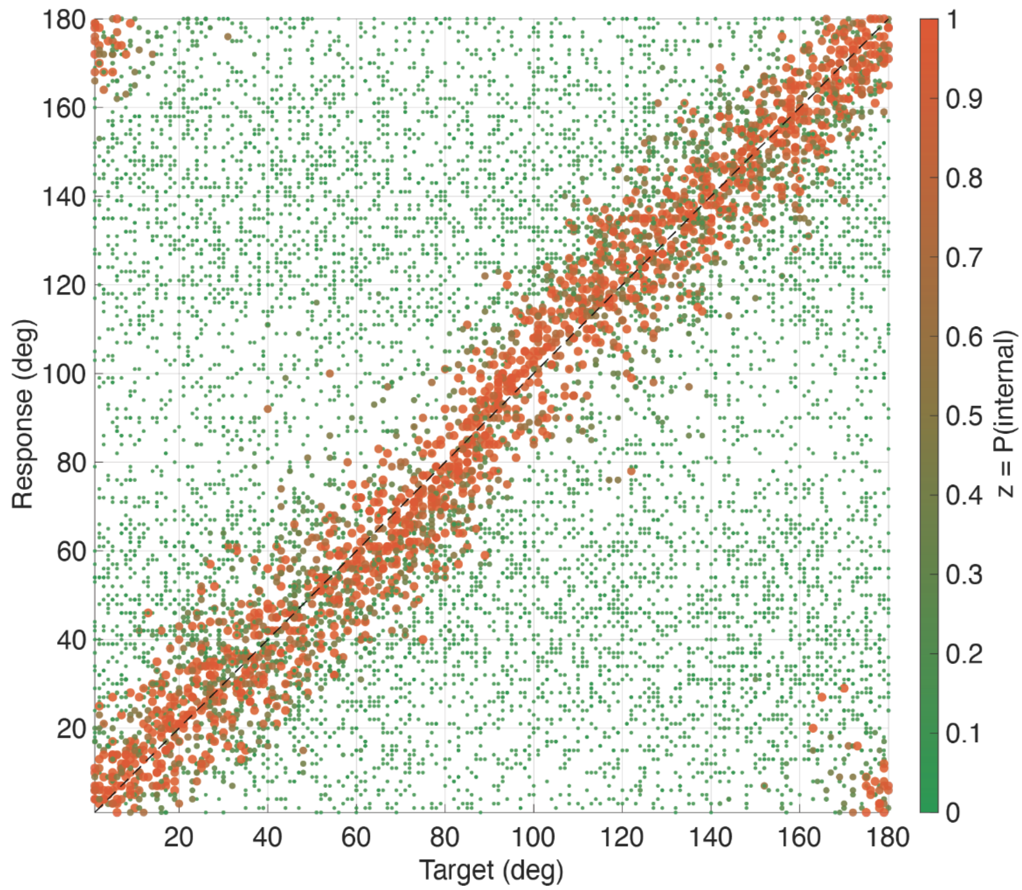


Figure 3. Experiment 2 0-ms trial guess response structure. Responses on 0-ms guess trials were distinctly non-uniform in all subjects. Each observer's 0-ms guess density was positively correlated to that same observer's induced-guess density from Experiment 1 (mean within-observer r = 0.444, 95% CI [0.199, 0.635]; sign test p = 0.0107), consistent with a shared guessing strategy across paradigms.

A Experiment 2: all responses



B Weighted guess density vs. precision

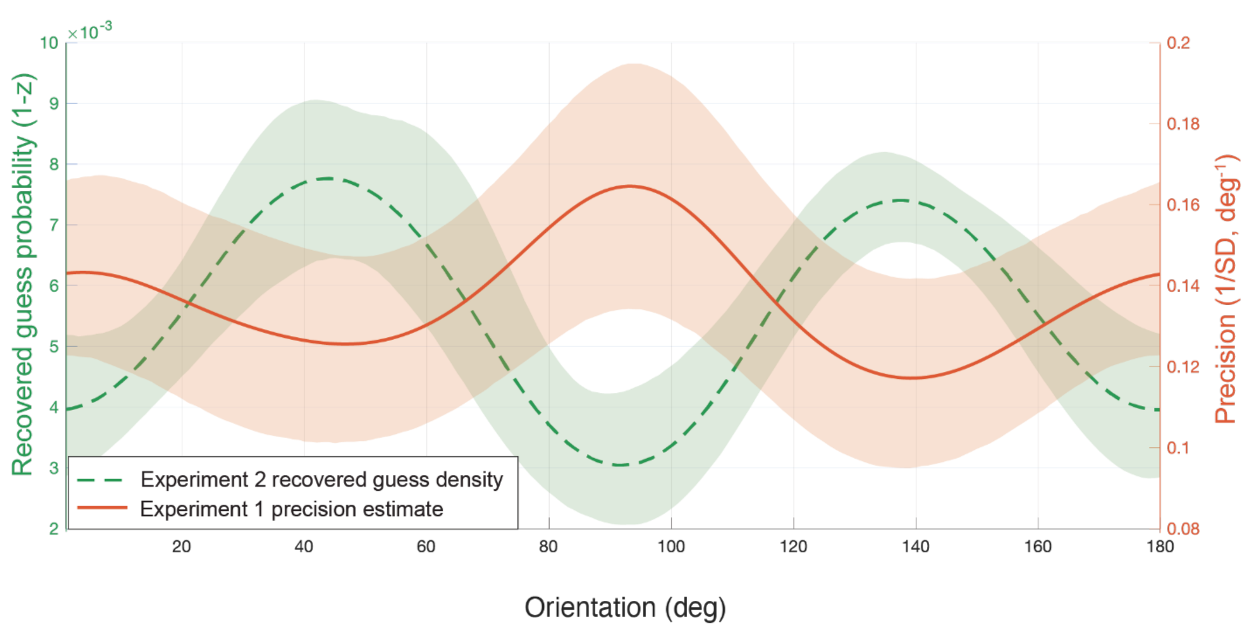
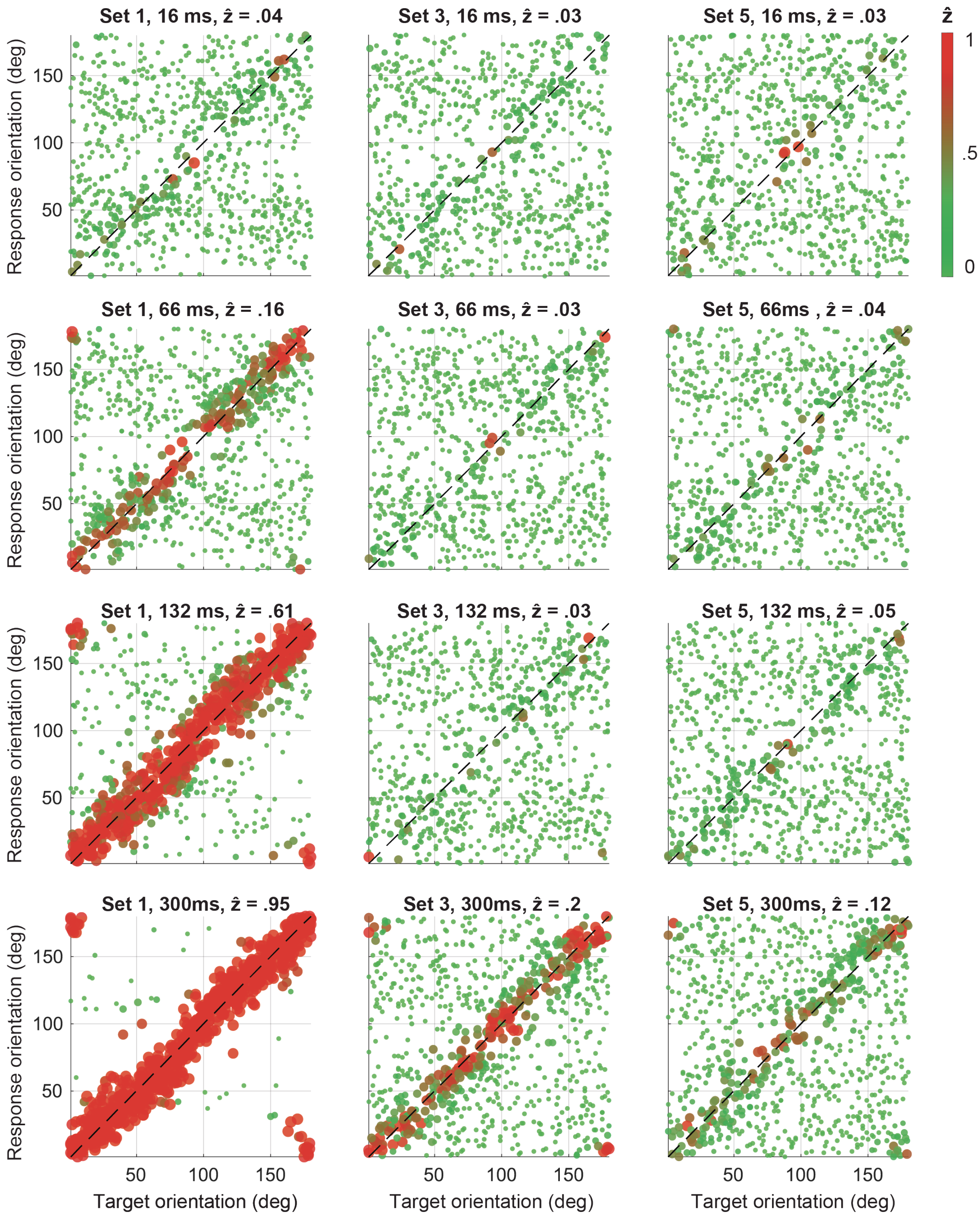


Figure 4. Recovered guess structure on stimulus-present trials tracks the complement of independently measured precision. Dashed curve (left y-axis) shows the group mean *recovered guess density* from Experiment 2 stimulus-present trials, reconstructed by weighting each response by its posterior probability of guessing, $1-z_i$ from the empirical-guess mixture model and estimating a circular KDE over the 180° axial orientation space. The solid curve (right y-axis) shows the group mean *precision landscape* from Experiment 1 precision trials, computed as 1/SD(θ), where SD(θ) is the smoothed target-orientation–dependent standard deviation of signed error.



Supplementary Figure 1.