Attention robustly dissociates objective performance and subjective visibility reports

Karen J. Tian*[1], Brian Maniscalco*[2], Michael L. Epstein[1], Angela Shen[2], Olenka Graham Castaneda[2], Taiga Kurosawa[2], Jennifer A. Motzer[1], Emil Olsson[2], Emily E. Russell[1], Meghan E. Walsh[1], Juneau Wang[1], Tugral Bek Awrang Zeb[2], Richard Brown[3], Victor A. F. Lamme[4], Hakwan Lau[5], Biyu J. He[6], Jan W. Brascamp[7], Ned Block[8,9], David Chalmers[9], Megan A. K. Peters[†2], Rachel N. Denison[†1]

[1] Department of Psychological and Brain Sciences, Boston University
[2] Department of Cognitive Sciences, University of California Irvine
[3] Department of Humanities, LaGuardia Community College; The Graduate Center, City University of New York
[4] Department of Psychology, University of Amsterdam
[5] Department of Biomedical Engineering and Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University; RIKEN Center for Brain Science, Wako, Japan
[6] New York University Grossman School of Medicine
[7] Department of Psychology, Michigan State University
[8] Department of Philosophy and Center for Neural Science, New York University
[9] Department of Psychology, New York University

*co-first authors, [†]co-senior authors

For correspondence:
ktian@bu.edu

Department of Psychological and Brain Sciences
Boston University
111 Cummington Mall
Boston, MA 02215
United States

Keywords: spatial attention, awareness, subjective inflation, visual perception, consciousness

Abstract

Attention generally enhances both visual performance and subjective appearance. Yet, at matched performance, unattended items can appear more visible than attended ones, a phenomenon called “subjective inflation.” Inflation, however, has only been narrowly tested near detection thresholds, making it unclear whether attention regularly dissociates objective and subjective aspects of perception with broad implications for everyday vision—where attention is usually unevenly distributed—and for studies of consciousness. Here, in four experiments, we tested inattentional inflation over varied stimulus and task conditions, spanning threshold to suprathreshold regimes. Using a new analytic approach to relate objective and subjective reports over full psychometric functions, we measured subjective inflation over wide ranges of matched performance. In all experiments, inattention inflated subjective stimulus visibility. But when subjective reports specified visibility of the task-relevant feature, we only found evidence for inflation at threshold. Thus, what we think we see may regularly dissociate from what we can visually discriminate.

Introduction

We generally think that when we attend to something, we will see it better. Indeed, attention enhances basic aspects of vision, including contrast sensitivity[1–5] and spatial resolution[6–8], to benefit performance on many tasks. Attention also enhances subjective appearance[9], making attended items look stronger[10–15], sharper[16,17], and more perceptually organized[18] than unattended ones. However, it has been reported that the effects of attention on objective and subjective aspects of perception do not proceed in lockstep[19,20]. Rather, at matched levels of task performance, unattended items can appear more visible than attended ones[21,22].

Subjective inflation—the idea that our phenomenal experience can be stronger, sharper, more vivid, or otherwise "inflated" above what task accuracy would suggest[23–27]—has been suggested to explain the apparent richness of the unattended periphery[28–30], in spite of its poorer sensory processing[31]. Subjective inflation also suggests that unique mechanisms may underlie objective vs. subjective aspects of perception, and this dissociation has been taken by some as evidence supporting "higher-order" theories of conscious perception. This class of theories proposes that subjective experience arises from downstream metacognitive representations[30,32–37], which can misrepresent the early-stage sensory processes governing visual performance. Against this view—and forming an enduring divide in theories of conscious perception[38,39]—first-order theories assert that sensory signals themselves are sufficient for subjective experience[40–42].

Despite its theoretical implications, empirical evidence for subjective inflation is limited. While there are reports of subjective inflation for unattended (vs. attended locations)[21,22], as well as for peripheral (vs. central) vision[23] and crowded (vs. singleton) conditions[43], these tests have been conducted near detection thresholds and typically at only one or two matched levels of performance. Compelling evidence of subjective inflation specific to inattention ("inattentional inflation") primarily comes from one study[21]. Others are partial, conceptual replications[28]. And yet some other studies have found only weak evidence[22] or even counter-evidence[44], leaving it unclear whether attention regularly dissociates objective and subjective aspects of perception in a way that could matter for everyday vision, where attention tends to be unevenly cast over clearly visible scenes.

The narrow scope of prior tests of subjective inflation was imposed in part by methodological constraints. One strategy to match performance across conditions of comparison is to physically titrate a single pair of stimuli. For example, in tests of inattentional inflation, the stimulus was made physically stronger in the unattended condition and weaker in the attended condition[21,22] with the goal of equating stimulus processing across different levels of attention. However, this approach isolates measurements of inflation to a single performance level, one that may be suboptimal for revealing effects of inflation, and relies on statistical null effects to assume performance equivalence across conditions, such that what looked like inflation may have instead reflected small, non-significant performance differences[45,46]. It is unclear, then, whether the inconsistency in prior tests of inattentional inflation is a consequence of methodological limitations or the fragility of the phenomenon.

To justify subjective inflation as an explanation for the apparent richness of perception across the visual field and as a motivating pillar for higher-order theories of conscious perception[34–36], it should withstand several key tests. First, subjective inflation should be tested beyond the visual detection of faint stimuli to understand whether inflation could underpin everyday, suprathreshold visual experience. Second, most demonstrations of inflation use grating stimuli[21,23,26], leaving it unclear the extent to which inflation generalizes to the perception of visual properties beyond low-level contrast sensitivity. Third, studies of inflation have not asked participants about the visibility of the particular visual feature that governs objective performance (the "task-relevant feature"), leaving open the possibility that previous findings of inflation were driven by objective and subjective reports accessing different stimulus information because the subjective task was underspecified.

Here in four experiments (total n=120), conceived of in an adversarial collaboration[47] including first- and higher-order theorists and led by theory-neutral laboratories, we tested the robustness of inattentional inflation over a large range of stimulus and task conditions in a high-powered experimental design (see **Supplementary Note 1** for further details on the adversarial collaboration). We manipulated covert spatial attention using a central precue and measured its effects on the objective discrimination and subjective perception of peripheral stimuli. Across experiments, the stimuli were of different types (gratings and texture-defined figure-ground stimuli) and strengths, spanning threshold to suprathreshold regimes. We used gratings to facilitate comparison to previous tests of subjective inflation and figure-ground stimuli, which tap into the mid-level visual process of texture segmentation[6,48], to test the phenomenon's generalizability. To index subjective strength in suprathreshold regimes, in which detection measures would not be revealing, we created a comparative visibility task that asked participants to judge the apparent strength of a target relative to a learned reference. The subjective measures probed visibility of the task-relevant feature in addition to visibility of the overall stimulus. To measure subjective inflation over full psychometric functions, we used a new analytic approach[46,49] to quantify subjective reports over common ranges of performance, circumventing statistical and practical pitfalls in matching performance using single stimulus pairs.

In all experiments, we found strong and consistent inattentional inflation of the overall stimulus: withdrawing attention impaired the objective discrimination of visual features more than it reduced reports of apparent stimulus strength. Inflation of the stimulus not only extended beyond threshold regimes, but was more pronounced in suprathreshold regimes. However, when subjective reports specified the task-relevant feature, inflation was only found in threshold regimes. All effects of inattentional inflation were simultaneously replicated at two experimental sites and using two independent analytic pipelines. The results show that attention decouples objective and subjective reports in many, but not all, contexts. Peculiarly, when attention is withdrawn, we report seeing more than visual performance would suggest.

Results

Preregistration and simultaneous replication

All four experiments, including their methods, planned primary analyses, and theory predictions, were preregistered on Open Science Framework (preregistration document[50]: https://osf.io/p3erc; detailed documents: https://osf.io/yur93/). For simultaneous replication, all experiments and analyses were conducted in parallel at two experimental sites. We refrained from sharing data between sites and performing the critical analyses for theory predictions (i.e., quantifying subjective reports as a function of objective performance across varying levels of attention) until both sites completed data collection for an experiment. All main results were consistent across both sites and were confirmed using two independent analysis pipelines. For simplicity, we report the combined data across sites in the main text, with additional site-specific details provided in the Supplementary Material.

Task protocol

In four experiments, human participants (n=30 per experiment with ~3000 trials per participant) performed a spatial attentional cueing task **(Figure 1, Supplementary Figure 1)**. On each trial, participants viewed up to four peripheral targets, which independently varied across 7 strengths, defined separately for two stimulus types. Targets were either contrast-defined gratings in noise (Experiments 1 and 2) or texture-defined figure-ground ovals, with textures composed of line elements (Experiments 3 and 4). Targets were parametrically manipulated in strength—via grating contrast or texture line length—to span near-threshold (Experiments 1 and 3) to suprathreshold (Experiments 2 and 4) regimes **(Figure 1b)**. Stimulus strengths were titrated per participant to be near detection thresholds (see Methods, Thresholding) or fixed across participants to be at suprathreshold values (see Methods, Stimuli). To further increase the grating visibility in the suprathreshold experiments, the noise contrast was decreased from 50% to 20%. The 2x2 design of stimulus type and strength regime, along with presenting 7 levels of stimulus strength within each experiment, equipped us to test for subjective inflation across broad stimulus conditions.

Before the targets appeared, a central precue directed covert attention to one or all target locations while central fixation was monitored **(Figure 1a)**. After the targets disappeared, a response cue indicated which single peripheral location to report. On most trials (60%), the response cue and precue matched (“valid” condition), so participants had incentive to direct attention to the precued location. On some trials (20%), the precue misdirected attention to a location the participant did not have to report (“invalid” condition). When the precue was spatially uninformative (20% of trials, “neutral” condition), participants were asked to distribute attention across all possible target locations.

Participants supplied simultaneously 1) an objective orientation report and 2) a subjective visibility report about the response-cued quadrant **(Figure 1c)**. The objective report was to discriminate the target orientation (i.e., “Was the grating tilted counterclockwise or clockwise from vertical?” or “Was the oval figure oriented vertically or horizontally along its major axis?”). In the threshold experiments, the gratings were orthogonally oriented (±45°) and figure-ground ovals clearly elongated (5° by 3° aspect ratio), so that the difficulty of orientation discrimination was driven largely by target detectability. In the suprathreshold experiments, a stimulus feature (the grating tilt or oval aspect ratio) was titrated per participant to ensure orientation discrimination remained challenging even when the target was easily detectable.

The subjective report was either to detect a near-threshold target (“Did you see or not see a stimulus?”, Experiments 1 and 3) or to compare the strength of a suprathreshold target to a learned reference (“Was the stimulus stronger or weaker than the reference?”, Experiments 2 and 4). We instructed participants to report subjective strength based on how visibly the gratings appeared to stand out from the noise or how visibly the texture-defined figures appeared to “pop-out” from the background, with excerpts of the instructions provided in the Supplementary Methods. In three experiments (1-3), participants additionally specified whether or not they saw the feature relevant for the objective task: the orientation of the grating or oval. Even if a target was not physically presented or consciously detected, participants nonetheless made a forced-choice guess about the orientation, so that every trial provided a concurrent objective and subjective measure, using a single key press to preclude post-choice effects[51,52], except in Experiment 2, in which participants reported the feature visibility judgment using a second key press.

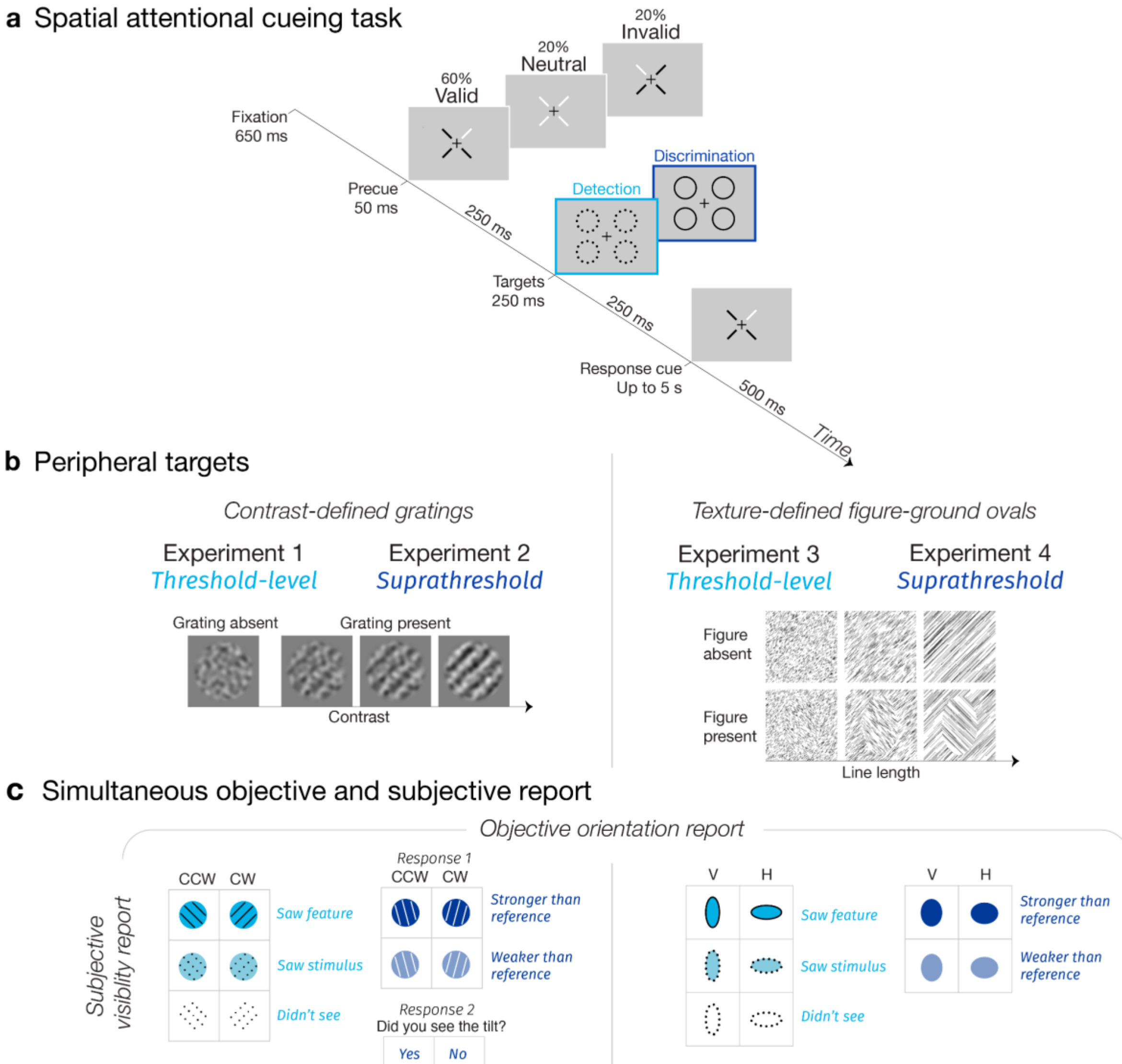


Figure 1. Spatial attentional cueing task. a) Trial timeline. A central precue before the targets directed attention covertly to one or all four target locations. A response cue after the targets indicated which peripheral quadrant to report, which most often matched the precued location (of the 80% of trials in which a single location was precued, the precue was 75% valid). **b)** Targets were either contrast-defined gratings (Experiments 1 and 2) or texture-defined figure-ground ovals (Experiments 3 and 4), which varied independently across 7 strengths in each quadrant. **c)** Participants made an objective orientation report and a subjective visibility report about the response-cued quadrant. The subjective visibility report was either to detect a threshold strength stimulus (Experiments 1 and 3) or to compare the visibility of a suprathreshold stimulus to a learned reference (Experiments 2 and 4). In three experiments (1-3), participants also specified whether or not they subjectively saw the task-relevant feature: the orientation of the grating or oval.

Objective and subjective reports increased with stimulus strength

We first assessed objective performance (**Figure 2a**) and subjective visibility reports (**Figure 2b, 3b**) separately as functions of stimulus strength. In all experiments, objective performance was the proportion correct in identifying the target orientation—p(correct discrimination). Chance performance was 50%. We operationalized subjective visibility as the proportion of trials participants reported seeing the target—p("saw stimulus")—which was presented half of the time in the detection experiments. In the suprathreshold experiments, we indexed subjective strength by asking participants to rate how strong the target appeared relative to a middle reference strength they had learned before the main experiment (see Methods, Reference training) and measured the proportion of trials they reported the test stimulus as stronger than the reference—p("test stronger"). We similarly quantified the proportion of trials in which participants reported seeing the task-relevant feature—p("saw feature") **(Figure 3b)**.

Whereas the contrast dependence of grating discriminability[3,5,53] and visibility[10–14] is well established, the effect of texture line length on the perception of texture-defined figures has not been tested without accompanying luminance confounds[54]. Here we ensured all textures had equal luminance (see Methods, Luminance calibration). Therefore, we first sought to confirm the effects of our stimulus manipulations on objective and subjective reports. We found that all objective **(Figure 2a)** and subjective measures **(Figure 2b, 3b)** increased with stimulus strength for both grating and texture targets (all p<0.001, full ANOVA tables are presented as **Supplementary Tables 1-3**) and spanned a wide range of performance and visibility levels, confirming the efficacy of our stimulus manipulations. Planned pairwise comparisons revealed each successive increment in stimulus strength significantly increased objective and subjective reports across all experiments (all p<0.004 after Holm's correction). Participants reported seeing the task-relevant feature significantly less often than they reported seeing the stimulus (Experiment 1 gratings: (F(1,28)=138.75, p<0.001, η_G^2=0.39; mean difference=26% [25,28]); Experiment 3 textures: (F(1,28)=108.61, p<0.001, η_G^2=0.39; mean difference=26% [24,28]). Thus, subjectively perceiving a stimulus did not guarantee subjectively perceiving its features, even for easily discriminable features (e.g., +45° vs. -45° tilted gratings).

Attention improved objective performance

To characterize the impact of attention on the full psychometric function, we fit Weibull functions to each participant's objective and subjective reports, separately, for each cue validity condition (see Methods, Psychometric function fits). Across all experiments, attention improved orientation discrimination performance (main effect of validity: F(2,222)=892.80, p<0.001, η_G^2=0.52, ε=0.70, **Supplementary Table 1**), consistent with previous reports[3,12,20]. Averaged across stimulus strengths, performance was highest on valid (87% [84,90]), intermediate on neutral (77% [74,79]), and lowest on invalid (66% [63,68]) trials. Reaction times were fastest for valid (0.55 s [0.53,0.58]), intermediate for neutral (0.87 s [0.84,0.90]), and slowest for invalid (0.97 s [0.94,0.99]) trials (F(6,666)=14.56, p<0.001, η_G^2=0.01, ε=0.35), ruling out speed-accuracy tradeoffs as driving performance improvements[55] **(Supplementary Figure 2)**. The benefit of attention on performance was greater at higher stimulus strengths, as revealed by an interaction of stimulus strength and validity (F(12,1332)=53.93, p<0.001, η_G^2=0.08, ε=0.75), consistent with multiplicative response gain[3]. These across-experiment effects of attention were also significant for each experiment individually. In the two detection experiments, attention improved performance even for stimuli reported as unseen[56] **(Supplementary Figure 3)** (main effect of validity: $F(2,112)=5.48$, $p=0.006$, $\eta_G^2=0.03$, ε=0.94, **Supplementary Table 4**), though we cannot be sure participants had no conscious experience of the stimulus for all such trials[57,58].

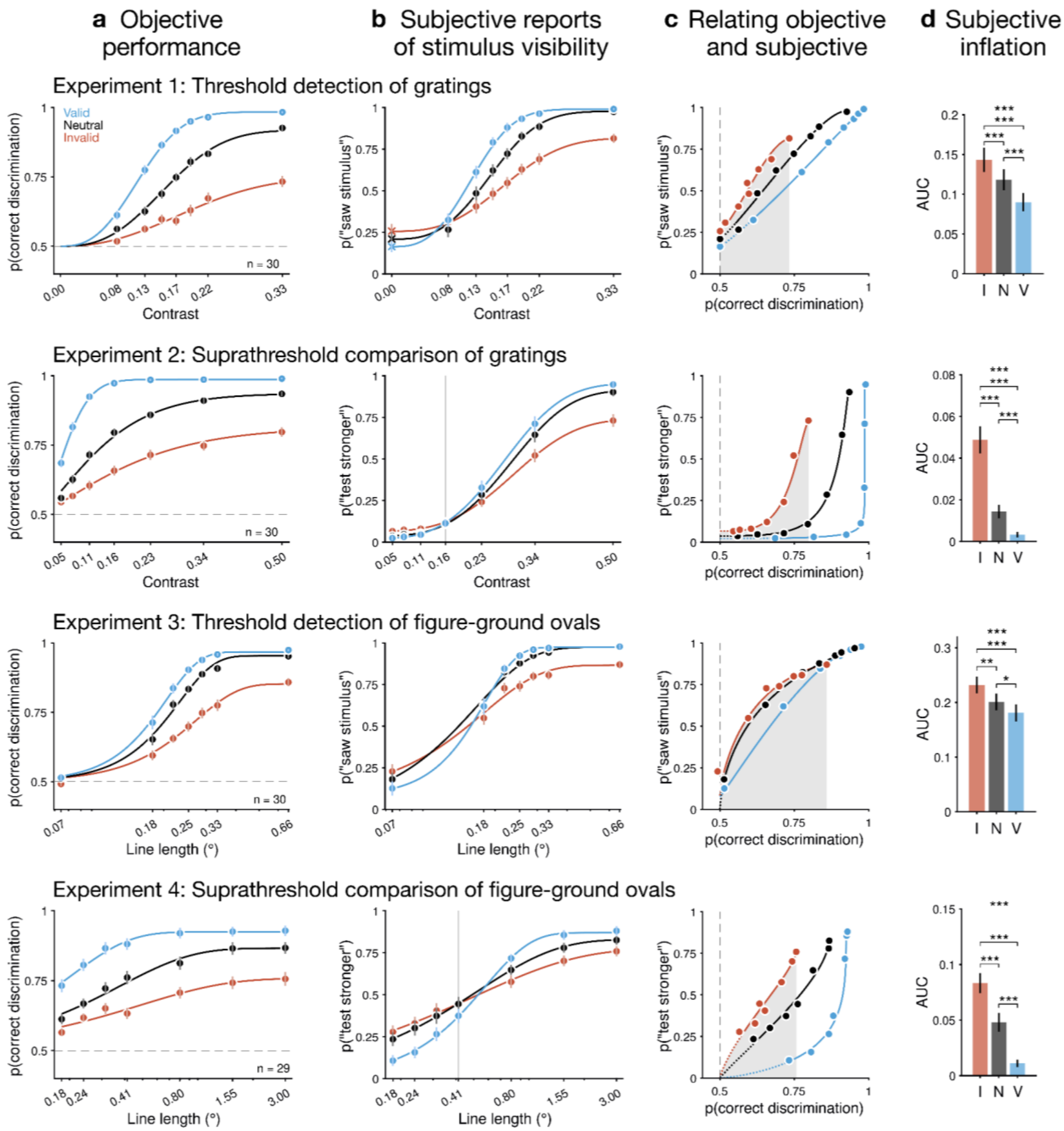


Figure 2. Inattentional inflation of stimulus visibility. In four experiments (one per row), **a)** objective performance increased with stimulus strength and with attention (blue = valid, gray = neutral, red = invalid). Contrast values indicate the grating contrast alone, not including the noise contrast, which was higher (50%) in Experiment 1 and lower (20%) in Experiment 2 to allow the gratings to be more visible. **b)** Subjective reports of stimulus visibility increased with stimulus strength, and attention increased the sensitivity to stimulus strength. Apart from when the grating contrast was zero in Experiment 1 (marked with x's), the data show responses to when the target was present. Target absent data were also collected in Experiment 3, plotted in Supplementary Figure 4. Solid vertical lines mark the reference

strength in the suprathreshold experiments. **c)** Subjective reports increased with objective performance. As a function of performance, subjective reports were higher for unattended than attended stimuli (subjective inflation). Gray-shaded regions mark the shared range of performance across attention conditions. For a range of matched performance, **d**) the area under the relative psychometric function[46] (AUC) was greater for unattended than attended stimuli (repeated measures ANOVA per experiment, all p<0.001). Group means are fit with Weibull functions (a,b) or their corresponding relative psychometric function (c) for visualization. The portion of the fit spanning the range of performance observed between the lowest and highest stimulus strengths is shown as a solid line; the portion extrapolated to chance performance is dotted. **p*<0.05, ***p*<0.01, ****p*<0.001. Data (total n=119; Experiments 1-3 each n=30, Experiment 4 n=29) are presented as mean values ±1 SEM.

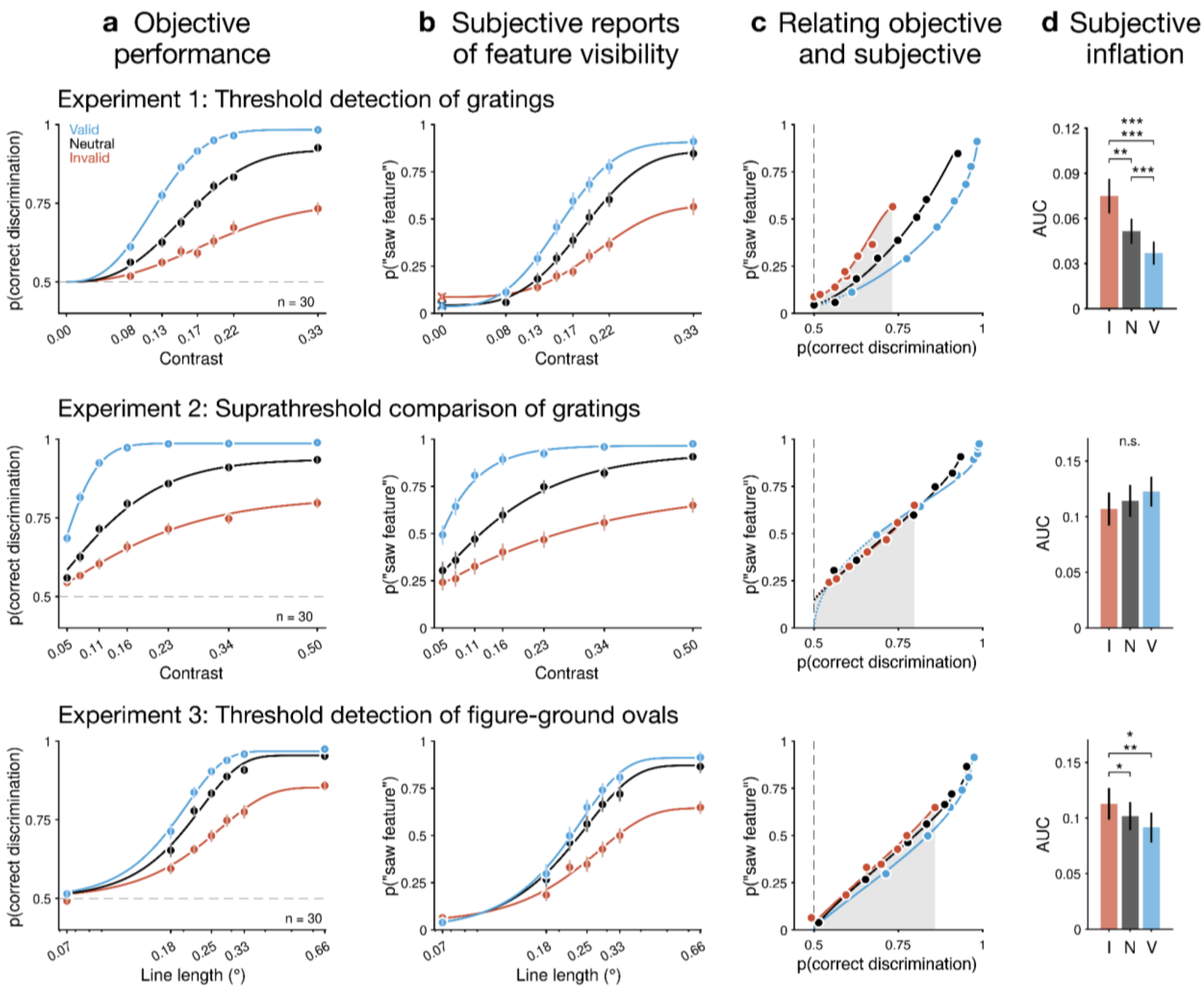


Figure 3. Inattentional inflation of task-relevant feature visibility. In three experiments (one per row), **a)** objective performance increased with stimulus strength and with attention (blue = valid, gray = neutral, red = invalid; same data as Figure 2a). **b)** Subjective reports of seeing the task-relevant feature (i.e., the orientation of the grating or oval) increased with stimulus strength. **c)** Subjective reports as a function of objective performance. For a matched range of performance (shaded gray region), subjective reports of the task-relevant feature were higher under inattention in Experiments 1 and 3. **d)** The AUC was greater for unattended than attended stimuli in threshold (repeated measures ANOVA per Experiment 1 and 3, all *p*<0.011) but not suprathreshold regimes (Experiment 2, non-significant trend in the opposite direction, *p*=0.118), so inattention inflated the task-relevant feature but only in threshold regimes. Group means are fit with Weibull functions (a,b) or their corresponding relative psychometric function (c) for visualization. The portion of the fit spanning the range of performance observed between the lowest and highest

stimulus strengths is shown as a solid line; the portion extrapolated to chance performance is dotted. *$p<0.05$, **$p<0.01$, ***$p<0.001$. Data (total n=90; n=30 per experiment) are presented as mean values ±1 SEM.

Attention increased the sensitivity of subjective reports to stimulus strength

Attention increased subjective reports of stimulus visibility overall **(Figure 2b)** (main effect of validity: $F(2,222)=82.33$, $p<0.001$, $\eta_G^2=0.05$, $\varepsilon=0.70$), which was driven by Experiments 1-3. In Experiment 4, attention primarily increased the sensitivity of subjective strength reports to line length. Attention also increased subjective reports of seeing the task-relevant feature **(Figure 3b)** ($F(2,168)=327.53$, $p<0.001$, $\eta_G^2=0.27$, $\varepsilon=0.63$), with significant effects observed in each experiment that included this measure (all $p<0.001$, **Supplementary Tables 2-3**). The attentional modulation of subjective reports was more pronounced at higher stimulus strengths, for both reports of stimulus visibility (interaction of stimulus strength and validity: $F(12,1332)=67.52$, $p<0.001$, $\eta_G^2=0.05$, $\varepsilon=0.53$) and task-relevant feature visibility ($F(12,1008)=71.71$, $p<0.001$, $\eta_G^2=0.06$, $\varepsilon=0.48$).

In fact, for weak or absent stimuli, inattention tended to elicit stronger subjective reports **(Figure 2b)**. In particular, in the threshold detection tasks, the probability of reporting seeing a stimulus when none was present (i.e., false alarms) was higher under inattention[21,22], for both grating **(Figure 2b,** row 1**)** ($F(2,56)=21.88$, $p<0.001$, $\eta_G^2=0.07$, $\varepsilon=0.58$) and texture stimuli **(Supplementary Figure 4)** ($F(2,56)=9.54$, $p<0.001$, $\eta_G^2=0.03$, $\varepsilon=0.61$). In the suprathreshold comparison tasks, the probability that the test stimulus was reported as appearing stronger than the reference, in cases when the test was in reality weaker, was again higher under inattention, for both grating **(Figure 2b,** row 2**)** ($F(2,56)=11.67$, $p=0.001$, $\eta_G^2=0.04$, $\varepsilon=0.67$) and texture stimuli **(Figure 2b,** row 4**)** ($F(2,54)=30.43$, $p<0.001$, $\eta_G^2=0.11$, $\varepsilon=0.75$). Thus, attention increased the sensitivity of subjective reports of stimulus visibility to the true physical strength of the stimulus, resulting in relatively lower reports for absent or weaker stimuli and higher reports for stronger stimuli when they were attended vs. unattended.

The strength-dependent effects of attention on subjective measures may seem at odds with reports of attention increasing perceived strength across the entire psychometric function[11,13,14,59]; however, methodological differences can reconcile the two patterns. In studies that report wholesale increases in subjective measures with attention, participants judged which of two simultaneous stimuli appeared stronger, so that enhancement of the cued stimulus and suppression of the non-cued stimulus together contributed to the visibility judgment. Using this approach, a stimulus of a given physical strength was more often deemed stronger when cued than non-cued. In our experiments, the visibility judgment involved comparing a single stimulus to an internal criterion for detecting a stimulus (Experiments 1 and 3) or an internal criterion of the reference strength (Experiments 2 and 4).

Apparent effects of attention on criterion setting may be better characterized as sensitivity effects

In a signal detection theory (SDT) framework[60,61], attention consistently improves detection sensitivity[20,21,62] (*d'*)—the ability to determine whether a target is present vs. absent—but its effects on criterion (*c*)—the propensity to report seeing a target—have been mixed[20,21,44]. Therefore we used SDT to assess sensitivity and criterion in the two detection experiments (Experiments 1 and 3).

In Experiment 1, there is one set of target-absent trials (when the grating contrast = 0) and multiple sets of target-present trials (when contrasts > 0). These trials are intermixed and indistinguishable aside from contrast differences, so it is not possible for participants to set more than one criterion per attention condition. Conversely, in Experiment 3, each line length has its own set of target-absent and target-present trials, so separate criteria can, in principle, be set for each line length. Here, to facilitate comparisons across experiments and to previous literature, we chose to analyze all data by estimating different criteria c_{YN} for each stimulus strength level (where c_{YN} specifies the yes-no criterion, commonly referred to simply as *c*, estimated per stimulus strength. See **Supplementary Note 2** for a complete discussion of the modeling and interpretation of criteria in these different experimental designs).

Although attention has previously been found to yield higher c_{YN} values in near-threshold tasks by Rahnev et al.[21], their computational model predicts that this effect should reverse at high enough stimulus strengths (**Supplementary Note 2**). Our findings confirm this prediction **(Figure 4a, Supplementary Figure 5b)** (interaction of validity and stimulus strength: $F(12,672)=61.69$, $p<0.001$, $\eta_G^2=0.04$, ε=0.61). Attention increased c_{YN} at weaker stimulus strengths but its effect gradually reversed, such that attention decreased c_{YN} at higher stimulus strengths. At matched detection sensitivity, the effect of attention was consistently to reduce c_{YN} **(Supplementary Figure 5c)**.

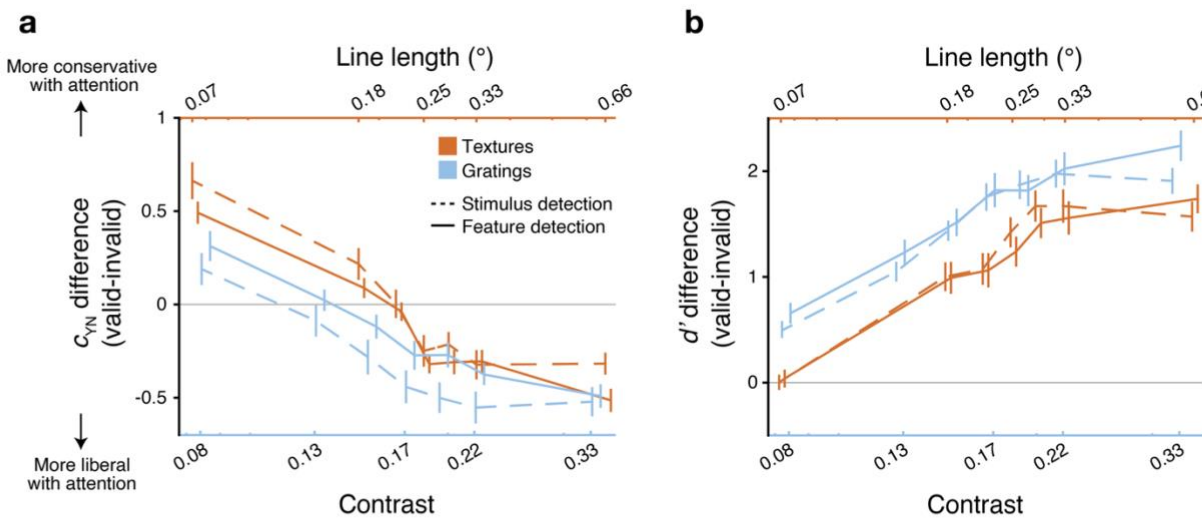


Figure 4. Attentional modulation of signal detection theory measures. a) Here we show the difference in the yes-no criterion c_{YN} between valid and invalid attention trials for detecting gratings (Experiment 1, blue) or texture-defined ovals (Experiment 3, orange) as a function of stimulus strength, which was controlled by either grating contrast (bottom x-axis) or texture line length (top x-axis). Attention made c_{YN} more conservative (positive difference) at lower stimulus strengths, which reversed (negative difference) at higher stimulus strengths (but see **Supplementary Note 2** for why these patterns may be best understood in terms of differences in sensitivity rather than criterion setting; in connection with this, note how the c_{YN} patterns mirror those of *d'* in panel b). **b)** Across both detection experiments, attention improved detection sensitivity (*d'*) and more so at stronger stimulus strengths. The effects of attention on SDT measures were consistent for detecting a stimulus as a whole (dashed lines) and for detecting a particular stimulus feature, the grating or oval orientation (solid lines). Data (total n=60; Experiments 1 and 3 each n=30) are presented as mean values ±1 SEM.

Meanwhile, attention increased detection sensitivity across the entire psychometric function (main effect of validity: $F(2,112)=296.54$; $p<0.001$, $\eta_G^2=0.40$, ε=0.74), and more so at higher stimulus strengths (interaction of stimulus strength and validity: $F(12,672)=56.13$, $p<0.001$, $\eta_G^2=0.10$, ε=0.77) (**Figure 4b, Supplementary Figure 5a**). As detailed in **Supplementary Note 2**, due to Experiment 1's task structure, there can only be one criterion per attention condition, and so c_{YN} values computed for each stimulus strength must reflect changes in sensitivity rather than criterion setting. A similar phenomenon likely occurs in Experiment 3, since the empirical relationship between c_{YN} and *d'* in that experiment closely mirrors the arithmetic relationship of these quantities in Experiment 1 (**Supplementary Figure 5c**). The measurement here of full psychometric functions thus suggests that the apparent effects of attention on criterion setting previously reported[21,22,44] may be attributable to the effects of attention on sensitivity.

Across both detection experiments, the effect of attention on SDT measures for detecting the task-relevant feature behaved the same as its effect on detecting the stimulus as a whole: attention increased detection sensitivity across the entire psychometric function **(Figure 4b, Supplementary Figure 5a)** (main effect of validity: $F(6,336)=397.97$, $p<0.001$, $\eta_G^2=0.64$, ε=0.72) and more so at higher stimulus strengths (interaction of validity and stimulus strength: $F(12,672)=47.89$, $p<0.001$, $\eta_G^2=0.10$, ε=0.74), while making c_{YN} more conservative or liberal, depending on stimulus strength **(Figure 4a, Supplementary Figure 5b)** (interaction of validity and stimulus strength: $F(12,672)=58.27$, $p<0.001$, $\eta_G^2=0.04$, ε=0.66). The effects of attention on the SDT measures were significant for each detection experiment individually and for both subjective reports types (all $p<0.001$, **Supplementary Tables 5-8**).

To confirm that these effects of attention on SDT measures did not depend on any unmet assumptions of equal variance of the internal noise and signal distributions, we also calculated the unequal variance measures: d_a and c_a (see Methods, Signal detection theory). The effects of attention on these measures were consistent with those found for *d'* and c_{YN} (**Supplementary Figure 6, Supplementary Tables 9-12**).

Inattentional inflation of stimulus visibility

After characterizing the effects of attention on objective and subjective measures separately, we turned to our main question: under what conditions, if any, does inattention lead to subjective inflation? To test for inattentional inflation, we determined whether subjective reports were higher for unattended items, when matched in performance to attended ones. For each attention condition, we plotted subjective reports as a function not of stimulus intensity but of discrimination performance. Leveraging the psychometric functions fitted separately to objective and subjective measures, we then constructed a "relative psychometric function" to describe their relation[46,49] (Methods, Relative psychometric function) **(Figure 2c)**.

Across all attention conditions, subjective reports increased monotonically and in most cases nonlinearly with increasing performance (**Figures 2c, 3c**), as has been noted in other studies collecting joint measures[46,57,58,63]. So although increases in performance with stimulus strength were accompanied by increases in visibility, the two changed at different rates as stimulus strength increased across its full range.

Attention changed the relation between objective and subjective reports, as shown by the divergence of the relative psychometric functions across validity conditions in most cases. To quantify the effect of attention on the relative psychometric function, we calculated for each validity condition the area under the curve for the subjective measure (on the y-axis) across a range of performance common to all conditions (on the x-axis) ("AUC," see Methods, Area under the relative psychometric function). A higher AUC indicates subjective reports in one condition were "inflated" over another, when the range of performance across conditions was matched (shaded gray regions in **Figures 2c, 3c**).

Across all experiments, the AUC for stimulus visibility **(Figure 2d)** was strongly and significantly modulated by attention: largest for invalid (0.13 [0.11,0.14]), intermediate for neutral (0.10 [0.08,0.11]), and smallest for valid (0.08 [0.06,0.09]) conditions ($F(2,222)=117.24$, $p<0.001$, $\eta_G^2=0.14$, ε=0.91), revealing robust inattentional inflation across the full relative psychometric function. Inattentional inflation was significant for each experiment individually (all $p<0.001$, **Supplementary Table 13**) and found for most participants **(Supplementary Figure 7)**.

Inattentional inflation of task-relevant feature visibility

To evaluate the possibility that subjective inflation arises from objective and subjective reports accessing different stimulus information, we asked participants to report whether or not they saw the feature relevant for the objective task (the orientation of the grating or oval) in three experiments (1-3). For near-threshold coarse discrimination tasks (as in Experiments 1 and 3), it is usually assumed that discriminating between very different stimulus features (e.g., +45° vs. -45° tilted gratings) relies on the same information as detecting the stimulus as a whole[64]. Perhaps as a result, studies of subjective inflation have never separately assessed inflation for the stimulus (e.g., "Did you see the stimulus?") vs. the task-relevant feature (e.g., "Did you see the stimulus orientation?").

Across all experiments, the AUC for feature visibility was significantly modulated by attention **(Figure 3d)**: largest for invalid (0.10 [0.08,0.11]), intermediate for neutral (0.09 [0.08,0.10]), and smallest for valid (0.08 [0.07,0.10]) conditions ($F(2,168)=6.54$, $p=0.004$, $\eta_G^2<0.01$, ε=0.78, **Supplementary Table 14**). This pattern indicates that even when subjective reports stipulated the task-relevant feature, the relationship to feature discriminability could nonetheless decouple with attention.

However, following a significant effect of experiment ($F(2,84)=7.26$, $p<0.001$, $\eta_G^2=0.13$) and an interaction of experiment and validity ($F(4,168)=7.67$, $p<0.001$, $\eta_G^2=0.02$, ε=0.78) on the "feature-visibility" AUC, we tested inflation of the task-relevant feature within each experiment. Inattention significantly inflated the task-relevant feature in threshold regimes, for both grating (Experiment 1: $F(2,56)=14.63$, $p<0.001$, $\eta_G^2=0.10$, ε=0.62) and texture (Experiment 3: $F(2,56)=5.27$, $p=0.011$, $\eta_G^2=0.02$, ε=0.89) stimuli (**Figure 3d**, top and bottom rows). But we found no evidence for inflation of the task-relevant feature in suprathreshold regimes for grating stimuli (Experiment 2: $F(2,56)=2.22$, $p=0.118$, $\eta_G^2<0.01$, ε=0.64); if anything the pattern went slightly in the opposite direction (**Figure 3d**, middle row). We could not assess inflation for the task-relevant feature in suprathreshold texture stimuli, as feature-specific subjective reports were not collected in Experiment 4. Thus, we only found evidence for inattentional inflation of the task-relevant feature in threshold regimes.

All effects of attention on the stimulus- and feature-level AUC were simultaneously replicated across two experimental sites (no significant interactions of site and validity, all $p>0.071$; **Supplementary Table 13-14**) **(Figure 5)**.

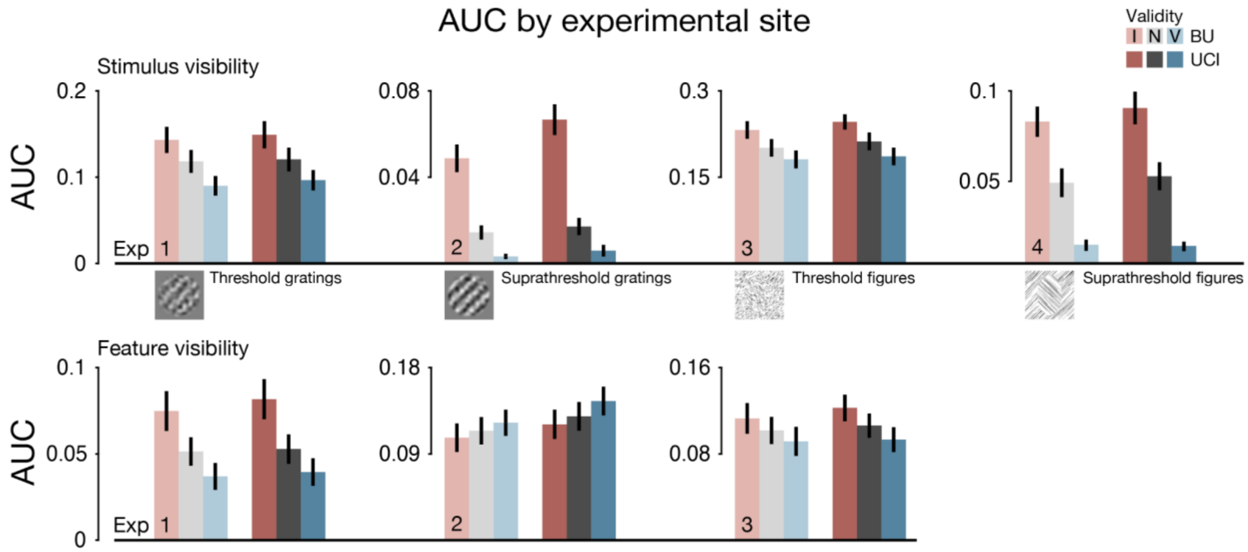


Figure 5. Inflation results replicated across experimental sites. The relative psychometric function AUC, quantifying subjective reports over a matched range of objective performance, at each experimental site (Boston University [BU] vs. University of California Irvine [UCI]). A larger AUC indicates one attention condition (blue = valid, gray = neutral, red = invalid) yielded "inflated" subjective reports over another, for a shared range of performance. Inattention inflated subjective reports of the overall stimulus (top) in all four experiments. Inattention inflated reports of the task-relevant feature (orientation, bottom) in threshold regimes (Experiments 1 and 3) but not suprathreshold regimes (Experiment 2). All effects of attention were simultaneously replicated across the two experimental sites (no significant interaction of site and validity for any experiment and visibility measure, all $p>0.071$). Data (n=15 per site and experiment, except Experiment 4 UCI n=13) are presented as mean values ±1 SEM.

Comparing inattentional inflation across stimulus types, stimulus strength regimes, and visibility measures

Finally, to compare inattentional inflation across different stimulus types, stimulus strength regimes, and visibility measures, we calculated an attentional modulation index (AMI) **(Figure 6)** as the AUC on valid trials subtracted from that on invalid trials, divided by their sum (see Methods, Attentional modulation index). Positive AMI values indicate inattentional inflation, negative values indicate inattentional deflation, and an AMI of zero indicates no dissociation of objective and subjective reports with attention. In all experiments, the AMI for stimulus visibility was significantly above zero ($F(1,110)=517.17$, $p<0.001$, $\eta_G^2=0.83$; mean AMI=0.51 [0.44, 0.58]), confirming strong and consistent inattentional inflation of the overall stimulus.

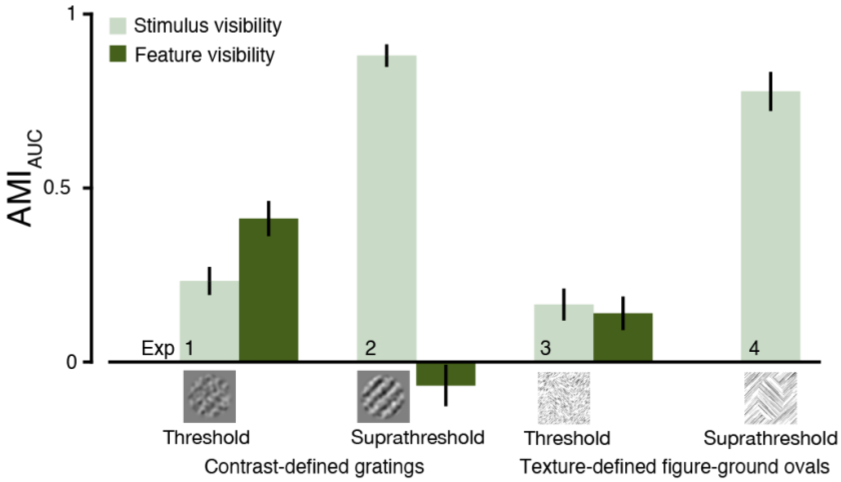


Figure 6. Comparing inattentional inflation across experiments. The AMI, quantifying the degree and direction with which objective and subjective reports dissociate with attention, was significantly greater than zero for stimulus visibility (light green) in all experiments (p<0.001), indicating robust and widespread inflation under inattention. Inflation of the overall stimulus visibility was more pronounced in suprathreshold (Experiments 2 and 4) than threshold-strength (Experiments 1 and 3) regimes (p<0.001). Inattention also inflated visibility of the task-relevant feature (dark green) but only in threshold regimes. Data (total n=118; Experiments 1-3 each n=30, Experiment 4 n=28) are presented as mean values ±1 SEM.

Comparing the AMIs for suprathreshold (Experiments 2 and 4) vs. threshold-strength (Experiments 1 and 3) stimuli revealed that stimulus-level inattentional inflation not only occurred beyond threshold vision, but was significantly more pronounced in suprathreshold regimes (main effect of strength regime: $F(1,110)$=197.65, p<0.001, η_G^2=0.64; suprathreshold AMI 0.83 [0.77, 0.90] > threshold AMI 0.20 [0.14, 0.26], **Supplementary Table 15**). This pattern was similar for grating and figure-ground stimuli (no significant effect of stimulus type: $F(1,110)$=3.58, p=0.061, η_G^2=0.03). Although the decoupling between objective and subjective reports was statistically more pronounced for the suprathreshold experiments, these results do not necessarily imply that attention had a larger effect on performance-matched visual experience at suprathreshold, given that participants were asked to make qualitatively different kinds of subjective judgments about their visual experience in the tasks at threshold vs. suprathreshold.

For near-threshold figure-ground stimuli (Experiment 3), the magnitude of inflation was similar for the overall stimulus and the task-relevant feature (no effect of visibility measure: p=0.707). But interestingly, for near-threshold grating stimuli (Experiment 1), inflation of feature visibility was stronger than inflation of stimulus visibility (main effect of visibility measure: $F(1,56)$=7.51, p=0.008, η_G^2=0.12; AMI feature=0.41 [0.31, 0.51], AMI stimulus=0.23 [0.15, 0.31]).

All measures of AMI were simultaneously replicated at two experimental sites (no significant effect of site, all p>0.459, Supplementary Table 15). Moreover, to ensure that AMI analysis results were replicable, separate analysis pipelines that computed the AMI from raw data were developed independently at each site. The sites communicated to align their analytic approaches conceptually but did not share code. There was no effect of analytic pipeline on AMI measures (all p>0.657, **Supplementary Table 16**) **(Supplementary Figure 8).** Therefore, the results were robust to differences in experimental setups and samples across sites and to any idiosyncrasies in analytic choices across pipelines.

Discussion

In four experiments manipulating voluntary spatial attention, we tested claims that the unattended visual periphery can appear “subjectively inflated” relative to objective performance. We put the phenomenon to several key tests. Inattentional inflation was examined in threshold to suprathreshold regimes, using different stimulus types, and over wide ranges of performance. When the subjective report concerned the overall stimulus, inattentional inflation withstood all tests: covertly unattended items were reported as appearing stronger and more visible than attended ones, over matched ranges of performance. However, when the subjective report specified the visibility of the feature relevant for the objective task, we found that inattentional inflation only survived tests in threshold regimes. The results indicate attention regularly, though not invariably, dissociates objective and subjective aspects of perception, thus preserving inattentional inflation as a possible explanation of the apparent richness of the unattended periphery and as a motivating observation for higher-order theories of consciousness.

Using a high-powered experimental design and improved methodology, we overcame limitations of previous studies to provide strong evidence for both the existence and extent of inattentional inflation. While subjective inflation has been an influential concept in the field[24,25,27–30,33,39,65–75] it has been based on relatively limited empirical evidence, even in a threshold regime[21,23,43]. Beyond threshold, tests of subjective inflation have not been reported, leaving it unknown whether inflation could affect the perception of clearly visible stimuli typical of normal viewing conditions.

We tested the robustness and generality of inattentional inflation in several ways. First, we used an analytic approach recently developed by our group[46] to relate objective and subjective reports over full psychometric functions. This strategy allowed us to measure subjective reports over large matched ranges of performance across conditions, which was not possible using previous “performance matching” approaches[21,22] that restricted tests of inflation to single performance levels. Second, we developed an approach to test inflation beyond threshold regimes, by having participants compare the subjective strength of a clearly visible target to a reference, while discriminating a target feature. Together, these methodological advances enabled tests of inattentional inflation across a full range of performance levels and stimulus strengths, revealing that inflation is strong, widespread, and replicable, though not without limits.

Our findings demonstrate that inattentional inflation generalizes to stimuli beyond those dependent on low-level visual properties, like contrast[21,23] and color[22] sensitivity. When the peripheral targets were texture-defined figure-ground stimuli, inattention inflated how strongly the figures appeared to “pop-out” from the background, showing that inflation can occur for stimuli defined by the mid-level visual property of texture segmentation. Inflation of these figures behaved similarly to that of simple gratings, indicating the phenomenon can operate at multiple levels of processing and may be commonplace in everyday vision. Moreover, figure-ground stimuli generate visual cortical recurrent processing[54,76–79], a candidate first-order substrate of conscious vision[40,41]. Assessing subjective experience using these particular stimuli can thus help arbitrate and refine theories of conscious perception.

Stimulus-level inflation also generalized to suprathreshold regimes. Indeed, of the conditions we tested, subjective reports of suprathreshold stimulus visibility exhibited the largest magnitude of inflation. One author's experience was that inflation effects are obvious at the single-trial level in the suprathreshold task, prompting an alternative framing of inflation that may inform future research (**Supplementary Note 3**). The striking suprathreshold inflation of perceived stimulus strength may contribute to the impression of a rich, intact visual world that extends across the visual field and beyond the focus of attention[28,29,33,66,80]. If the subjective strength of unattended items can far exceed what might be predicted from feature discriminability in suprathreshold scenarios, this discrepancy may help explain the sense of surprise in inattentional[81] and change[82] blindness demonstrations, in which people fail to notice a salient stimulus or stimulus alteration but feel certain they should have, given their subjective impressions. It may suggest that unless explicitly instructed, people make such metacognitive assessments based on overall stimulus visibility rather than the visibility of a particular stimulus feature.

If objective and subjective reports derive from different stimulus features[83,84], their dissociation can be straightforwardly explained. For example, motion in the periphery could signal someone’s approach without indicating their identity, and such different information sources could be differentially affected by attention[53]. A strong test of inflation should therefore constrain subjective reports to the stimulus feature relevant for the objective task. But previous studies have not attempted to do so. When we asked participants to report the visibility of the feature discriminated in the objective task, subjective reports were still inflated under inattention in threshold regimes. However, in the suprathreshold task where these reports were available (Experiment 2 with grating stimuli), we did not find inflation of the task-relevant feature; instead, the relative psychometric functions were statistically indistinguishable across attention conditions. We do note that these were the only reports to be collected with a second keypress, so it is possible that post-decisional effects or extra cognitive load uniquely affected these subjective judgments in a way that obscured potential inflation effects. Thus, while we found that inattention does not inflate feature visibility in some suprathreshold scenarios, feature-level inflation was robust and consistent near the detection threshold.

A prominent model of inattentional inflation at threshold is based on signal detection theory[21]. According to this model, separable components of the internal signal govern subjective reports and objective performance, allowing their dissociation. A stimulus is reported to be visible when its internal signal magnitude exceeds a threshold, whereas its discriminability depends on the signal-to-noise ratio. Inattention is hypothesized to decrease the response magnitude but increase the variability of the internal signal[85], which can increase the likelihood of crossing a fixed detection threshold. Studies have shown dissociable effects of neural variability[86,87] on objective and subjective aspects of perception, lending empirical support to the model. Its dissociable computational mechanisms may map onto dissociable aspects of subjective perceptual experience (**Supplementary Note 3**).

This signal detection model[21] can reproduce several key phenomena in our data regarding how perceptual metrics change for unattended stimuli relative to attended stimuli in detection tasks: false alarm rates increase; both subjective reports and objective performance become less sensitive to stimulus strength; and at matched performance, stimuli are reported as more visible. Additionally, the model can account for the yes-no criterion being more conservative under attention at low stimulus strengths, and it predicts that the criterion becomes more liberal under attention at high stimulus strengths. (See **Supplementary Note 2** for further elaboration of this model prediction, and for an argument that this yes-no criterion effect may be better understood not as an effect on criterion setting *per se*, but rather as an indirect reflection of more primary sensitivity effects.) Here, by measuring the full psychometric function, we not only confirm this prediction, but also provide a unifying explanation for apparent contradictions previously reported in the literature, wherein attention to peripheral stimuli has been variously found to correspond to conservative[21,22,26], neutral[20], or liberal[44] shifts in the apparent detection criterion, depending on the task design and stimuli chosen. Although the model of Rahnev et al.[21] qualitatively accords with several of our findings in the threshold regime, further work is required to extend this model and other contending models to determine their ability to capture feature-specific reports and suprathreshold regimes, and to quantitatively fit the entire dataset.

According to higher-order theories of conscious perception, subjective inflation arises from an overestimation of the strength of the sensory signals governing performance, driven by a putative implicit metacognitive mechanism[30]. By proposing two levels of representation, higher-order views can account for the decoupling between objective and subjective aspects of perception. But these views do not immediately explain why or under what circumstances they will come apart. In the case of attention, higher-order views do not in general explain why withdrawing attention should inflate subjective strength, when the first-order states are presumably comparable, yielding matched performance. That said, one proposal posits a higher-order Bayesian observer that misrepresents the distribution of sensory noise in the periphery or under inattention, leading to subjective inflation[67,73].

Some proponents of higher-order theories have interpreted the fixed threshold of visibility in the signal detection theory model[21] as a higher-order mechanism, likely carried out in prefrontal areas[88]. In this view, the higher-order areas gate the entry into phenomenal consciousness, and the fixed criterion reflects a stable higher-order threshold for conscious perception[32,33]. However, the signal detection model is also compatible with some versions of a first-order view holding that the sensory signal is sufficient to generate conscious perception. The criterion could then be interpreted in different ways consistent with some versions of a first-order view[71]. For example, the criterion could reflect the sensory signals themselves exceeding some threshold that is not set by any higher-order mechanism. Inflation could then arise from attention changing the sensory signals, as in the signal detection model of Rahnev et al.[21] Or, the criterion could be decisional rather than perceptual, governing whether the participant reports seeing the stimulus rather than whether they actually see it[69,89,90], although the resistance of subjective inflation to feedback and reward gives some reason to think the phenomenon is perceptual[21,23]. Overall, while failure to find inattentional inflation would have challenged a motivation of higher-order theories—and the predictions made by higher-order theorists in the current adversarial collaboration—the behavioral demonstration of inflation is not on its own necessarily incompatible with first-order theories (**Supplementary Note 1**).

Importantly, given the clear data obtained here, any theory of subjective awareness must now take seriously the phenomenon of inattentional inflation. First-order theories must account for the widespread occurrence of inflation, while higher-order theories must account for the conditions under which it is absent. The current dataset, testing large numbers of experimental conditions with high trial counts, provides by far the most comprehensive extant data on subjective inflation, and theories of awareness should seek to explain the detailed pattern of these data. The measurement of full psychometric functions is likely to strongly constrain model fits, allowing model identification that would not be possible with single pairs of performance-matched stimuli. The current dataset, which we have documented and made publicly available to invite model fitting by the wider community, can serve as a benchmark dataset for the field.

Going forward, inattentional inflation presents an opportunity to identify the neural processes specific to subjective perception. Typically, the neural signals that track subjective strength tightly correlate with stimulus processing that supports objective performance, confounding their interpretation[45,91,92]. But by reliably dissociating objective and subjective aspects of perception, inattentional inflation may be a valuable approach to isolate the neural signals that uniquely covary with subjective strength while objective performance is controlled. Different theories of conscious perception make different predictions about which neural processes will correlate specifically with subjective strength (e.g., first-order processes in sensory areas[42,90], higher-order processes in prefrontal areas[30], global ignition[93]; for reviews[38,39,94]), making inattentional inflation a promising tool—along with other methods for decoupling objective and subjective aspects of perception[58,86,87,95–104] to adjudicate among these theories.

The idea that our introspection can be but a dubious authority on our own visual performance has long been recognized. Clinical cases in which visual cortical damage leads to ignorance of residual capacity (blindsight)[105] or incapacity (Anton's syndrome)[106] have shown powerful, sometimes permanent dissociations between subjective reports of visual awareness and the objective ability to discriminate visual features. Our findings show that even in healthy observers, following temporary, voluntary fluctuations of visuospatial attention, objective and subjective aspects of perception routinely come apart. What we think we can see therefore may not accurately reflect how well we can distinguish visual features, particularly at the threshold of vision.

Methods

Participants

One hundred twenty healthy adult humans (88 females and 32 males, ages 19-36, based on self-report) participated across four experiments, which was the preregistered target sample size. Fifteen participants per experiment participated at each of two research sites: Boston University (BU) and the University of California, Irvine (UCI). Each participant completed an average of ~3400 trials (range of 2288-4771 trials) across four to six 1.5-hour-long visits on separate days, for both a highly powered sample size and reliable measurements for each participant. Authors A.S., T.K., J.A.M., E.O., E.E.R., M.E.W., and J.W. participated in the experiments. All other participants were naive to the study design. All participants provided informed consent, and the University Committee on Activities Involving Human Subjects at BU and the Institutional Review Board at UCI approved the experimental protocols. All participants had normal or corrected-to-normal vision and were monetarily compensated for their time. No sex or gender-based analyses were performed, and we did not consider sex or gender in the study design, as neither sex nor gender played a role in our research questions.

Procedure

Participants were seated in a dark room 75 cm from a computer monitor. Stimuli were generated on Linux at BU and Windows at UCI using MATLAB and Psychophysics Toolbox[107–109]. Stimuli were displayed on a VIEWPixx LCD monitor (VPixx Technologies Inc., QC, Canada) with a resolution of 1920 x 1080 pixels and a refresh rate of 120 Hz at BU and a CRT monitor (NEC MultiSync FE2111SB) with a resolution of 1280 x 1024 pixels and a refresh rate of 60 Hz at UCI. To linearize the contrasts, the displays were calibrated using a Konica Minolta LS-100 Luminance Meter (Konica Minolta, Tokyo, Japan). Participants had their head position stabilized in a chin-and-head rest. Gaze position was continuously recorded at a sampling frequency of

1000 Hz using an EyeLink 1000 (SR Research Ltd., ON, Canada) at BU and 500 Hz using a LiveTrack Lightning eye-tracker (Cambridge Research Systems, Ltd., UK) at UCI.

Task

In all four experiments, participants performed a spatial attentional cueing task while fixating on a central cross **(Figure 1, Supplementary Figure 1)**. The screen was divided into four quadrants. On each trial, participants viewed up to four peripheral targets, which varied independently across seven strengths. Only one quadrant was relevant for the report, as indicated by a post-stimulus response cue. Before the targets, one or all four arms of a black central precue, each 1° long by 0.1° wide, flashed white (50 ms) to direct covert attention focally to one quadrant (80% of trials) or in a distributed fashion across all quadrants (20% of trials; neutral attention condition). The focal precue matched the response cue with 75% validity, to incentivize using the spatial information provided by the precue, leading to 60% valid and 20% invalid attention trials overall. The precue appeared 300 ms before the targets, to allow for the deployment of covert attention to the cued location[6,12]. The targets were presented for 250 ms. A response cue 500 ms after target onset instructed participants to make both a forced-choice objective orientation report and a subjective visibility report about the response-cued quadrant **(Figure 1c)**.

The subjective report was either to indicate visibility of a near-threshold target (Experiments 1 and 3) or to compare the strength of a suprathreshold target to that of a learned reference (Experiments 2 and 4). In experiments 1-3, participants also reported whether or not they saw the feature relevant for the objective task: the stimulus orientation. All objective and subjective reports were made using a single keypress, except the feature visibility report in Experiment 2, which was made using a second keypress. For the subjective reports, participants were instructed to report how the stimuli appeared to them and not, for example, their confidence in the orientation judgment or what they considered the stimulus contingencies to be. See Supplementary Methods for task instruction excerpts.

When a response was registered, the fixation cross lightened to gray for 500 ms before the next trial was initiated. Otherwise, no feedback was provided. If no valid keypress was registered within 5 s following the response cue, the trial was aborted and not repeated. Participants rarely failed to make a valid keypress within the response window (on average 2.9 trials, SD=5.9). Between trials, a gray screen containing only the fixation cross appeared for 500 ms. The gray was a mid-gray (45.6 cd/m^2 at BU, 60.7 cd/m^2 at UCI) in the experiments with grating stimuli. In Experiment 3 with texture stimuli, 7 participants had the same mid-gray blank screen luminance. For the remaining participants in Experiments 3 and 4, the gray was lightened (78.7 cd/m^2 at BU, 93.2 cd/m^2 at UCI) to match the average texture luminance, with the goal of preventing eye-tracking issues due to luminance changes between trial stages.

Trials were grouped into consecutive runs of 560 trials. Within each run, precue validity and the location of the response-cued quadrant were counterbalanced. For each permutation of precue validity and response-cue location, stimulus strength and identity at the response-cued location were also counterbalanced. The presentation order of these counterbalanced conditions was pseudo-randomized within each run. On invalid trials, the location of the precued quadrant was randomly selected as one of the three locations not probed by the response cue. The stimulus properties at the non-response-cued quadrants were pseudo-counterbalanced, so that their marginal probabilities were controlled for within a run. Breaks were offered after every block of

112 trials, for 5 blocks per run. The attention manipulation and target timings were identical across experiments. However, the stimuli themselves and the nature of the task report differed from experiment to experiment.

Stimuli

Visual targets were contrast-defined gratings embedded in noise (Experiments 1 and 2) or texture-defined figure-ground ovals (Experiments 3 and 4). Targets were present either half the time at threshold strengths (Experiments 1 and 3) or all the time at suprathreshold strengths (Experiments 2 and 4). The experiments are reported in a different order than they were collected; the order of data collection adhered to the order in the preregistration: 1) threshold detection of figure-ground ovals (Experiment 3), 2) suprathreshold comparison of figure-ground ovals (Experiment 4), 3) threshold detection of gratings (Experiment 1), and 4) suprathreshold comparison of gratings (Experiment 2).

Contrast-defined gratings. Gratings were luminance-modulated sinusoids with a spatial frequency of 1 cycle per degree. Gratings were centered at 5° eccentricity. In Experiment 1, the gratings were presented at low contrasts, calibrated per participant to be near threshold visibility when added to noise pedestals, and oriented ±45° from vertical. In Experiment 2, the gratings were presented at suprathreshold contrasts—fixed across participants from 5% to 50% in 7 log steps—and oriented about vertical at individual tilt thresholds (see Methods, Thresholding).

Noise pedestals. Gratings were added pixelwise to noise pedestals, then placed in a circular aperture 5° in diameter with a cosine edge subtending 0.5° that gradually faded to the background gray. To generate each noise patch, Gaussian noise was bandpass filtered around the grating spatial frequency ±1 octave. The filtered noise was then centered at mid-gray (matching the background luminance) and scaled to the desired contrast, which was 50% in the Experiment 1 and lowered to 20% in Experiment 2 to allow the superimposed gratings to have higher signal to noise ratio, thus increasing their visibility. Noise patches that by chance deviated from the desired mean contrast by more than 2% were regenerated.

Texture-defined figure-ground ovals. The texture "background" was made of parallel lines, oriented either ±45°, on which an oval "figure" delineated by orthogonal lines could appear (**Supplementary Figure 1**). Texture lines were black on a white background. The entire screen (21.6° by 38.4° at BU, 23.8° by 29.8° at UCI) was filled with textures, with the exception of a gray circle (2° radius) at the center of the screen containing the fixation cross and cues (see Methods, Online fixation monitoring). Each quadrant of the screen (5.4° by 9.6° at BU, 6.0° by 7.4° at UCI) was filled with a unique background texture, made of many randomly placed lines of uniform length and orientation. On a given trial, the background texture lines across quadrants were of the same orientation but could vary in length. Texture lines were truncated at the quadrant boundaries and the perimeter of the central circle so that no lines crossed these boundaries. The number of lines drawn scaled with line length, so that the proportion of black pixels (which controlled the mean luminance of the display) was fixed to 0.15 across all textures (see Methods, Luminance calibration).

The oval "figure" was centered at 8° eccentricity within a quadrant and oriented either vertically or horizontally along the major axis. Each oval was filled with a unique texture of many randomly placed lines, which were of the same length as the background texture lines but orthogonal in orientation. We ensured the ovals had buffers of at least 3.9° (at BU) or 1.7° (at UCI) between their perimeter and the edge of the screen, to minimize edge interactions from providing information about the oval's orientation. The buffer values correspond to the smallest distance between the edge of a vertically oriented oval and the top or bottom of the screen. Texture lines of the background and figure were truncated at the oval perimeter to provide a clean figure-ground signal (see Figure 1b for examples).

In the threshold detection experiment (Experiment 3), when the oval figures appeared, they were clearly elongated (5° by 3° in aspect ratio) but made difficult to segment from the ground by presenting the texture lines at short lengths, calibrated per participant to be near texture segmentation thresholds. In the suprathreshold experiment (Experiment 4), the texture lines were longer in length—fixed across participants from 0.18° to 3° in 7 common log steps—so that the oval figure was more clearly segmented from the ground, but the oval aspect ratio was calibrated per participant (i.e., made more circular) so that the orientation was difficult to discern.

Random textures. Random textures consisted of scattered lines, each randomly oriented and drawn at any of the 7 possible line lengths. The mean luminance of the random textures (controlled by the proportion of black pixels) was fixed to 0.15, matching the luminance of the figure-ground textures (see Methods, Luminance calibration). Random textures were presented before the targets, starting from fixation (see **Supplementary Figure 1** for timing), with the intent of reducing target-evoked contrast transients in future neuroimaging versions of the study and to reduce pupil-dilation effects at target onset.

Luminance calibration. To control the mean luminance of the textures, we performed a calibration procedure prior to running the experiments. The calibration procedure was run separately on each stimulus presentation computer at each site to ensure consistency regardless of differences in display properties. The purpose of the calibration procedure was to determine, for a given line length, how many randomly placed texture lines should be drawn to a quadrant of the screen to ensure that the average proportion of black pixels in the quadrant achieved a target value. (It is not straightforward to derive an analytical solution for this problem, given that randomly placed lines may frequently overlap and be truncated at the quadrant boundaries.) The procedure ensured that all textures had equal luminance regardless of differences in line length and computer display properties. We chose a target proportion of 0.15 black pixels to achieve qualitative similarity of appearance to texture stimuli previously used in the literature.

The calibration procedure proceeded as follows. For a given line length L and number of lines N, many background textures were drawn. For each texture generated in this way, the proportion P of black pixels drawn to the quadrant was computed. The expected value of the proportion of black pixels $\underline{P}$ was then computed as the average of P across texture instances. For a given value of L, the procedure was repeated for many values of N, generating samples of a function $\underline{P} = F(N|L)$ describing how the expected proportion of black pixels $\underline{P}$ depends on number of lines drawn N for a given value of line length L. Interpolation of these samples was used to find the value N_T achieving the desired target value of $\underline{P} = 0.15$ for the specified value of L.

This procedure was then repeated for many values of L, generating samples of a function $N_T = G(L)$ describing how N_T, the number of lines drawn to achieve $\underline{P} = 0.15$, depends on L. Interpolation of these samples was used to find the value N_T achieving the desired target value of $\underline{P} = 0.15$ for any arbitrary value of L. This second interpolation was necessary to determine in real time how many lines to draw to a quadrant for the many possible values of L that could be probed during the thresholding procedure of Experiment 3.

Since the structure of the random textures differed considerably from that of the oriented background textures, a similar calibration procedure was separately performed for the random textures to ensure proper luminance calibration at $\underline{P} = 0.15$. The line lengths of random textures were fixed values determined by the thresholding results for each participant, so this procedure only needed to interpolate over samples of a function $\underline{P} = H(N)$ and therefore could be performed quickly following the thresholding block to properly calibrate random textures for each participant before the main experiment.

Thresholding

To calibrate stimulus properties so that the objective task was appropriately challenging, each participant completed a thresholding procedure prior to the main experiment. The thresholding task was identical to the main task, except a stimulus feature was continuously adjusted between trials using the QUEST Bayesian adaptive staircasing procedure[110] and all precues were neutral. The adjusted feature in the detection experiments was stimulus strength (i.e., grating contrast or texture line length) and in the suprathreshold experiments was the grating tilt or oval aspect ratio. Three independent thresholding tracks, each adaptively approaching a discrimination accuracy of 75%, were interleaved for a total of 240 trials (80 per track) contributing to threshold estimates. On each trial, one of the three tracks was selected pseudo-randomly to set the strength of the response-cued stimulus to the current threshold estimate of that track, and subsequently that track had its threshold estimate updated using the participant's accuracy on that trial. If the individual tracks did not qualitatively converge, participants repeated the thresholding procedure. There was no statistically significant difference in the median thresholded values by experimental site (factors of experiment and site, no effects of site, all $p>0.144$, **Supplementary Table 17**), suggesting the two experimental setups and samples were reasonably similar.

During the thresholding procedure for the detection experiments (Experiments 1 and 3), while target-absent trials were included to mimic the structure of the main task, only target-present trials contributed to QUEST estimates. Likewise, although participants made a subjective visibility report simultaneously with their orientation report, as in the main experiment, only the orientation report contributed to threshold estimates. Using the thresholding track with the median threshold estimate, 5 middle stimulus strengths were selected to yield orientation discrimination accuracies of 60%, 67.5%, 75%, 82.5%, and 90%. The lowest and highest stimulus strengths were selected to yield near-floor and -ceiling performance. To select comparable values at each extreme, the lowest stimulus strength was chosen to be midway between the minimum possible strength (i.e., 0% contrast or a line length of 3 pixels) and the strength yielding near-chance (51%) accuracy on a common logarithm scale. The highest stimulus strength was chosen to be the line length yielding 90% accuracy plus d, where d was defined as the distance on a common log scale between the lowest stimulus strength and the strength yielding 60% accuracy. However, if in Experiment 1 this procedure resulted in a maximum grating contrast that exceeded 1 when added to noise, the thresholding procedure was repeated. In Experiment 3, if the procedure resulted in a maximum line length that exceeded the length of the oval minor axis (3°), the maximum line length was instead set to 3°.

During the thresholding procedure for the suprathreshold experiments (Experiments 2 and 4), stimuli were fixed to the middle (16% contrast or 0.41° line length) of the 7 possible target strengths (5% to 50% contrast or 0.18° to 3° line length in log steps). Because the targets were always presented at a fixed, easily visible stimulus strength, no subjective report was required. Instead, the task was made challenging by adjusting the grating orientation or oval aspect ratio to yield discrimination accuracy of 75%. The mean thresholded grating tilt was 8.9° (SD=6.6°). The mean thresholded oval aspect ratio as a proportion of the maximum aspect ratio was 0.67 (SD=0.31), with 1 corresponding to a maximally elongated oval (5:3° aspect ratio) and 0 a perfect circle. While the aspect ratio was adjusted, the area of the oval was held constant.

Online fixation monitoring

To ensure that effects of attention could not be attributed to saccades towards the precued location, participants were instructed to maintain fixation on a small black cross (0.35° wide) displayed within a gray aperture (2° radius) in the center of the screen. Before each session, a calibration sequence converted raw gaze position into degrees of visual angle. The start of every trial was contingent upon fixation for 500 ms within a 2° allowance. If fixation was not acquired within 10 s, participants were shown a message to keep their gaze locked on the center of the screen. If fixation was unable to be acquired within this time limit on two consecutive trials, the calibration sequence was repeated. After acquisition, fixation was enforced until the onset of the response cue. If fixation was lost before then, due either to a saccade or a blink, the trial was stopped and repeated at the end of the block. Any trial interrupted by a fixation break was not completed and therefore not part of the behavioral data used for analysis. On average, participants broke fixation on about 5% of trials.

Task training

To learn the task, participants first completed a self-guided instructions walk-through, followed by practice trials. Participants completed a sequence of practice trials with increasing difficulty: 1) slowed-down trials, 2) full-speed trials with auditory feedback, and 3) full-speed trials with no feedback. 1) The slowed-down trials were identical to trials in the main task, except stimulus timings were slowed down by a factor of 4. Participants had unlimited time to respond and received trial-by-trial feedback about both the objective and subjective reports in text to check their target and response mappings (e.g., “At the cued location, the stimulus was: counterclockwise (-45°). Your response was: counterclockwise (-45°), saw a grating but not its orientation”). 2) On trials with auditory feedback, the pitch of a 250 ms tone indicated orientation discrimination accuracy (correct=784 Hz, G5; incorrect=523 Hz, C5). Tones were enveloped by cosine ramps 10 ms wide. When the response-cued target was absent, the orientation judgment was irrelevant, so the auditory feedback played a high tone (G5) with 75% chance and a low tone (C5) otherwise, matching the expected average accuracy after thresholding. 3) The full-speed trials with no feedback mimicked trials in the main task.

For all practice stages, cueing validity, the response-cued quadrant, and stimulus properties were randomly chosen on each trial. Each practice block was 10 trials. The experimenter observed the participant during the practice trials to verify comprehension of the task and to answer any questions between blocks. Participants were allowed to repeat any of these blocks of practice trials until they felt comfortable with the task.

Reference training

In the suprathreshold experiments (Experiments 2 and 4), the subjective report was to judge the visibility of the target relative to a reference. To learn the reference strength, participants completed 120 trials identical to the main task, except all precues were neutral and the response-cued target was fixed to a "reference" strength, which was set to the middle value (i.e., 0.17 grating contrast or 0.41° line length) of the 7 pre-defined stimulus strengths in the main task (see Methods, Stimuli). Rather than make a subjective strength report, participants were instructed to notice and internalize the reference strength. Participants then completed 112 trials incorporating the attentional precue, individually thresholded grating tilt or oval aspect ratio, and subjective report comparing the visibility of the target to the newly learned reference, mimicking the main task.

Retention and exclusion

Datasets from participants who completed fewer than 2240 trials (i.e., 4 runs) were considered incomplete and not included in the analyses reported here. Across the 4 experiments at both sites, 21 participants with partial datasets were excluded. All partial datasets were due to participant dropout. We recruited additional participants until we reached the preregistered target sample size with complete datasets. To encourage retention, a monetary completion bonus was introduced during the third experiment at BU and the fourth experiment at UCI.

We excluded trials from analysis if no valid keypress was made before the response window timed out. One participant's data from Experiment 4 was excluded from all analyses due to chance performance on the main task despite sensible behavior during the thresholding procedure. Another participant's data from Experiment 4 was excluded from the AMI analyses because of chance performance on invalid trials regardless of stimulus strength, resulting in AUCs of 0 across all attention conditions, for which the AMI could not be computed. Although AUCs of 0 across attention conditions do not meaningfully contribute to the AUC analysis, the effect of attention on AUC was consistent whether or not this subject was included, so we included their data in the AUC analysis presented here.

Data analysis

Psychometric function fits

We fit Weibull functions to characterize objective performance $P_1 = F_1(x;\theta_1)$ and subjective reports $P_2 = F_2(x;\theta_2)$ as functions of stimulus strength x, separately for each participant and attention condition. Here, P denotes an objective or subjective measure as a probability, like p(correct), p(saw stimulus), or p(saw feature), and θ denotes the parameters of the Weibull function: α (threshold), β (slope), γ (guess rate), and λ (lapse rate). The full form of the Weibull function is given by:

$$P_n = F_n(x;\theta_n) = \gamma_n + (1-\gamma_n-\lambda_n)\left[1-e^{-(x/\alpha_n)^{\beta n}}\right] \tag{1}$$

Maximum likelihood estimation (MLE) of θ for each dataset was conducted using the Palamedes toolbox[111], treating trial-level responses as outcomes of a Bernoulli process. All parameters were free, except the guess rate for orientation discrimination performance γ_1, which was fixed to 0.5 corresponding to chance discriminability.

Parameters were optimized in a two-stage process: first, the best parameter estimates over the following predefined grid were identified, which then served as the starting point for Nelder-Mead optimization.

α: $[0.05,3]$ in increments of 0.05
β: $[10^{-1}, 10^{1}]$ in increments of $10^{0.1}$
λ: $[0,0.1]$ in increments of 0.01
γ_2 : $[0,1]$ in increments of 0.1

Relative psychometric function

To relate objective and subjective reports over their full psychometric functions, we created a general *relative psychometric function (RPF) analysis framework*[46]. An RPF is a function that characterizes how the values of one traditional psychometric function are related to the values of another when both depend on a common independent variable (such as stimulus strength). Applying the RPF framework to the current analysis, we expressed subjective reports P_2 as a function of objective performance P_1 using an RPF of the form $P_2 = R(P_1)$. The formula for *R* depends on the formulae for $F_1(x)$ and $F_2(x)$; here, since we used Weibulls to characterize both functions, we call the resulting RPF the *Weibull RPF*, denoted by R_W. Using equation (1), we can solve for $x = F_1^{-1}(P_1)$ and plug this into the equation for F_2 to arrive at the mathematical form of R_W:

$$P_2 = R_W(P_1; \theta_1, \theta_2) = \gamma_2 + (1 - \lambda_2 - \gamma_2)\left[1 - e^{-\left(\left(\frac{\alpha_2}{\alpha_1}\right)^{-\beta_2}\left(ln\left(\frac{1-\lambda_1-\gamma_1}{1-\lambda_1-P_1}\right)\right)^{\beta_2/\beta_1}\right)}\right] \quad (2)$$

We used R_W to mathematically characterize the relationship between subjective reports and objective performance, given the parameters θ_1 and θ_2. In turn, these parameters were found via independent MLE fits for $F_1(x; \theta_1)$ and $F_2(x; \theta_2)$, respectively. Full details of the relative psychometric function analysis framework can be found in Maniscalco et al.[46]

Area under the relative psychometric function

To quantify subjective reports over a given range of objective performance, following procedures established by Maniscalco et al.[46], we calculated the area under the relative psychometric function ("area under the curve", AUC). We approximated the cumulative integral of R_W using the trapezoidal method, partitioning P_1 into segments of 0.01 and computing P_2 at each value using equation (2). We constrained $P_1$to a matched range across attention conditions (shaded gray regions in **Figures 2c, 3c**) per participant. For all attention conditions of each participant, we set the lower bound of P_1 to 0.5 (i.e., chance performance), capitalizing on the fact that psychometric function fits ensured defined values for P_2 at $P_1 = 0.5$. Separately for each participant, we set the upper bound of P_1 to the maximum fitted P_1 value *within* each attention condition that was minimal *across* conditions. This maximized the P_1 interval used to compute AUC for each participant, subject to the constraint that AUC be computed using the same P_1 interval for each within-participant condition. This constraint ensured that any across-condition modulation in AUC was due only to differences in P_2 over a fixed P_1 interval. To measure subjective inflation, we compared the AUC across varying levels of attention. A larger AUC indicates higher reported visibility for a given condition over a common range of objective performance.

Attentional modulation index

To compare the degree and direction with which subjective reports dissociate from objective performance with attention across different experiments, we calculated an attentional modulation index (AMI) for each participant as the difference in AUC between invalid and valid conditions as a proportion of their summed AUC.

$$\mathrm{AMI}_{\mathrm{AUC}} = \frac{\mathrm{AUC}_{\mathrm{Invalid}} - \mathrm{AUC}_{\mathrm{Valid}}}{\mathrm{AUC}_{\mathrm{Invalid}} + \mathrm{AUC}_{\mathrm{Valid}}} \tag{3}$$

Signal detection theory

We calculated signal detection theory measures[61,112] of sensitivity (*d'*) and criterion (c_{YN}) using equations (4-5) for each participant and attention condition, where z is the inverse of the normal cumulative distribution function and c_{YN} denotes the criterion for the yes-no discrimination task (see **Supplementary Note 2**). To handle conditions with hit (H) or false alarm (F) rates of 0 or 1, for which the signal detection measures are indeterminate, we applied a log-linear correction[113], adding 0.5 to all response condition counts and 1 to all stimulus condition counts. These measures assume equal variance of the internal signal and noise distributions.

$$d' = z(H) - z(F) \tag{4}$$

$$c_{YN} = -0.5(z(H) + z(F)) \tag{5}$$

We also calculated the signal detection variants d_a and c_a using equations (6-7). These measures relax the equal variance assumption, but require estimating s, a slope parameter of the z-transformed receiver operator characteristic (ROC) curve (hit vs. false alarm rates). We leveraged our two subjective measures (of seeing the stimulus and of seeing the task-relevant feature) to estimate s **(Supplementary Figure 6c)**, taking the slope of the line between the two points in ROC space.

$$d_a = \sqrt{\frac{2}{1+s^2}}(z(H) - s \times z(F)) \tag{6}$$

$$c_a = \frac{-\sqrt{2}s}{\sqrt{1+s^2}(1+s)}(z(H) + z(F)) \tag{7}$$

Statistical analysis

Mixed ANOVAs were performed in R (v4.3.1; R Core Team 2023) and MATLAB to evaluate: 1) the effects of stimulus strength and attention on objective and subjective reports, including signal detection theory measures, 2) the effects of attention on the AUC, and 3) the effects of stimulus and task conditions on the AMI of the AUC. The within-subject factors included attention (valid, neutral, or invalid, with respect to the match between the precue and response cue), stimulus strength level, and visibility report type (of the overall stimulus vs. the task-relevant feature). The between-subjects factors included experimental site (BU vs. UCI) and experiment (1-4). When applicable, experiment was factorized into stimulus type (gratings vs. texture-defined figures) and task type (threshold detection vs. suprathreshold visibility comparison). All statistical tests were two-sided. Effect sizes are reported as generalized eta-squared. When Mauchly's test indicated violations of sphericity assumptions[114], we confirmed that all significant *F* tests remained significant after Greenhouse-Geisser correction[115], with the estimated degree of sphericity violation reported as epsilon (ε). Planned analyses were conducted per experiment. Planned comparisons were made between stimulus types, task types, visibility report types, and all pairs of cueing validity and stimulus strengths, when applicable.

Acknowledgments

This research was supported by the Templeton World Charity Foundation Accelerating Research on Consciousness Initiative TWCF 0567 to B.J.H., J.W.B., N.B., D.C., M.A.K.P., and R.N.D., National Defense Science and Engineering Graduate Fellowship to K.J.T., startup funding from the University of California Irvine and support from the Canadian Institute of Advanced Research Brain, Mind, & Consciousness program to M.A.K.P., and startup funding from Boston University to R.N.D. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank Sarah Cormiea, Angus Chapman, Jeffrey Nestor, and Brian Odegaard for helpful discussions at different stages of the project and Giancarlo Arzu and Lizbeth Romero for assistance with data collection.

Data availability

The data will be made publicly available upon publication.

Code availability

The code to run the experiments is available at https://github.com/denisonlab/twcf_expt1_code and the presented analyses is available at https://github.com/denisonlab/twcf_expt1_analysis.

The toolbox for fitting the relative psychometric function is available at https://github.com/CNClaboratory/RPF, with additional documentation provided in the accompanying paper by Maniscalco et al.[46]

Author contributions

K.J.T., B.M., M.L.E., A.S., M.A.K.P., and R.N.D. conceptualized and designed the experiments with feedback from R.B., V.A.F.L., H.L., B.J.H., J.W.B., N.B., and D.C. Theory predictions as to the experimental outcomes were supplied by V.A.F.L., R.B., and H.L. K.J.T. and B.M. programmed the experiments. K.J.T., M.L.E., A.S., O.G.C., T.K., J.A.M., E.O., E.E.R., M.E.W., J.W., T.B.A.Z., and M.A.K.P. collected the data. K.J.T. and B.M. analyzed and visualized the data with input and supervision from M.A.K.P. and R.N.D. K.J.T., B.M., M.A.K.P., and R.N.D. wrote the manuscript. M.L.E., A.S., J.A.M., H.L., R.B., B.J.H., J.W.B., N.B., and D.C. provided feedback on the manuscript draft. All authors approved the final version.

Competing interests

The authors declare no competing interests.

References

1. Jigo, M. & Carrasco, M. Differential impact of exogenous and endogenous attention on the contrast sensitivity function across eccentricity. *J Vision* **20**, 11 (2020).

2. Carrasco, M. Visual attention: the past 25 years. *Vision Res* **51**, 1484–1525 (2011).

3. Herrmann, K., Montaser-Kouhsari, L., Carrasco, M. & Heeger, D. J. When size matters: attention affects performance by contrast or response gain. *Nat Neurosci* **13**, 1554–1559 (2010).

4. Cameron, E. L., Tai, J. C. & Carrasco, M. Covert attention affects the psychometric function of contrast sensitivity. *Vis. Res.* **42**, 949–967 (2002).

5. Lu, Z.-L. & Dosher, B. A. External noise distinguishes attention mechanisms. *Vis. Res.* **38**, 1183–1198 (1998).

6. Yeshurun, Y., Montagna, B. & Carrasco, M. On the flexibility of sustained attention and its effects on a texture segmentation task. *Vis. Res.* **48**, 80–95 (2008).

7. Yeshurun, Y. & Carrasco, M. The locus of attentional effects in texture segmentation. *Nat. Neurosci.* **3**, 622–627 (2000).

8. Carrasco, M. & Yeshurun, Y. Covert attention effects on spatial resolution. *Prog. Brain Res.* **176**, 65–86 (2009).

9. Carrasco, M. & Barbot, A. Spatial attention alters visual appearance. *Curr. Opin. Psychol.* **29**, 56–64 (2019).

10. Luo, Z.-X., Pan, W.-N., Zeng, X.-J., Gong, L.-Y. & Cai, Y.-C. Endogenous attention enhances contrast appearance regardless of stimulus contrast. *Atten., Percept., Psychophys.* **86**, 1883–1896 (2024).

11. Carrasco, M., Ling, S. & Read, S. Attention alters appearance. *Nat Neurosci* **7**, 308–313 (2004).

12. Liu, T., Abrams, J. & Carrasco, M. Voluntary attention enhances contrast appearance. *Psychol. Sci.* **20**, 354–362 (2008).

13. Störmer, V. S., McDonald, J. J. & Hillyard, S. A. Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proc. Natl. Acad. Sci.* **106**, 22456–22461 (2009).

14. Itthipuripat, S., Chang, K.-Y., Bong, A. & Serences, J. T. Stimulus visibility and uncertainty mediate the influence of attention on response bias and visual contrast appearance. *J. Vis.* **19**, 8 (2019).

15. Anton-Erxleben, K., Abrams, J. & Carrasco, M. Evaluating comparative and equality judgments in contrast perception: attention alters appearance. *J. Vis.* **10**, 6–6 (2010).

16. Abrams, J., Barbot, A. & Carrasco, A. Voluntary attention increases perceived spatial frequency. *Atten., Percept., Psychophys.* **72**, 1510–1521 (2010).

17. Gobell, J. & Carrasco, M. Attention alters the appearance of spatial frequency and gap size. *Psychol. Sci.* **16**, 644–651 (2004).

18. Barbot, A., Liu, S., Kimchi, R. & Carrasco, M. Attention enhances apparent perceptual organization. *Psychon. Bull. Rev.* **25**, 1824–1832 (2018).

19. Wilimzig, C., Tsuchiya, N., Fahle, M., Einhäuser, W. & Koch, C. Spatial attention increases performance but not subjective confidence in a discrimination task. *J. Vis.* **8**, 7 (2008).

20. Chica, A. B., Lasaponara, S., Chanes, L., Valero-Cabré, A., Doricchi, F., Lupiáñez, J. & Bartolomeo, P. Spatial attention and conscious perception: the role of endogenous and exogenous orienting. *Atten., Percept., Psychophys.* **73**, 1065–1081 (2011).

21. Rahnev, D., Maniscalco, B., Graves, T., Huang, E., de Lange, F. P. & Lau, H. Attention induces conservative subjective biases in visual perception. *Nat Neurosci* **14**, 1513–1515 (2011).

22. Li, M. K., Lau, H. & Odegaard, B. An investigation of detection biases in the unattended periphery during simulated driving. *Atten. Percept. Psychophys.* **80**, 1325–1332 (2018).

23. Solovey, G., Graney, G. G. & Lau, H. A decisional account of subjective inflation of visual perception at the periphery. *Atten Percept Psychophys* **77**, 258–271 (2015).

24. Knotts, J., Odegaard, B., Lau, H. & Rosenthal, D. Subjective inflation: phenomenology's get-rich-quick scheme. *Curr Opin Psychology* **29**, 49–55 (2019).

25. Knotts, J. D., Michel, M. & Odegaard, B. Defending subjective inflation: an inference to the best explanation. *Neurosci. Conscious.* **2020**, niaa025 (2020).

26. Odegaard, B., Chang, M. Y., Lau, H. & Cheung, S.-H. Inflation versus filling-in: why we feel we see more than we actually do in peripheral vision. *Philosophical Transactions Royal Soc B Biological Sci* **373**, 20170345 (2018).

27. Abid, G. Deflating inflation: the connection (or lack thereof) between decisional and metacognitive processes and visual phenomenology. *Neurosci. Conscious.* **2019**, niz015 (2019).

28. Okubo, L. & Yokosawa, K. Attentional allocation and the pan-field color illusion. *J. Vis.* **23**, 13 (2023).

29. Okubo, L., Miyoshi, K., Yokosawa, K. & Nishida, S. Inattentional noise leads to subjective color uniformity across the visual field. *PsyArXiv* (2024).

30. Brown, R., Lau, H. & LeDoux, J. E. Understanding the higher-order approach to consciousness. *Trends Cogn Sci* **23**, 754–768 (2019).

31. Azzopardi, P. & Cowey, A. Preferential representation of the fovea in the primary visual cortex. *Nature* **361**, 719–721 (1993).

32. Lau, H. C. A higher order Bayesian decision theory of consciousness. *Prog. Brain Res.* **168**, 35–48 (2007).

33. Lau, H. & Rosenthal, D. Empirical support for higher-order theories of conscious awareness. *Trends Cogn Sci* **15**, 365–373 (2011).

34. Brown, R. The HOROR theory of phenomenal consciousness. *Philos. Stud.* **172**, 1783–1794 (2015).

35. LeDoux, J. E. & Brown, R. A higher-order theory of emotional consciousness. *Proc. Natl. Acad. Sci.* **114**, E2016–E2025 (2017).

36. Lau, H. Consciousness, metacognition, and perceptual reality monitoring. *PsyArXiv* (2019).

37. Gennaro, R. J. *Higher-Order Theories of Consciousness*. (John Benjamins Press, 2004).

38. Seth, A. K. & Bayne, T. Theories of consciousness. *Nat Rev Neurosci* **23**, 439–452 (2022).

39. Mudrik, L., Boly, M., Dehaene, S., Fleming, S. M., Lamme, V., Seth, A. & Melloni, L. Unpacking the complexities of consciousness: theories and reflections. *Neurosci. Biobehav. Rev.* **170**, 106053 (2025).

40. Lamme, V. A. F. & Roelfsema, P. R. The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* **23**, 571–579 (2000).

41. Lamme, V. A. F. Visual functions generating conscious seeing. *Front. Psychol.* **11**, 83 (2020).

42. Lamme, V. A. F. Towards a true neural stance on consciousness. *Trends Cogn. Sci.* **10**, 494–501 (2006).

43. Odegaard, B., Li, M. & Lau, H. Subjective inflation in the unattended periphery in a naturalistic environment. *J. Vis.* **18**, 530 (2018).

44. Vernet, M., Japee, S., Lokey, S., Ahmed, S., Zachariou, V. & Ungerleider, L. G. Endogenous visuospatial attention increases visual awareness independent of visual discrimination sensitivity. *Neuropsychologia* **128**, 297–304 (2019).

45. Morales, J., Odegaard, B. & Maniscalco, B. The neural substrates of conscious perception without performance confounds. in *Anthology of Neuroscience & Philosophy* (eds. Brigard, F. de & Sinnott-Armstrong, W.) (MIT Press, 2022).

46. Maniscalco, B., Castaneda, O. G., Odegaard, B., Morales, J., Rajananda, S., Denison, R. & Peters, M. The relative psychometric function: a general analysis framework for relating psychological processes. *PsyArXiv* (2024).

47. Latham, G. P., Erez, M. & Locke, E. A. Resolving scientific disputes by the joint design of crucial experiments by the antagonists: application to the Erez–Latham dispute regarding participation in goal setting. *J. Appl. Psychol.* **73**, 753–772 (1988).

48. Nothdurft, H. C. Orientation sensitivity and texture segmentation in patterns with different line orientation. *Vis. Res.* **25**, 551–560 (1985).

49. Peters, M. A. K. Introspective psychophysics for the study of subjective experience. *Cereb. Cortex* **8**, 49–57 (2025).

50. Block, N., Brascamp, J., Chalmers, D., Denison, R., He, B. J. & Peters, M. An adversarial collaboration to test predictions of first-order and higher-order theories of consciousness [preregistration]. *Open Science Framework* (2020).

51. Navajas, J., Bahrami, B. & Latham, P. E. Post-decisional accounts of biases in confidence. *Curr. Opin. Behav. Sci.* **11**, 55–60 (2016).

52. Wierzchoń, M., Paulewicz, B., Asanowicz, D., Timmermans, B. & Cleeremans, A. Different subjective awareness measures demonstrate the influence of visual identification on perceptual awareness ratings. *Conscious. Cogn.* **27**, 109–120 (2014).

53. Lee, D. K., Itti, L., Koch, C. & Braun, J. Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* **2**, 375–381 (1999).

54. Supèr, H., Spekreijse, H. & Lamme, V. A. F. Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat Neurosci* **4**, 304–310 (2001).

55. Heitz, R. P. The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Front. Neurosci.* **8**, 150 (2014).

56. Koch, C. & Tsuchiya, N. Attention and consciousness: two distinct brain processes. *Trends Cogn Sci* **11**, 16–22 (2007).

57. Peters, M. A. K. & Lau, H. Human observers have optimal introspective access to perceptual processes even for visually masked stimuli. *Elife* **4**, e09651 (2015).

58. Peters, M. A. K., Fesi, J., Amendi, N., Knotts, J. D., Lau, H. & Ro, T. Transcranial magnetic stimulation to visual cortex induces suboptimal introspection. *Cortex* **93**, 119–132 (2017).

59. Itthipuripat, S., Phangwiwat, T., Wiwatphonthana, P., Sawetsuttipan, P., Chang, K.-Y., Störmer, V. S., Woodman, G. F. & Serences, J. T. Dissociable neural mechanisms underlie the effects of attention on visual appearance and response bias. *J. Neurosci.* **43**, 6628–6652 (2023).

60. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics*. (Wiley, 1966).

61. Macmillan, N. A. & Creelman, C. D. *Detection theory: a user's guide*. vol. 2nd ed. (Lawrence Erlbaum, 2005).

62. Hawkins, H. L., Hillyard, S. A., Luck, S. J., Mouloua, M., Downing, C. J. & Woodward, D. P. Visual attention modulates signal detectability. *J. Exp. Psychol.: Hum. Percept. Perform.* **16**, 802–811 (1990).

63. Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L. & Tallon-Baudry, C. Cueing attention after the stimulus Is gone can retrospectively trigger conscious perception. *Curr. Biol.* **23**, 150–155 (2013).

64. Thomas, J. P. & Gille, J. Bandwidths of orientation channels in human vision. *J. Opt. Soc. Am.* **69**, 652 (1979).

65. Cohen, M. A., Ostrand, C., Frontero, N. & Pham, P.-N. Characterizing a snapshot of perceptual experience. *J Exp Psychology Gen* **150**, 1695–1709 (2021).

66. Cohen, M., Sung, S. & Alaoui, Z. Familiarity alters the bandwidth of perceptual awareness. *J Cogn Neurosci* **1**, 1546–1556 (2024).

67. Winter, C. J. & Peters, M. A. K. Variance misperception under skewed empirical noise statistics explains overconfidence in the visual periphery. *Atten., Percept., Psychophys.* **84**, 161–178 (2022).

68. Otten, M., Pinto, Y., Paffen, C. L. E., Seth, A. K. & Kanai, R. The uniformity illusion. *Psychol. Sci.* **28**, 56–68 (2016).

69. Lee, J. L., Denison, R. & Ma, W. J. Challenging the fixed-criterion model of perceptual decision-making. *Neurosci Conscious* **2023**, (2023).

70. Pruitt, J., Knotts, J. D. & Odegaard, B. Consistent metacognitive efficiency and variable response biases in peripheral vision. *J. Vis.* **24**, 4 (2024).

71. Block, N. *The border between seeing and thinking*. (Oxford University Press, 2023).

72. Ge, Y., Taschereau-Dumouchel, V., Lin, Q., Moharramipour, A., Sun, Z. & Lau, H. Representation of visual uniformity in the lateral prefrontal cortex. *bioRxiv* (2024).

73. Michel, M. The perceptual reality monitoring theory. in *Scientific Theories of Consciousness: The Grand Tour* (eds. Herzog, M., Schurger, A. & Doerig, A.) (Cambridge University Press).

74. Brown, R. The brain and its states. in *Being in Time: Dynamical Models of Phenomenal Experience.* (eds. Edelman, S., Fekete, T. & Zach, N.) (John Benjamins Press, 2012).

75. Brown, R. The myth of phenomenological overflow. *Conscious. Cogn.* **21**, 599–604 (2012).

76. Kirchberger, L., Mukherjee, S., Schnabel, U. H., Beest, E. H. van, Barsegyan, A., Levelt, C. N., Heimel, J. A., Lorteije, J. A. M., Togt, C. van der, Self, M. W. & Roelfsema, P. R. The essential role of recurrent processing for figure-ground perception in mice. *Sci. Adv.* **7**, eabe1833 (2021).

77. Self, M. W., van Kerkoerle, T., Supèr, H. & Roelfsema, P. R. Distinct roles of the cortical layers of area V1 in figure-ground segregation. *Curr. Biol.* **23**, 2121–2129 (2013).

78. Poort, J., Raudies, F., Wannig, A., Lamme, V. A. F., Neumann, H. & Roelfsema, P. R. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* **75**, 143–156 (2012).

79. Lamme, V. The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* **15**, 1605–1615 (1995).

80. Cohen, M. A., Dennett, D. C. & Kanwisher, N. What is the bandwidth of perceptual experience? *Trends Cogn. Sci.* **20**, 324–335 (2016).

81. Mack, A. & Rock, I. *Inattentional blindness*. (MIT Press, 1998).

82. Simons, D. J. Current approaches to change blindness. *Vis. Cogn.* **7**, 1–15 (2000).

83. Tversky, A. & Kahneman, D. Judgment under uncertainty: heuristics and biases. *Science* **185**, 1124–1131 (1974).

84. Michel, M. How (not) to underestimate unconscious perception. *Mind Lang.* **38**, 413–430 (2022).

85. Cohen, M. R. & Maunsell, J. H. R. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* **12**, 1594–1600 (2009).

86. Rahnev, D. A., Bahdo, L., de Lange, F. P. & Lau, H. Prestimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception. *J. Neurophysiol.* **108**, 1529–1536 (2012).

87. Rahnev, D. A., Maniscalco, B., Luber, B., Lau, H. & Lisanby, S. H. Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. *J. Neurophysiol.* **107**, 1556–1563 (2012).

88. Ko, Y. & Lau, H. A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. *Philos. Trans. R. Soc. B: Biol. Sci.* **367**, 1401–1411 (2012).

89. Denison, R. N., Block, N. & Samaha, J. What do models of visual perception tell us about visual phenomenology? in *Neuroscience and Philosophy* (eds. Brigard, F. D. & Sinnott-Armstrong, W.) (MIT Press, 2022).

90. Block, N. Two neural correlates of consciousness. *Trends Cogn. Sci.* **9**, 46–52 (2005).

91. Aru, J., Bachmann, T., Singer, W. & Melloni, L. Distilling the neural correlates of consciousness. *Neurosci. Biobehav. Rev.* **36**, 737–746 (2012).

92. Chalmers, D. What is a neural correlate of consciousness? in *Neural Correlates of Consciousness: Conceptual and Empirical Questions* (ed. Metzinger, T.) (MIT Press, 2000).

93. Dehaene, S. & Naccache, L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* **79**, 1–37 (2001).

94. He, B. J. Towards a pluralistic neurobiological understanding of consciousness. *Trends Cogn Sci* **27**, 420–432 (2023).

95. Samaha, J., Iemi, L. & Postle, B. R. Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Conscious. Cogn.* **54**, 47–55 (2017).

96. Lau, H. C. & Passingham, R. E. Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc National Acad Sci* **103**, 18763–18768 (2006).

97. Koizumi, A., Maniscalco, B. & Lau, H. Does perceptual confidence facilitate cognitive control? *Atten., Percept., Psychophys.* **77**, 1295–1306 (2015).

98. Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E. & Lau, H. Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn Neurosci* **1**, 165–175 (2010).

99. Maniscalco, B. & Lau, H. Manipulation of working memory contents selectively impairs metacognitive sensitivity in a concurrent visual discrimination task. *Neurosci. Conscious.* **2015**, niv002 (2015).

100. Maniscalco, B. & Lau, H. The signal processing architecture underlying subjective reports of sensory awareness. *Neurosci. Conscious.* **2016**, niw002 (2016).

101. Maniscalco, B., McCurdy, L. Y., Odegaard, B. & Lau, H. Limited cognitive resources explain a trade-off between perceptual and metacognitive vigilance. *J. Neurosci.* **37**, 1213–1224 (2017).

102. Zylberberg, A., Fetsch, C. R. & Shadlen, M. N. The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *eLife* **5**, e17688 (2016).

103. Zylberberg, A., Barttfeld, P. & Sigman, M. The construction of confidence in a perceptual decision. *Front. Integr. Neurosci.* **6**, 79 (2012).

104. Maniscalco, B., Peters, M. A. K. & Lau, H. Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. *Atten., Percept., Psychophys.* **78**, 923–937 (2016).

105. Weiskrantz, L., Warrington, E. K., Sanders, M. D. & Marshall, J. Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain* **97**, 709–728 (1974).

106. Michel, M., Gao, Y., Mazor, M., Kletenik, I. & Rahnev, D. When visual metacognition fails: widespread anosognosia for visual deficits. *Trends Cogn. Sci.* **28**, 1066–1077 (2024).

107. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision* **10**, 437–442 (1997).

108. Kleiner, M., Brainard, D. & Pelli, D. What's new in Psychtoolbox-3? *Perception* (2007).

109. Brainard, D. H. The Psychophysics Toolbox. *Spatial Vision* **10**, 433–436 (1997).

110. Watson, A. B. & Pelli, D. G. QUEST: A Bayesian adaptive psychometric method. *Percept Psychophys* **33**, 113–120 (1983).

111. Prins, N. & Kingdom, F. A. A. Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes Toolbox. *Front Psychol* **9**, 1250 (2018).

112. Maniscalco, B. & Lau, H. Signal detection theory analysis of type 1 and type 2 data: meta-d', response-specific meta-d', and the unequal variance SDT model. in *The Cognitive Neuroscience of Metacognition* (eds. Fleming, S. M. & Frith, C. D.) 25–66 (Springer, 2014).

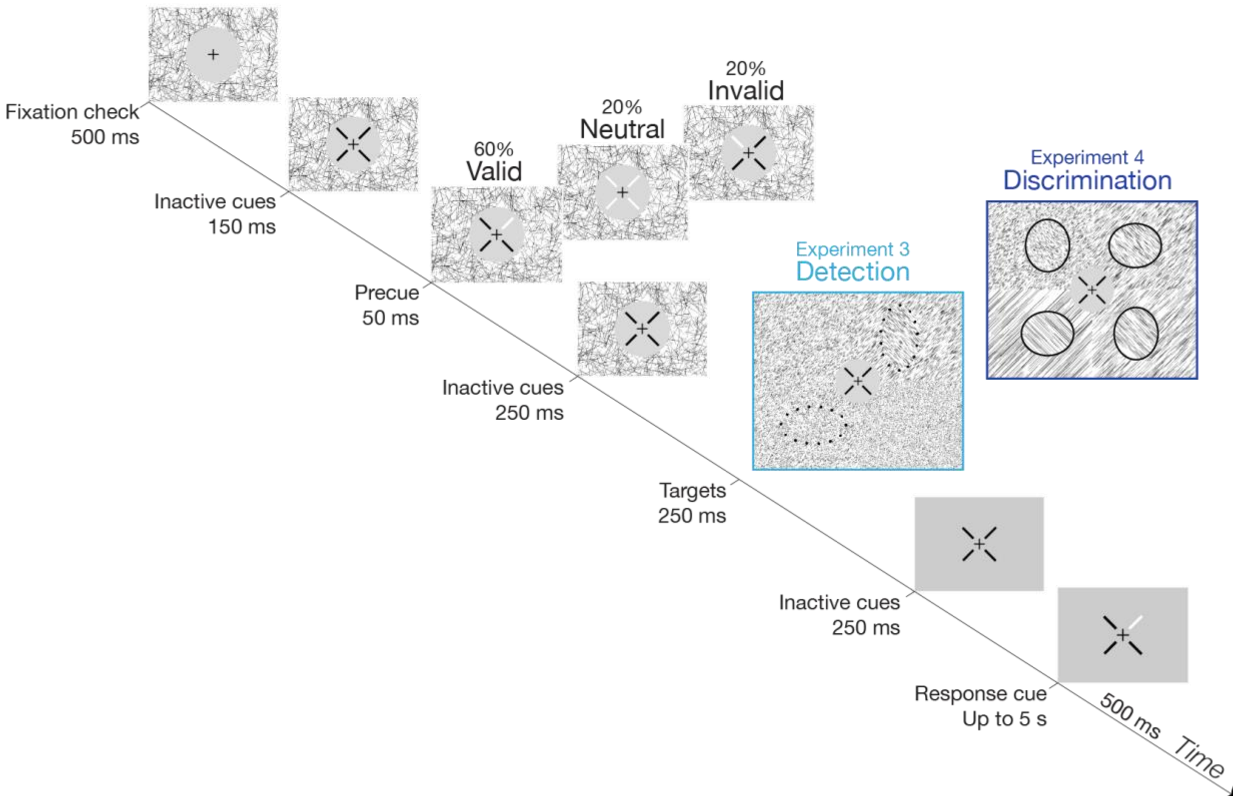
113. Stanislaw, H. & Todorov, N. Calculation of signal detection theory measures. *Behav. Res. Methods, Instrum., Comput.* **31**, 137–149 (1999).

114. Mauchly, J. Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics* **11**, 204–209 (1940).

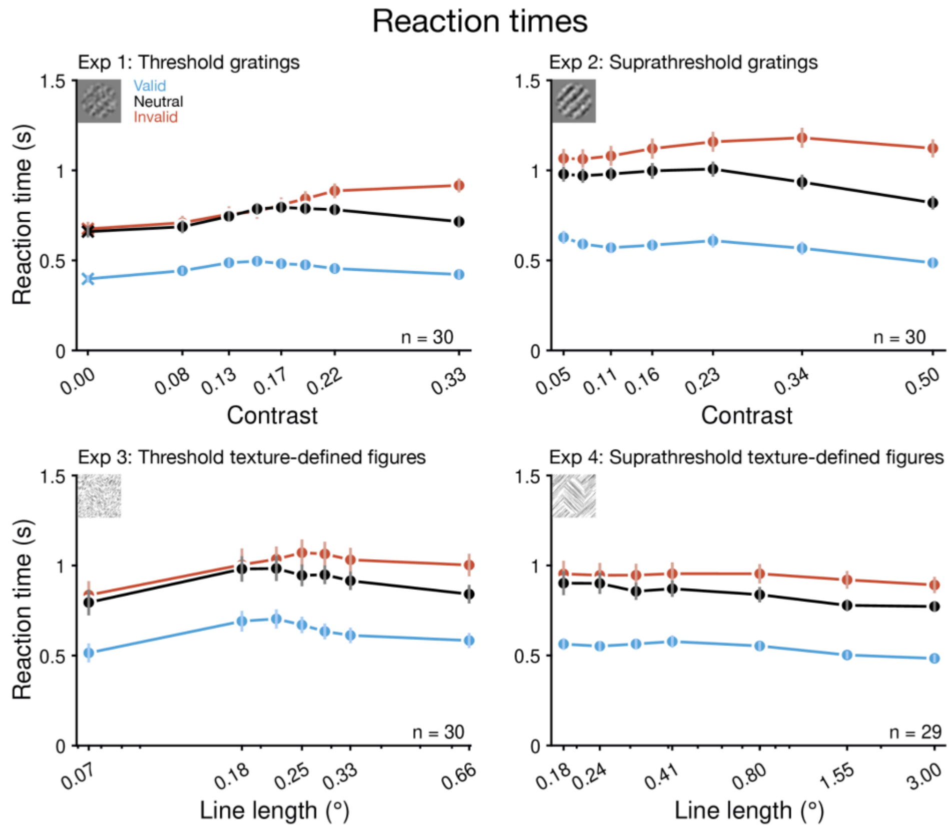
115. Greenhouse, S. W. & Geisser, S. On methods in the analysis of profile data. *Psychometrika* **24**, 95–112 (1959).

Supplementary Information

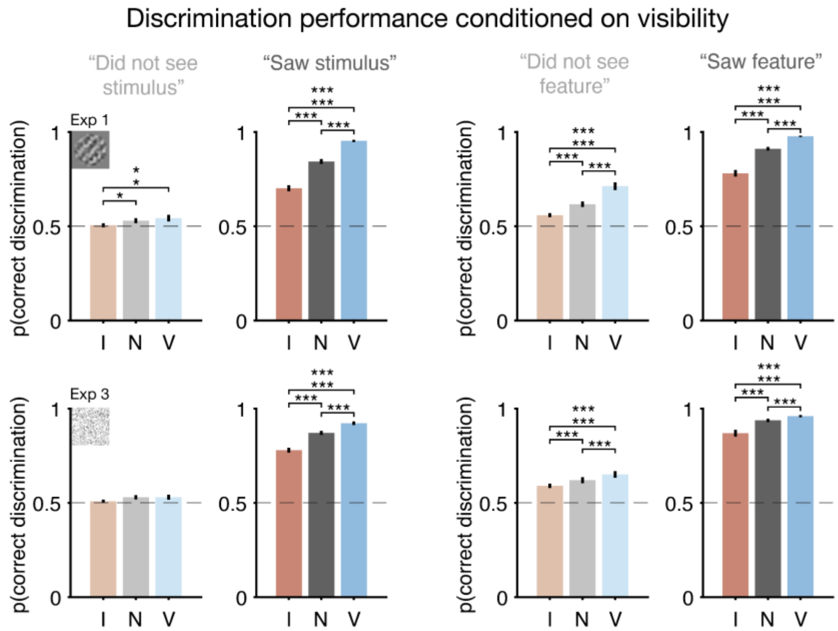
Supplementary Figures



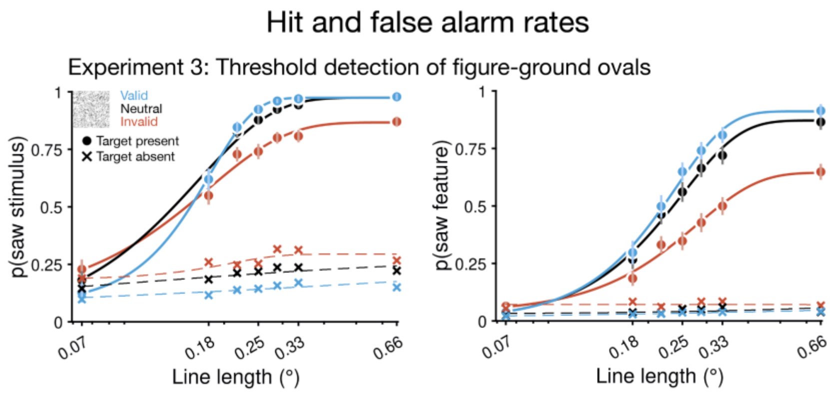
Supplementary Figure 1. Detailed task timeline for figure-ground experiments. Expansion of Figure 1, to detail the trial sequence of the experiments featuring texture-defined figure-ground stimuli (Experiments 3 and 4). Each trial was initiated by a fixation check. From fixation until the offset of the target stimuli, the central fixation cross, attentional precues, and inactive cues were presented in a circular gray aperture 4° in diameter. The circular aperture was placed atop a random texture composed of lines of different orientations and lengths until target onset. The target display was divided into quadrants. Each quadrant was filled with a unique “background” texture, composed of lines of a single orientation (-45° or +45°) and length, in which an oval “figure” delineated by orthogonal lines could appear (Experiment 3) or always appeared (Experiment 4).



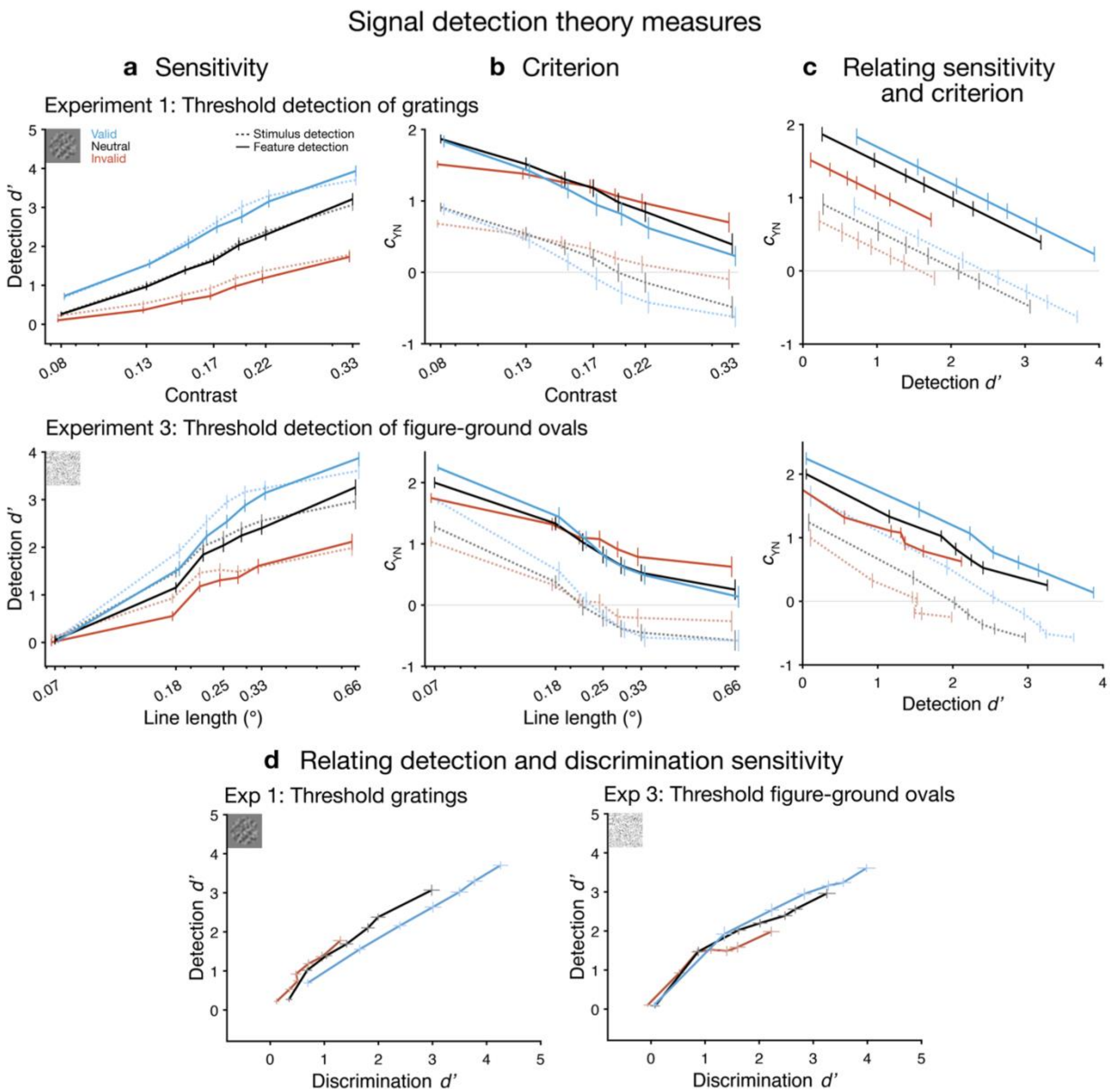
Supplementary Figure 2. Reaction times. In all experiments, reaction times were fastest on trials when attention was validly cued (blue), intermediate when it was distributed neutrally (gray), and slowest when it was invalidly cued (red). Thus attentional benefits to performance were not driven by speed-accuracy tradeoffs. Data (total n=120; Experiments 1-3 each n=30, Experiment 4 n=29) are presented as mean values ±1 SEM.



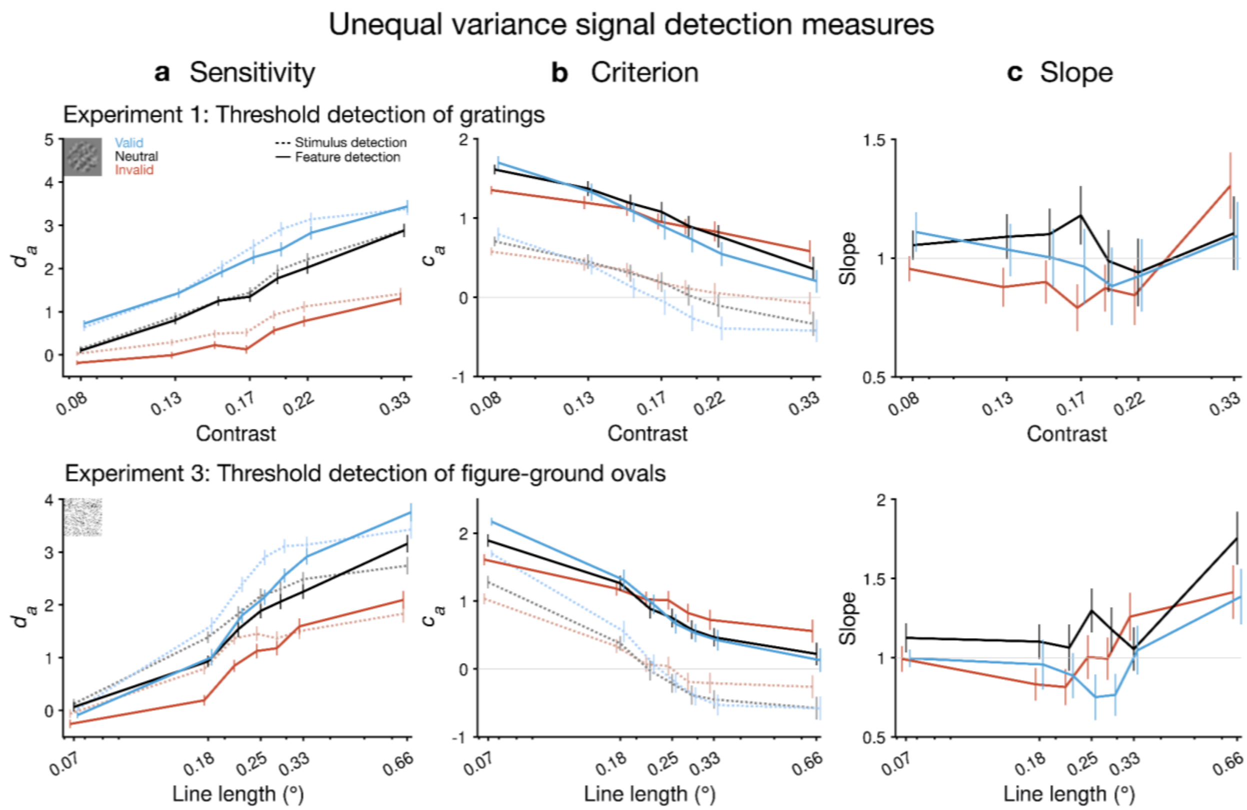
Supplementary Figure 3. Discrimination performance conditioned on subjective visibility in the detection experiments. Experiment 1 (top) and 3 (bottom). The attentional cue modulated performance even when the stimuli were reported as unseen. *p<0.05, **p<0.01, ***p<0.001. Data (total n=60; Experiments 1 and 3 each n=30) are presented as mean values ±1 SEM.



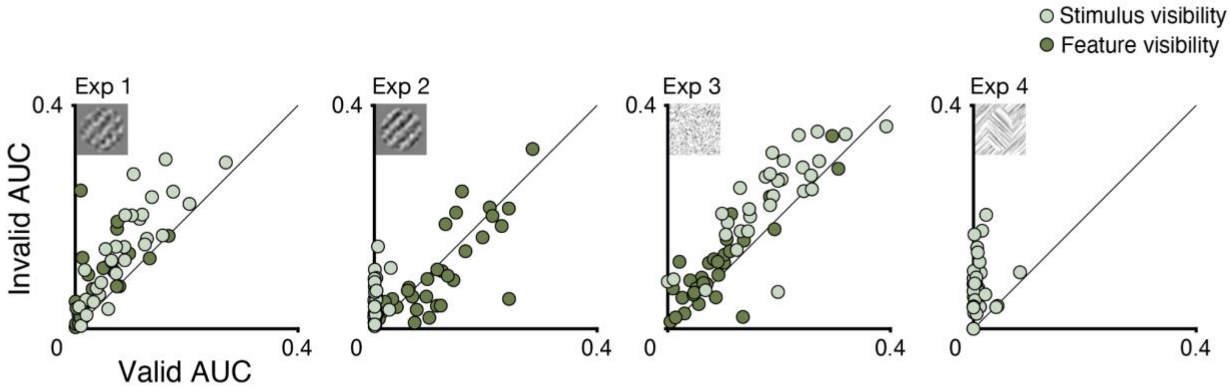
Supplementary Figure 4. Hit and false alarm rates for threshold figure-ground experiment. In the figure-ground experiment, line lengths were matched between target present and absent conditions, allowing hits and false alarm rates to be estimated separately at each stimulus strength. Inattention (red) both decreased hit rates (i.e., reports of seeing a texture-defined figure when there was one; solid lines) except at the lowest line length and increased false alarm rates (i.e., reports of seeing a texture-defined figure when there was none; dashed lines). These patterns were consistent whether the subjective report pertained to detection of any figure within the texture (left) or detection of a particular figure feature (the oval orientation; right). Group means are fit with Weibull functions for visualization. Data (Experiment 3 n=30) are presented as mean values ±1 SEM.



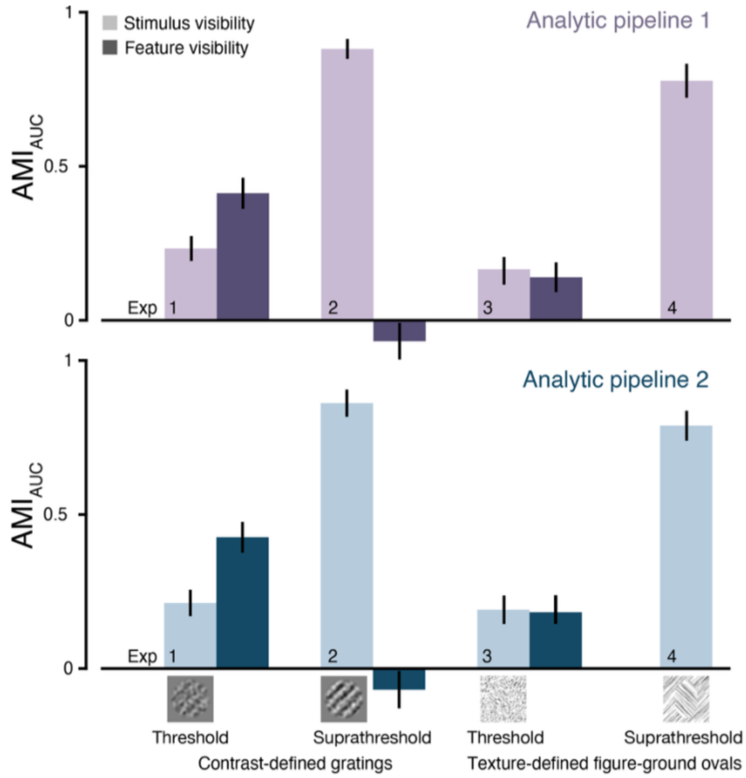
Supplementary Figure 5. Signal detection theory measures. a) Sensitivity (*d'*) in detecting gratings (top) and texture-defined figures (bottom) increased with stimulus strength and with attention (valid = blue, neutral = gray, invalid = red). **b)** Criterion (c_{YN}) was more liberal for reports of seeing an overall stimulus (dashed lines) than reports of seeing a particular stimulus feature (orientation, solid lines). Although participants were more conservative in reporting they saw the task-relevant feature than the overall stimulus, detection sensitivity was interestingly similar for the two subjective report types. At matched stimulus strengths, the influence of attention on criterion was strength-dependent: attention induced a more conservative criterion at lower stimulus strengths, but a more liberal criterion at higher stimulus strengths. See the main text and **Supplementary Note 2** for an argument that these changes in criterion are driven by changes in *d'*. **c)** At matched performance, the effect of attention was to consistently induce a conservative criterion. **d)** Here for comparison to the p(seen) vs. p(correct) plots in **Figure 2c** and **Figure 3c**, we plot signal detection sensitivity for detection vs. discrimination. Detection sensitivity—the ability to distinguish when the target was present vs. absent—was not systematically higher for unattended trials at matched discriminabilities across experiments. Data (total n=60; Experiments 1 and 3 each n=30) are presented as mean values ±1 SEM. Horizontal and vertical error bars in panel d indicate SEMs for discrimination and detection *d'* respectively.



Supplementary Figure 6. Unequal variance signal detection theory measures. The effects of stimulus strength and attention on the unequal variance measures of **a)** detection sensitivity d_a and **b)** criterion c_a were consistent with their effects on the equal variance signal detection measures (Supplementary Figure 5). That is, attention increased sensitivity across all stimulus strengths, but differentially impacted criterion, depending on stimulus strength. **c)** The slopes of the zROC did not deviate substantially from 1, indicating equal variance assumptions were not violated in most cases, apart from at the highest stimulus strengths. Data (total n=60; Experiments 1 and 3 each n=30) are presented as mean values ±1 SEM.



Supplementary Figure 7. Consistency of inattentional inflation across participants. Each dot shows an individual participant's AUC calculated from trials in which the attention cue was invalid (y-axis) vs. valid (x-axis). For most participants, the AUC for stimulus visibility (light green) was larger when the cue was invalid (i.e., above the unity line), indicating inattentional inflation. The AUC for feature visibility (dark green) was also larger under inattention for most participants, but only in Experiments 1 and 3 in threshold regimes.



Supplementary Figure 8. AMI of AUC replicated across analytic pipelines. The attentional modulation index of the AUC, quantifying the degree and direction with which objective and subjective reports dissociate with attention, as analyzed by two separate analytic pipelines, developed independently at the two study sites, BU (top) and UCI (bottom). AMIs greater than zero indicate inattentional inflation and less than zero indicate inattentional deflation. AMIs did not significantly differ by pipeline (all p>0.657, Supplementary Table 16). Thus, all measures of inattentional inflation were simultaneously replicated by two independent analytic pipelines and robust to any idiosyncrasies within pipelines. Data (total n=118; Experiments 1-3 each n=30 per experiment, Experiment 4 n=28) are presented as mean values ±1 SEM.

Supplementary Notes

Supplementary Note 1

This experiment was conceived as part of an adversarial collaboration between first-order and higher-order theories of consciousness, as described in the preregistration[50]. The specific theories under consideration were Recurrent Processing Theory[42] (RPT), a first-order theory, and two higher-order theories: Perceptual Reality Monitoring theory[36] (PRM) and Higher-Order Representation of a Representation theory[34,35] (HOROR).

The full adversarial collaboration includes two suites of experiments, involving subjective inflation and change blindness, each with two phases: psychophysics and fMRI. The current manuscript describes the outcomes of the psychophysics phase of the subjective inflation experiments. As described in the preregistered predictions table[50], only the fMRI phase has the potential to pose a serious challenge to one theory that would require a revision of the theory—denoted as a “fail” outcome in the predictions table. However, the theorists representing each theory also made predictions for the behavioral phase. Here, predictions unsupported by experimental outcomes would prompt reconsideration of some aspect of the theory but would not invalidate a core theoretical component—denoted as a “challenge” outcome in the predictions table.

The theorists representing the higher-order theories tested here (PRM and HOROR) predicted that the behavioral experiments would show inattentional subjective inflation. An outcome in which no behavioral experiment showed inflation would have challenged the higher-order theories by invalidating a motivating pillar of these theories. If inflation had not been supported in the current rigorous test, it would have undermined the idea that subjective experience regularly exceeds objective performance, reducing the empirical basis for the theoretical separation of sensory processing and conscious experience inherent to higher-order theories.

The theorist representing the first-order theory tested here (RPT) predicted in contrast that the behavioral experiments would not show inattentional subjective inflation. This prediction stemmed from the absence of a mechanism within RPT to generate inflation: in RPT, recurrent sensory processing is proposed to generate both conscious experience and performance in sensory tasks. However, given that inflation can in principle be generated by first-order signal detection models (see Discussion), a behavioral finding of inflation would not invalidate a core theoretical component of RPT. Rather, the current findings of robust inflation prompt RPT and other first-order theories to develop or incorporate more explicit mechanisms to account for decouplings of objective and subjective reports.

Supplementary Note 2

Overview of the criterion analysis of Figure 4

In **Figure 4a**, we plot the difference between detection criteria for attended and unattended trials as a function of stimulus strength, following the similar analysis in Figure 2b of Rahnev et al.[21] for their Experiment 2. Rahnev et al.[21] probed detection of peripherally presented gratings at four levels of near-threshold grating contrast when these stimuli were either validly or invalidly cued, making their experimental design similar to that of our Experiment 1. The results in our **Figure 4a** replicate those of Rahnev et al.'s[21] Figure 2b in showing that 1) detection criteria for peripheral gratings are higher (more conservative) for attended trials than for unattended trials at low grating contrasts, and 2) the difference between criteria for attended vs. unattended trials decreases as grating contrast increases. Moreover, our results confirm a prediction made by Rahnev et al.'s[21] computational model of their data by showing that if high enough contrasts are probed, the criterion difference reverses, such that detection criteria are increasingly lower (more liberal) for attended trials as contrast increases—though that study did not note that prediction and instead focused on the finding that attention induces conservative detection criteria. We also extend the results of Rahnev et al.[21] by 1) showing these criterion patterns hold not just for stimulus detection (i.e., reporting that a grating was visible), but also for feature detection (i.e., reporting that the grating's tilt was visible); and 2) showing that the same patterns for stimulus and feature detection criteria hold for texture-defined ovals (Experiment 3).

Distinguishing the criterion in yes-no discrimination tasks vs. two-response classification tasks

The analysis of **Figure 4a** requires important conceptual context. Our Experiment 1 uses a detection task featuring a single set of target-absent trials (corresponding to grating contrast = 0) and multiple sets of target-present trials (corresponding to contrasts > 0). The observer must provide two binary classifications for each stimulus, corresponding to stimulus detection (“saw grating” vs. “did not see grating”) and feature detection (“saw grating tilt” vs. “did not see grating tilt”). For simplicity, in the following discussion we will consider only the stimulus detection task, but all considerations similarly apply to modeling the feature detection task. In signal detection theory (SDT), the stimulus detection task of Experiment 1 is treated as a *two-response classification task*[61] in which the observer sets a single criterion on a decision axis to classify more than two stimulus categories (here, target-absent trials and target-present trials at multiple contrasts) into two classes.

Importantly, the criterion in two-response classification tasks is computed and interpreted differently from the criterion used to model the more common *yes-no discrimination task*[61]. Here we clarify this distinction to prevent potential confusions that might arise from failing to do so, and to facilitate correct interpretation of our analysis in **Figure 4a**.

In the yes-no discrimination task, the observer is presented with a stimulus from one of two stimulus categories (e.g., target-absent and target-present) and must provide a binary classification (e.g., “yes, saw target” or “no, did not see target”). In the yes-no SDT model, the two stimulus categories generate normal distributions of evidence along some decision axis. The observer sets a criterion on the decision axis such that they respond “yes” for any trial yielding evidence that surpasses the criterion, and “no” otherwise. Let us call the criterion for the yes-no discrimination task the *yes-no criterion*, or c_{YN}. Its formula is given by

$$c_{\mathrm{YN}} = -0.5\,(z(H) + z(F)) \tag{S1}$$

where H and F correspond to hit rate and false alarm rate, respectively, and z is the inverse of the normal CDF. c_{YN} is measured relative to a coordinate system whose zero occurs at the location on the decision axis where the two stimulus distributions intersect (i.e., have equal likelihood). This is a point of zero response bias in the sense that setting the criterion here yields an equal error rate for "yes" and "no" responses (i.e., equal false alarm rate and miss rate). It follows that the sign of the yes-no criterion can be interpreted in terms of bias in the decision-making strategy. Setting the criterion above the zero-bias point ($c_{\mathrm{YN}} > 0$) is a conservative strategy that prioritizes decreasing false alarm rate at the expense of increasing miss rate, whereas setting it below the zero-bias point ($c_{\mathrm{YN}} < 0$) is a liberal strategy that prioritizes decreasing miss rate at the expense of increasing false alarm rate.

The SDT model of the two-response classification task is identical to that of the yes-no task, except that there are more than two stimulus distributions. It follows that for this model, there is not a unique zero-bias point. Every possible pairing of stimulus distributions yields a different location at which the paired distributions intersect, and these correspond to different zero-bias points that are specific to each pairing. It follows that the criterion used to model such tasks, the *two-response classification criterion* or c_{2RC}, cannot be computed and interpreted relative to a zero-bias point in a way analogous to the yes-no criterion; instead, a new choice must be made for the zero of the measurement scale. One reasonable approach is to compute the location of c_{2RC} relative to the mean of the noise distribution[a]:

$$c_{2RC} = -z(F) \tag{S2}$$

Whatever convention for the zero point is chosen, the sign of c_{2RC} cannot be interpreted as reflecting bias in the decision-making strategy in the same way as c_{YN}.

Interpreting yes-no criterion effects in Experiment 1

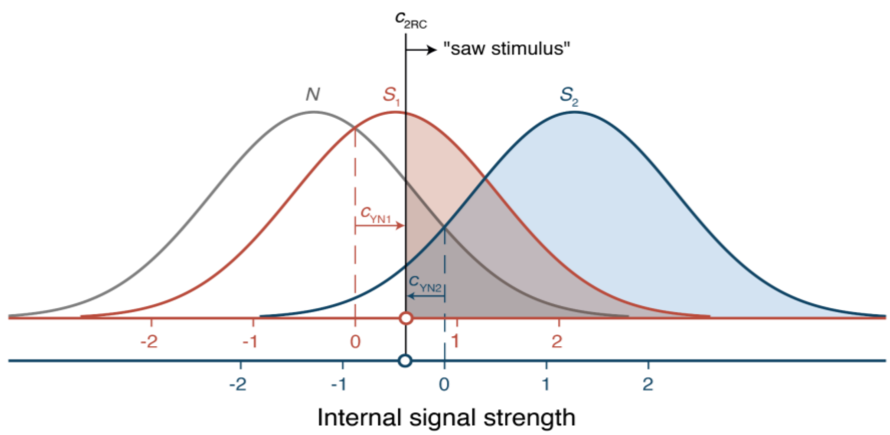
Bearing these distinctions between c_{YN} and c_{2RC} in mind, we nonetheless chose to analyze the criterion for the two-response classification task of Experiment 1 using a yes-no analysis framework. We explain the rationale for this choice in the following section; here, we examine why this choice leads to results for Experiment 1 that are trivial and potentially misleading if not properly understood.

Applying a two-response classification SDT model to Experiment 1 would involve computing a single c_{2RC} value for each attention condition of each participant, where this value corresponds to the single criterion that determines stimulus detection responses for all stimuli. Instead, for each attention condition of each participant, we applied the yes-no discrimination SDT model

[a] Technically, a two-response classification task does not necessarily have to include target-absent or "noise" trials. In this case, the criterion could be computed relative to the mean of the weakest stimulus distribution, with p("yes") for these stimuli being the analogue of false alarm rate.

separately for all possible pairings of the target-present trials at each level of contrast with the common set of target-absent trials. This yielded separate values of c_{YN} at every contrast. We observed that c_{YN} decreased with contrast (**Supplementary Figure 5b**), and that the rate of this decrease depended on attention (**Figure 4a**).

Taken at face value, this would seem to suggest that participants employed an increasingly liberal criterion-setting strategy as contrast increased, and that this criterion setting effect was modulated by attention. However, this cannot possibly be the case, since each c_{YN} within an attention condition is computed from the same false alarm rate arising from the same set of target-absent trials, and therefore must correspond to the *same criterion*, i.e., the fixed two-response classification criterion. This single underlying criterion is assigned different values in the pairwise yes-no analyses due to the fact that the zero-bias point for a given pairing of noise and signal distributions (i.e., the location where they intersect) differs for each level of contrast (**Supplementary Figure 9**). As contrast increases, so does the mean of the corresponding target-present distribution, causing a rightward shift in the zero-bias point and a corresponding decrease in the computed value for c_{YN}. Thus, within each attention condition, the observer sets only one criterion which applies across all contrasts (fixed c_{2RC}), and this single criterion exhibits different relationships to the zero-bias point for each pairing of noise and signal distributions (yielding different c_{YN} values)—but these do not reflect substantive changes in criterion setting *per se*.



Supplementary Figure 9. Signal detection theory (SDT) analysis schematic. Schematic for SDT analysis of a two-response classification task with one set of target-absent trials (corresponding to the noise distribution N) and two sets of target-present trials with different stimulus strengths (corresponding to the signal distributions S_1 and S_2). A single two-response classification criterion c_{2RC} determines one false alarm rate from N and two hit rates from S_1 and S_2, respectively. However, when SDT analysis for a yes-no discrimination task is conducted separately for each stimulus strength, it yields two different values (c_{YN1} and c_{YN2}) for the location of the same underlying criterion c_{2RC}. These values differ because the yes-no criterion is measured relative to the location where the noise and signal distributions intersect, and this zero-bias point depends on the mean of the signal distribution, which changes with signal strength (compare the red and blue axis coordinates derived from S_1 and S_2, respectively). Here, these different coordinate systems render c_{2RC} as positive (i.e., conservative) in c_{YN1} and as negative (i.e., liberal) in c_{YN2}. As the signal distribution mean increases, the zero-bias point shifts rightwards, leading to increasingly negative values for the yes-no criterion c_{YN} despite the underlying two-response classification criterion c_{2RC} being fixed. Thus, although measurements for c_{YN} become more negative as stimulus strength increases, this is best understood not as an actual shift in criterion setting *per se*, but rather as a systematic change in how the fixed criterion c_{2RC} relates to the changing zero-bias point as sensitivity increases.

Because these changes in c_{YN} are driven entirely by changes in signal distribution mean, there is a simple relationship between c_{YN} and sensitivity (*d'*). With every unit increase in *d'*, the zero-bias point shifts rightwards by 0.5 units (since it is located at the mean of the noise and signal distribution means). This entails that c_{YN} is linearly related to *d'* with a slope of -0.5 (**Supplementary Figure 5c**, top row). As a consequence, changes in c_{YN} with contrast trivially reflect changes in *d'* with contrast (compare blue plots in **Figure 4** panels **a** and **b**; also top row of **Supplementary Figure 5** panels **a** and **b**). Thus, in two-response classification tasks, not only do changes of c_{YN} with contrast not reflect changes in criterion setting, in fact they are best understood as sensitivity effects insofar as they indirectly reflect changes in *d'*. Likewise, the decreasing difference between c_{YN} for attended and unattended trials with increasing contrast (**Figure 4a**, blue plots) does not reflect an attentional modulation of criterion setting across contrasts, but rather reflects that *d'* increases more rapidly with contrast for attended stimuli (**Figure 4b**, blue plots).

Interpreting yes-no criterion effects in Experiment 3 and Rahnev et al. (2011)

If the c_{YN} results for Experiment 1 do not inform us about criterion setting but rather are trivial reflections of the effects of contrast and attention on *d'*, what motivates using the yes-no analysis framework here? The answer is that this analysis as applied to Experiment 1 provides a benchmark for comparison to the c_{YN} results of our Experiment 3 and Rahnev et al.'s[21] Experiment 2. For reasons discussed below, the c_{YN} analysis is not trivial *a priori* for either of these experiments. However, if these experiments nonetheless exhibit patterns in the c_{YN} results similar to Experiment 1, this would suggest that the findings of these experiments might best be interpreted in a similar way—i.e., as reflecting not criterion-setting effects *per se*, but rather sensitivity effects.

Although the stimulus detection task for texture-defined oval stimuli in Experiment 3 is structurally similar to the task of Experiment 1 in many ways, for the purposes of the present discussion there is an important difference. In Experiment 3, the presence or absence of the oval is determined by line orientation, which is independent of the stimulus strength manipulation of line length. It follows that each level of line length has its own set of target-absent trials (all lines parallel) and target-present trials (figure lines orthogonal to ground lines). These data are most naturally analyzed not with a two-response classification SDT model as in Experiment 1, but rather with separate yes-no SDT models applied to each line length. Because the yes-no SDT analyses conducted at each line length do not use a common set of target-absent trials, it is not necessarily the case that the c_{YN} computed at each line length merely reflects a single underlying criterion measured relative to different coordinate systems, as is the case in Experiment 1. Rather, they may reflect real changes in criterion setting induced by perceptible changes in line length.

However, false alarm rates in Experiment 3 exhibited only slight changes as a function of line length (**Supplementary Figure 4**), approximating the constant false alarm rate from the common set of stimulus-absent trials in Experiment 1. This implies that changes in c_{YN} as a function of line length were driven primarily by changes in hit rate, i.e., by changes in the mean of the signal distribution with stimulus strength—again approximating the sensitivity-driven effects in Experiment 1. These patterns in the Experiment 3 data resulted in criteria being very nearly linear with sensitivity and with a slope of approximately -0.5 (**Supplementary Figure 5c**, bottom row), in close approximation to the (trivially) perfectly linear psychometric functions of slope -0.5 in Experiment 1 (**Supplementary Figure 5c**, top row).

These data suggest that although the c_{YN} patterns in Experiment 3 are not trivial *a priori* as they are in Experiment 1, nonetheless they closely approximate the patterns in Experiment 1 and therefore may be best understood in the same way—i.e., as reflecting effects of attention on sensitivity, not criterion setting. On this interpretation, the observer sets a fixed criterion for all line lengths within an attention condition. Provided that the means and variances of the noise distributions are roughly constant across line lengths, this fixed criterion would yield approximately constant false alarm rates[b] and increasing hit rates across line lengths. This would create a situation similar to that of the two-response classification task, where c_{YN} values decrease with stimulus strength as a side-effect of the zero-bias point increasing with *d'* despite the underlying criterion being fixed (**Supplementary Figure 9**).

Experiment 2 of Rahnev et al.[21] used a grating detection task similar to our Experiment 1, which we argued above is best modeled as a two-response classification task. However, there is an important difference in how these two experiments were conducted. In our Experiment 1, target-absent trials and target-present trials of all contrasts were interleaved randomly across trials. In Rahnev et al.'s[21] Experiment 2, each block of trials contained a mix of target-absent trials and target-present trials of a fixed contrast; grating contrast varied across but not within different blocks. This design provides a natural way to pair separate sets of target-absent trials with target-present trials at each contrast, which justifies treating the data as a series of independent yes-no tasks at each contrast rather than as an omnibus two-response classification task using a common set of target-absent trials. In turn, this structure allows for the possibility that changes in c_{YN} with contrast reflect real changes in criterion setting. For instance, observers might adjust their detection criterion in each block due to perceptible differences in grating contrast. Nonetheless, Rahnev et al.[21] observed that false alarm rates for attended and unattended stimuli were roughly constant across contrasts (their Supplementary Figure 4), suggesting that the effects of contrast and attention on c_{YN} they observed (their Figure 2b) may be better understood as an indirect reflection of effects of contrast and attention on *d'*, per the above considerations.

In fact, this idea is well accommodated by the analysis of Rahnev et al.[21] Although they analyzed their empirical data in terms of c_{YN} effects, they proposed a deeper computational model that explains the data in terms of how attention influences sensitivity, not criterion. According to their model, attention both boosts signal magnitude and decreases the variance of perceptual evidence. Under the assumption that the same criterion is used for attended and unattended stimuli, this model predicts that false alarm rates are higher under inattention[c]. Provided that the model parameters are tuned in the appropriate way, the model can yield a higher c_{YN} value for attended stimuli at low stimulus strengths, in agreement with empirical data. Furthermore, given the model assumptions that 1) the distribution means for unattended stimuli increase more slowly with stimulus strength than those for attended stimuli and 2) the distribution variances for unattended stimuli are always higher than variances for attended ones at each stimulus strength, it follows that sensitivity (*d'*) increases with stimulus strength more slowly for unattended stimuli. The model's assumption of a fixed criterion entails that c_{YN} decreases with stimulus strength for all attention conditions, and its prediction that *d'* increases

---

[b] The slight increase in stimulus detection false alarm rates with line length (**Supplementary Figure 4**, left panel) could arise from a fixed criterion if the mean and/or variance of the noise distributions increases slightly with line length.

[c] Provided that the criterion location exceeds the mean of the noise distribution, i.e., provided that false alarm rates are less than 0.5.

with stimulus strength more slowly for unattended stimuli entails that the decrease of c_{YN} with stimulus strength will correspondingly be slower for unattended stimuli. The slower rate of decrease for c_{YN} under inattention entails that the difference between criteria for attended and unattended stimuli decreases as stimulus strength increases from low levels, eventually yielding a crossover such that c_{YN} is lower for attended stimuli at high enough stimulus strengths. This is consistent with the pattern in our data (**Figure 4a**) as well as Rahnev et al.'s[21] data (their Figure 2b) and model fits (their Supplementary Figure 6b, where the model's predicted crossover begins to emerge at the highest plotted stimulus strength, in a straightforward extrapolation of the patterns observed at lower stimulus strengths.)

Since Rahnev et al.'s[21] experiment used low-contrast gratings, their empirical results stopped just short of exhibiting this crossover effect and their interpretation emphasized the effect of attention on making c_{YN} conservative at low stimulus strengths. However, importantly, their computational model explained these patterns in c_{YN} as arising from processes best captured by a more complex SDT-based model in which criterion is fixed and attention modulates sensitivity via effects on signal strength and variance. Thus, their modeling approach aligns with the above considerations that these patterns are best understood as effects of stimulus strength and attention on sensitivity rather than criterion-setting.

Although Rahnev et al.[21] modeled a single fixed criterion across all stimulus strengths and attention conditions, their data and ours could also be consistent with a criterion that is fixed across stimulus strengths within each attention condition, but varies with attention[69]. The present dataset will provide an opportunity for further testing and development of SDT models to better understand how attention interacts with perception and perceptual decision making.

Supplementary Note 3

Here, we include discussion of one author's (B.M.'s) experiences of this task that may suggest an alternative framing for subjective inflation. We include these observations as they may be helpful in informing future research—both theoretical and empirical.

Subjective inflation is typically framed as the surprising and unintuitive finding that, for matched levels of feature discrimination performance, subjective visibility is higher when not attending to the stimulus. However, turning this framing on its head yields the potentially more intuitive idea that, at matched levels of subjective visibility, feature discrimination performance is higher when attending to the stimulus. The general idea that strong overall visibility does not necessarily entail fine perceptual discrimination is familiar from the everyday observation that stimuli perceived in the visual periphery can feel highly prominent and yet elude attempts to discern fine-grained stimulus features. This suggests a more general principle that in impoverished viewing conditions, the subjective experience of perceptual prominence does not necessarily guarantee the subjective experience of perceptual sharpness.[d] It is perhaps not so counterintuitive to suppose that this principle might apply to cases of inattention, making fine discrimination of a strongly visible unattended stimulus more difficult than it would be for a similarly visible attended stimulus.

This framing accords well with the subjective experience of one of the authors (B.M.) when testing the task for suprathreshold texture-defined ovals (Experiment 4). Considering trials where attended and unattended stimuli appeared similar to the reference strength and thus had similarly strong levels of overall "pop-out" from the background (i.e., similar levels of "perceptual prominence"), this matched level of stimulus visibility did not feel nearly as useful for making fine discrimination judgments about the orientation of its nearly circular shape in the absence of attention (i.e., different levels of "perceptual sharpness"). This difference stood out as being strikingly obvious at the single trial level in a way that was not the case for the threshold stimuli. This is likely due to the fact that, to achieve threshold discrimination performance for clearly visible suprathreshold stimuli, the differences in stimulus features to be discriminated must be far more subtle than they are for stimuli near the detection threshold, which has the effect of making discrimination for suprathreshold stimuli feel far more challenging. This configuration maximizes the contrast between these dissociable aspects of perception, with detection being very easy and discrimination being very hard, which in turn may be a formula for making inflation effects stand out as much as possible. The net effect of all this was to make inflation effects (framed as inattentional deficits in feature discrimination for matched-visibility stimuli) seem obvious to the author at the single-trial level in a way that felt analogous to how one can clearly detect an object in the visual periphery and yet struggle to discern its fine-grained features.

[d] "Perceptual prominence" and "perceptual sharpness" are introduced here as terms to refer to phenomenological aspects of visual experience that are familiar from everyday life. The subjective reports participants made about seeing the stimulus (Experiments 1 and 3) or judging stimulus strength relative to a reference (Experiments 2 and 4) can be seen as operationalizations of the participants' experiences of the perceptual prominence of the response-cued stimuli in these experiments, and similarly reports about seeing stimulus features (Experiments 1, 2, and 3) can be seen as operationalizations of their experiences of the perceptual sharpness of these stimuli. These phenomenological and operational concepts are distinct from, but related to, the computational concepts of signal strength and signal-to-noise ratio in the model discussed below.

Of course, this anecdotal experience is entirely consistent with the data analyzed in the main manuscript, taking "perceptual prominence" and "perceptual sharpness" to correspond to overall stimulus visibility and feature visibility, respectively. Consider the plot of subjective stimulus visibility vs. objective feature discrimination performance for Experiment 4 (**Figure 2c**, bottom panel). Taking any horizontal slice through this plot shows that at matched levels of subjective visibility, discrimination performance is considerably higher when the stimulus is attended than when unattended. Although feature visibility data were not collected for Experiment 4, feature visibility data for Experiment 2 (which used suprathreshold gratings) were closely related to discrimination accuracy in a way that did not depend on attention (**Figure 3c**, middle panel). Provided that feature visibility data in Experiment 4 would have been similar, these results jointly demonstrate that for matched levels of overall stimulus visibility for suprathreshold stimuli, subjective feature visibility (and objective feature discrimination) is considerably higher under attention.

The motivation for including this anecdotal report is not the experienced phenomena per se, as these are entirely consistent with effects already demonstrated or suggested in the data. Rather, what is notable is the force with which single trial observations of suprathreshold inflation stimuli suggested to the author how this alternative framing of inflation effects naturally and intuitively accords with similar experiences familiar from everyday, suprathreshold vision—contra the typical framing of inflation effects as unintuitive and surprising. It is possible that deeper consideration of this alternative framing could lead to insightful new research angles and perhaps stimulate new paradigms for understanding inflation effects and their relationship to similar phenomena in everyday suprathreshold vision.

This framing accords well with the signal detection model of Rahnev et al.[21] As explained in greater detail in the Discussion and in **Supplementary Note 2**, this model explains inflation effects as resulting (in part) from higher levels of noise in perceptual representations of unattended stimuli. At matched levels of signal-to-noise ratio (and so, matched levels of objective performance), the noisier perceptual evidence for unattended stimuli has higher overall magnitude (and so higher subjective visibility). But turning this framing on its head per the above, the model also predicts that at matched levels of subjective visibility, perceptual evidence for unattended stimuli has similar overall magnitude to those of attended stimuli but is noisier, leading to lower signal-to-noise ratio for feature discrimination and thus poorer discrimination performance.

The dissociable mechanisms of overall evidence magnitude and signal-to-noise ratio in the model suggest an analogy with the dissociable subjective experiences of perceptual prominence and perceptual sharpness discussed above. Perceptual prominence maps naturally onto evidence magnitude, as both are magnitude-based concepts. Perceptual sharpness might also seem to map naturally onto signal-to-noise ratio as both are precision-based concepts, but further consideration raises some complexities. The signal detection model attributes feature detection reports to perceptual evidence exceeding a feature detection criterion, and so would seem to associate perceptual sharpness with evidence magnitude rather than signal-to-noise ratio per se. Additionally, signal-to-noise ratio in signal detection theory (SDT) pertains to signal precision across trials, whereas perceptual sharpness pertains to the precision of subjective experience for a single percept (a within-trial phenomenon), and SDT does not model within-trial noise.

Despite these conceptual and computational distinctions, it may be the case that within-trial perceptual sharpness and across-trial signal-to-noise ratio are closely related for suprathreshold stimuli, as we observed that in this regime reports of feature detection were closely correlated with objective feature discrimination performance in a way that was not modulated by attention (**Figure 3c**, middle panel). Ideally, a model of these phenomena would capture this empirical relationship while simultaneously resolving the associated conceptual and computational tensions noted above. For instance, such a model might link perceptual sharpness to within-trial uncertainty in perceptual processing, and in turn link this within-trial uncertainty to across-trial signal-to-noise ratio and thus objective discrimination performance. Alternatively, if one accepts that perceptual sharpness can be adequately characterized as a magnitude-based concept, or provides an account of how evidence magnitude and within-trial uncertainty are linked, then principles from the signal detection model of Rahnev et al.[21] might be sufficient to characterize our findings in the suprathreshold stimulus experiments without raising these conceptual tensions. One natural way to link evidence magnitude to within-trial uncertainty in SDT is via the likelihood ratio between the signal and noise distributions occurring at a given value of evidence magnitude, as evidence associated with a higher likelihood ratio entails better differentiation between signal and noise. However, the Rahnev et al.[21] model posits that decision criteria apply to raw evidence magnitude despite the likelihood ratio of these magnitudes differing across attention conditions, which raises complications for linking the two.

Although the modeling ideas discussed above may have intuitive appeal, a more definitive account would require extending any candidate model to simultaneously account for stimulus detection (related to “perceptual prominence”), feature detection (related to “perceptual sharpness”), and objective discrimination performance across all levels of attention and stimulus strength, and demonstrating that the model in question can account for the data well while also outperforming competing models.

Supplementary Methods

Task instruction excerpts

Experiment 1: Threshold detection of gratings

In this experiment you will be asked to make judgments about visual stimuli. Specifically, you will be pointed or "cued" to a particular location on the screen and will have to respond:

1. Did you see a grating embedded in the noise patch in the cued location?
2. If yes, did you see what direction the grating was oriented?
3. Was the grating oriented counterclockwise (-45 deg) or clockwise (+45 deg) from vertical?

We ask separate questions about whether you saw a grating and whether you saw its orientation because sometimes you may clearly see some kind of grating in the noise, without being able to see clearly the exact orientation of this grating.

About the "saw grating" / "saw orientation" judgments

> When we ask about whether you saw a grating and its orientation, we're interested to know about what your actual visual experience was like. Did it actually look like there was a grating embedded in the noise patch in the response-cued location? If so, could you actually see whether the grating was oriented counterclockwise or clockwise from vertical?
>
> Don't try to answer this question based on any information other than what your actual visual experience of the patch was like.

About the "counterclockwise" / "clockwise" judgment:

> Sometimes you may not be very sure whether the grating was oriented counterclockwise or clockwise from vertical. This may be especially the case when you didn't see a grating to begin with!
>
> On trials where the response-cued quadrant didn't contain a grating to begin with, obviously the counterclockwise / clockwise judgment has no meaning. However, just because you didn't see a grating doesn't mean there wasn't one there! In cases where a grating was objectively present but you didn't see it, the orientation judgment is still meaningful and your response may be meaningful too, even if it "feels" like a wild guess.
>
> For this reason, we ask that you take the counterclockwise / clockwise question seriously on each trial and give the best response you can, even when you didn't see a grating. On such trials, try to make the orientation judgment *as if* a grating had been presented but you just didn't see it. If you're not sure about the grating's orientation, just make your best gut instinct guess without spending too much time deliberating.
>
> If you have to guess, it's important that you don't guess based on a strategy (like always responding the opposite of what you chose last, or always responding "counterclockwise"). Try to keep your "counterclockwise" and "clockwise" guesses roughly balanced, based on your best hunch. If you have no idea at all, try to pick randomly, as if you were flipping a coin.

Stimulus frequencies and dependencies

> Overall, each quadrant is equally likely to contain a grating. When a grating is present, it is equally likely to be oriented counterclockwise or clockwise from vertical.
>
> It is very important to note that grating presence and orientation in each quadrant is completely independent of grating presence and orientation in the other quadrants. In other words, knowing what was shown in one quadrant gives you no information whatsoever about what was shown in any other quadrant.

As a consequence, your responses should always be based *only* on what you saw at the response-cued quadrant, and should never be influenced by what you saw at any of the quadrants that were not cued.

Supplementary Tables

Supplementary Table 1. Objective performance (related to Figure 2a).

	Effect	DF_n	DF_d	*F*	*p*	η^2_G	ε	*p*[GG]
	All experiments							
1	Site	1	111	5.87	**0.017**	0.02	-	-
2	Exp	3	111	2.69	0.056	0.03	-	-
3	Strength	6	666	748.05	**<0.001**	0.61	0.48	**<0.001**
4	Att	2	222	892.80	**<0.001**	0.52	0.70	**<0.001**
5	Site:Exp	3	111	2.67	0.051	0.03	-	-
6	Site:Strength	6	666	0.18	0.983	<0.01	0.48	0.907
7	Exp:Strength	18	666	13.79	**<0.001**	0.08	0.48	**<0.001**
8	Site:Att	2	222	2.78	0.064	<0.01	0.70	0.084
9	Exp:Att	6	222	14.78	**<0.001**	0.05	0.70	**<0.001**
10	Strength:Att	12	1332	53.93	**<0.001**	0.08	0.75	**<0.001**
11	Site:Exp:Strength	18	666	0.93	0.538	0.01	0.48	0.495
12	Site:Exp:Att	6	222	0.48	0.820	<0.01	0.70	0.756
13	Site:Strength:Att	12	1332	2.05	**0.018**	<0.01	0.75	**0.032**
14	Exp:Strength:Att	36	1332	4.45	**<0.001**	0.02	0.75	**<0.001**
15	Site:Exp:Strength:Att	36	1332	1.78	**0.003**	0.01	0.75	**0.009**
	Experiment 1							
1	Site	1	28	0.17	0.681	<0.01	-	-
2	Strength	6	168	158.04	**<0.001**	0.62	0.56	**<0.001**
3	Att	2	56	315.97	**<0.001**	0.64	0.69	**<0.001**
4	Site:Strength	6	168	1.01	0.418	0.01	0.56	0.396
5	Site:Att	2	56	1.65	0.202	<0.01	0.69	0.210
6	Strength:Att	12	336	18.56	**<0.001**	0.16	0.61	**<0.001**
7	Site:Strength:Att	12	336	1.54	0.109	0.02	0.61	0.154
	Experiment 2							
1	Site	1	28	11.39	**0.002**	0.11	-	-
2	Strength	6	168	249.79	**<0.001**	0.72	0.30	**<0.001**
3	Att	2	56	356.47	**<0.001**	0.70	0.63	**<0.001**
4	Site:Strength	6	168	0.41	0.873	<0.01	0.30	0.645
5	Site:Att	2	56	2.03	0.142	0.01	0.69	0.158
6	Strength:Att	12	336	30.10	**<0.001**	0.19	0.55	**<0.001**
7	Site:Strength:Att	12	336	3.35	**<0.001**	0.03	055	**0.003**
	Experiment 3							
1	Site	1	28	0.60	0.445	<0.01	-	-
2	Strength	6	168	219.99	**<0.001**	0.74	0.38	**<0.001**
3	Att	2	56	112.19	**<0.001**	0.37	0.62	**<0.001**
4	Site:Strength	6	168	0.80	0.573	0.10	0.38	0.469
5	Site:Att	2	56	0.73	0.487	<0.01	0.62	0.427
6	Strength:Att	12	336	12.90	**<0.001**	0.01	0.62	**<0.001**
7	Site:Strength:Att	12	336	1.48	0.131	0.01	0.62	0.173
	Experiment 4							
1	Site	1	28	3.30	0.080	0.08	-	-
2	Strength	6	168	148.88	**<0.001**	0.33	0.39	**<0.001**
3	Att	2	56	173.33	**<0.001**	0.38	0.75	**<0.001**
4	Site:Strength	6	168	0.59	0.741	<0.01	0.39	0.587
5	Site:Att	2	56	0.05	0.948	<0.01	0.75	0.904
6	Strength:Att	12	336	7.80	**<0.001**	0.03	0.51	**<0.001**
7	Site:Strength:Att	12	336	1.58	0.095	<0.01	0.51	0.154

Factors of experiment, attention, and stimulus strength are abbreviated as “Exp,” “Att,” and “Strength.” The statistical tests used in this table were repeated measures ANOVAs. For all ANOVA tables, when Mauchly’s test indicated violation of sphericity, Greenhouse-Geisser epsilon and corrected p-values are shown. Significant p-values are bolded.

Supplementary Table 2. Objective performance for stimuli reported as “unseen” (related to Supplementary Figure 3).

	Effect	DF_n	DF_d	*F*	p	η_G^2	ε	p[GG]
	Experiments 1 and 3							
1	Site	1	56	1.67	0.201	0.02	-	-
2	Expt	1	56	0.05	0.817	<0.01	-	-
3	Att	2	112	5.48	**0.005**	0.03	0.94	**0.006**
4	Site:Expt	1	56	0.05	0.818	<0.01	-	-
5	Site:Att	2	112	2.27	0.108	0.01	0.94	0.112
6	Expt:Att	2	112	0.41	0.664	<0.01	0.94	0.651
7	Site:Expt:Att	2	112	2.72	0.070	0.02	0.94	0.074

Supplementary Table 3. Subjective reports of stimulus visibility (related to Figure 2b).

	Effect	DF_n	DF_d	*F*	*p*	η^2_G	ε	*p*[GG]
	All experiments							
1	Site	1	111	0.13	0.720	<0.01	-	-
2	Exp	3	111	87.35	**<0.001**	0.52	-	-
3	Strength	6	666	544.71	**<0.001**	0.65	0.34	**<0.001**
4	Att	2	222	82.33	**<0.001**	0.05	0.70	**<0.001**
5	Site:Exp	3	111	0.19	0.902	<0.01	-	-
6	Site:Strength	6	666	0.57	0.753	<0.01	0.34	0.570
7	Exp:Strength	18	666	34.72	**<0.001**	0.26	0.34	**<0.001**
8	Site:Att	2	222	2.57	0.079	<0.01	0.70	0.099
9	Exp:Att	6	222	25.78	**<0.001**	0.05	0.70	**<0.001**
10	Strength:Att	12	1332	67.52	**<0.001**	0.05	0.53	**<0.001**
11	Site:Exp:Strength	18	666	1.51	0.081	0.02	0.34	0.175
12	Site:Exp:Att	6	222	3.22	**0.005**	<0.01	0.70	**0.013**
13	Site:Strength:Att	12	1332	2.07	0.016	<0.01	0.53	0.050
14	Exp:Strength:Att	36	1332	13.93	**<0.001**	0.03	0.53	**<0.001**
15	Site:Exp:Strength:Att	36	1332	0.75	0.859	<0.01	0.53	0.769
	Experiment 1							
1	Site	1	28	0.02	0.889	<0.01	-	-
2	Strength	6	168	168.54	**<0.001**	0.57	0.31	**<0.001**
3	Att	2	56	98.32	**<0.001**	0.23	0.61	**<0.001**
4	Site:Strength	6	168	0.11	0.998	<0.01	0.31	0.883
5	Site:Att	2	56	8.13	**0.001**	0.02	0.61	**0.005**
6	Strength:Att	12	336	24.87	**<0.001**	0.07	0.50	**<0.001**
7	Site:Strength:Att	12	336	1.02	0.434	<0.01	0.50	0.416
	Experiment 2							
1	Site	1	28	0.84	0.367	0.01	-	-
2	Strength	6	168	220.87	**<0.001**	0.81	0.28	**<0.001**
3	Att	2	56	20.41	**<0.001**	0.02	0.81	**<0.001**
4	Site:Strength	6	168	0.20	0.977	<0.01	0.28	0.783
5	Site:Att	2	56	2.32	0.108	<0.01	0.81	0.119
6	Strength:Att	12	336	22.76	**<0.001**	0.08	0.26	**<0.001**
7	Site:Strength:Att	12	336	0.55	0.885	<0.01	0.26	0.660
	Experiment 3							
1	Site	1	28	0.06	0.807	<0.01	-	-
2	Strength	6	168	171.40	**<0.001**	0.73	0.37	**<0.001**
3	Att	2	56	22.98	**<0.001**	0.08	0.59	**<0.001**
4	Site:Strength	6	168	1.34	0.240	0.02	0.37	0.269
5	Site:Att	2	56	0.90	0.414	<0.01	0.59	0.368
6	Strength:Att	12	336	17.98	**<0.001**	0.06	0.43	**<0.001**
7	Site:Strength:Att	12	336	1.52	0.115	<0.01	0.43	0.185
	Experiment 4							
1	Site	1	28	0.13	0.721	<0.01	-	-
2	Strength	6	168	102.84	**<0.001**	0.60	0.23	**<0.001**
3	Att	2	56	2.29	0.111	0.01	0.79	0.124
4	Site:Strength	6	168	2.65	0.018	0.04	0.23	0.102
5	Site:Att	2	56	0.18	0.834	<0.01	0.79	0.782
6	Strength:Att	12	336	46.17	**<0.001**	0.09	0.38	**<0.001**
7	Site:Strength:Att	12	336	1.33	0.199	<0.01	0.38	0.259

Supplementary Table 4. Subjective reports of task-relevant feature visibility (related to Figure 3b).

	Effect	DF_n	DF_d	*F*	*p*	η^2_G	ε	*p*[GG]
	All experiments							
1	Site	1	84	0.01	0.915	<0.01	-	-
2	Exp	2	84	11.92	**<0.001**	0.15	-	-
3	Strength	6	504	457.96	**<0.001**	0.53	0.36	**<0.001**
4	Att	2	168	327.53	**<0.001**	0.27	0.63	**<0.001**
5	Site:Exp	2	84	0.47	0.626	0.01	-	-
6	Site:Strength	6	504	0.37	0.888	<0.01	0.36	0.699
7	Exp:Strength	12	504	7.15	**<0.001**	0.03	0.36	**<0.001**
8	Site:Att	2	168	0.50	0.606	<0.01	0.63	0.523
9	Exp:Att	4	168	13.83	**<0.001**	0.03	0.63	**<0.001**
10	Strength:Att	12	1008	71.71	**<0.001**	0.06	0.48	**<0.001**
11	Site:Exp:Strength	12	504	0.89	0.561	<0.01	0.36	0.480
12	Site:Exp:Att	4	168	0.52	0.724	<0.01	0.63	0.641
13	Site:Strength:Att	12	1008	0.80	0.654	<0.01	0.48	0.567
14	Exp:Strength:Att	24	1008	7.81	**<0.001**	0.01	0.48	**<0.001**
15	Site:Exp:Strength:Att	24	1008	1.35	0.120	<0.01	0.48	0.189
	Experiment 1							
1	Site	1	28	0.56	0.461	0.83	-	-
2	Strength	6	168	209.62	**<0.001**	0.57	0.35	**<0.001**
3	Att	2	56	136.45	**<0.001**	0.27	0.67	**<0.001**
4	Site:Strength	6	168	1.07	0.385	<0.01	0.35	0.353
5	Site:Att	2	56	2.12	0.130	<0.01	0.67	0.148
6	Strength:Att	12	336	37.09	**<0.001**	0.01	0.45	**<0.001**
7	Site:Strength:Att	12	336	1.19	0.287	<0.01	0.45	0.315
	Experiment 2							
1	Site	1	28	0.19	0.669	<0.01	-	-
2	Strength	6	168	116.73	**<0.001**	0.41	0.25	**<0.001**
3	Att	2	56	116.52	**<0.001**	0.39	0.59	**<0.001**
4	Site:Strength	6	168	0.64	0.702	<0.01	0.25	0.493
5	Site:Att	2	56	0.01	0.989	<0.01	0.59	0.941
6	Strength:Att	12	336	21.46	**<0.001**	0.05	0.26	**<0.001**
7	Site:Strength:Att	12	336	0.98	0.468	<0.01	0.26	0.408
	Experiment 3							
1	Site	1	28	0.23	0.637	<0.01	-	-
2	Strength	6	168	153.98	**<0.001**	0.60	0.38	**<0.001**
3	Att	2	56	101.66	**<0.001**	0.16	0.67	**<0.001**
4	Site:Strength	6	168	0.57	0.751	<0.01	0.38	0.590
5	Site:Att	2	56	0.03	0.972	<0.01	0.67	0.922
6	Strength:Att	12	336	29.59	**<0.001**	0.06	0.59	**<0.001**
7	Site:Strength:Att	12	336	1.37	0.178	<0.01	0.59	0.219

Supplementary Table 5. Stimulus detection sensitivity (related to Supplementary Figure 5a).

	Effect	DF_n	DF_d	*F*	*p*	η_G^2	ε	*p*[GG]
	All experiments							
1	Site	1	56	1.40	0.241	0.01	-	-
2	Exp	1	56	2.81	0.099	0.03	-	-
3	Strength	6	336	350.59	**<0.001**	0.59	0.42	**<0.001**
4	Att	2	112	296.54	**<0.001**	0.40	0.74	**<0.001**
5	Site:Exp	1	56	<0.01	0.944	<0.01	-	-
6	Site:Strength	6	336	0.44	0.854	<0.01	0.42	0.693
7	Exp:Strength	6	336	10.97	**<0.001**	0.04	0.42	**<0.001**
8	Site:Att	2	112	0.35	0.703	<0.01	0.74	0.639
9	Exp:Att	2	112	3.13	0.048	<0.01	0.74	0.063
10	Strength:Att	12	672	56.13	**<0.001**	0.10	-	-
11	Site:Exp:Strength	6	336	0.59	0.739	<0.01	0.42	0.594
12	Site:Exp:Att	2	112	0.32	0.725	<0.01	0.74	0.660
13	Site:Strength:Att	12	672	1.39	0.164	<0.01	-	-
14	Exp:Strength:Att	12	672	1.32	0.201	<0.01	-	-
15	Site:Exp:Strength:Att	12	672	2.30	**0.007**	<0.01	-	-
	Experiment 1							
1	Site	1	28	0.77	0.389	0.02	-	-
2	Strength	6	168	243.16	**<0.001**	0.60	0.38	**<0.001**
3	Att	2	56	186.97	**<0.001**	0.48	0.80	**<0.001**
4	Site:Strength	6	168	0.06	0.999	<0.01	0.38	0.956
5	Site:Att	2	56	0.67	0.517	<0.01	0.80	0.485
6	Strength:Att	12	336	45.43	**<0.001**	0.11	0.61	**<0.001**
7	Site:Strength:Att	12	336	1.15	0.319	<0.01	0.61	0.333
	Experiment 3							
1	Site	1	28	0.64	0.431	0.01	-	-
2	Strength	6	168	149.19	**<0.001**	0.59	0.40	**<0.001**
3	Att	2	56	115.62	**<0.001**	0.33	0.66	**<0.001**
4	Site:Strength	6	168	0.74	0.616	<0.01	0.40	0.503
5	Site:Att	2	56	0.03	0.966	<0.01	0.66	0.910
6	Strength:Att	12	336	21.49	**<0.001**	0.10	0.68	**<0.001**
7	Site:Strength:Att	12	336	2.14	**0.014**	0.01	0.68	**0.032**

Supplementary Table 6. Stimulus detection criterion (related to Supplementary Figure 5b).

	Effect	DF_n	DF_d	*F*	*p*	η^2_G	ε	*p*[GG]
	All experiments							
1	Site	1	56	0.22	0.641	<0.01	-	-
2	Exp	1	56	0.53	0.470	<0.01	-	-
3	Strength	6	336	352.02	**<0.001**	0.38	0.40	**<0.001**
4	Att	2	112	11.15	**<0.001**	0.01	0.63	**0.001**
5	Site:Exp	1	56	<0.01	0.987	<0.01	-	-
6	Site:Strength	6	336	2.36	0.030	<0.01	0.40	0.088
7	Exp:Strength	6	336	24.81	**<0.001**	0.04	0.40	**<0.001**
8	Site:Att	2	112	7.13	**0.001**	0.01	0.63	**0.006**
9	Exp:Att	2	112	9.35	**<0.001**	0.01	0.63	**0.002**
10	Strength:Att	12	672	61.69	**<0.001**	0.04	0.61	**<0.001**
11	Site:Exp:Strength	6	336	2.56	0.019	<0.01	0.40	0.071
12	Site:Exp:Att	2	112	1.90	0.155	<0.01	0.63	0.171
13	Site:Strength:Att	12	672	1.19	0.288	<0.01	0.61	0.308
14	Exp:Strength:Att	12	672	2.05	**0.018**	<0.01	0.61	**0.036**
15	Site:Exp:Strength:Att	12	672	0.90	0.550	<0.01	0.61	0.523
	Experiment 1							
1	Site	1	28	0.09	0.773	<0.01	-	-
2	Strength	6	168	243.16	**<0.001**	0.27	0.38	**<0.001**
3	Att	2	56	15.00	**<0.001**	0.04	0.64	**<0.001**
4	Site:Strength	6	168	0.06	0.999	<0.01	0.38	0.956
5	Site:Att	2	56	6.84	**0.002**	0.02	0.64	**0.008**
6	Strength:Att	12	336	45.43	**<0.001**	0.03	0.61	**<0.001**
7	Site:Strength:Att	12	336	1.15	0.319	<0.01	0.61	0.333
	Experiment 3							
1	Site	1	28	0.15	0.703	<0.01	-	-
2	Strength	6	168	171.11	**<0.001**	0.51	0.37	**<0.001**
3	Att	2	56	3.18	0.049	<0.01	0.61	0.076
4	Site:Strength	6	168	3.22	**0.005**	0.02	0.37	**0.041**
5	Site:Att	2	56	1.05	0.357	<0.01	0.61	0.328
6	Strength:Att	12	336	26.73	**<0.001**	0.06	0.48	**<0.001**
7	Site:Strength:Att	12	336	1.00	0.448	<0.01	0.48	0.425

Supplementary Table 7. Feature detection sensitivity (related to Supplementary Figure 5a).

	Effect	DF_n	DF_d	*F*	*p*	η^2_G	ε	*p*[GG]
	All experiments							
1	Site	1	56	<0.01	0.950	<0.01	-	-
2	Exp	1	56	1.81	0.184	0.01	-	-
3	Strength	6	336	397.97	**<0.001**	0.64	0.37	**<0.001**
4	Att	2	112	248.90	**<0.001**	0.44	0.72	**<0.001**
5	Site:Exp	1	56	<0.01	0.983	<0.01	-	-
6	Site:Strength	6	336	0.18	0.982	<0.01	0.37	0.856
7	Exp:Strength	6	336	6.77	**<0.001**	0.03	0.37	**0.001**
8	Site:Att	2	112	0.48	0.618	<0.01	0.72	0.556
9	Exp:Att	2	112	5.70	**0.004**	0.02	0.72	**0.010**
10	Strength:Att	12	672	47.89	**<0.001**	0.10	0.74	**<0.001**
11	Site:Exp:Strength	6	336	1.92	0.077	<0.01	0.37	0.146
12	Site:Exp:Att	2	112	1.93	0.150	<0.01	0.72	0.163
13	Site:Strength:Att	12	672	1.97	**0.025**	<0.01	0.74	**0.042**
14	Exp:Strength:Att	12	672	1.39	0.165	<0.01	0.74	0.190
15	Site:Exp:Strength:Att	12	672	1.91	**0.031**	<0.01	0.74	**0.049**
	Experiment 1							
1	Site	1	28	<0.01	0.976	<0.01	-	-
2	Strength	6	168	264.35	**<0.001**	0.64	0.31	**<0.001**
3	Att	2	56	146.49	**<0.001**	0.54	0.77	**<0.001**
4	Site:Strength	6	168	2.35	0.033	0.02	0.31	0.109
5	Site:Att	2	56	1.74	0.185	0.01	0.77	0.193
6	Strength:Att	12	336	35.14	**<0.001**	0.12	0.55	**<0.001**
7	Site:Strength:Att	12	336	0.93	0.522	<0.01	0.56	0.485
	Experiment 3							
1	Site	1	28	<0.01	0.954	<0.01	-	-
2	Strength	6	168	172.17	**<0.001**	0.64	0.39	**<0.001**
3	Att	2	56	102.73	**<0.001**	0.33	0.65	**<0.001**
4	Site:Strength	6	168	0.42	0.867	<0.01	0.39	0.693
5	Site:Att	2	56	0.53	0.594	<0.01	0.65	0.520
6	Strength:Att	12	336	19.06	**<0.001**	0.09	0.70	**<0.001**
7	Site:Strength:Att	12	336	2.48	**0.004**	0.01	0.70	**0.012**

Supplementary Table 8. Feature detection criterion (related to Supplementary Figure 5b).

	Effect	DF_n	DF_d	*F*	*p*	η_G^2	ε	*p*[GG]
	All experiments							
1	Site	1	56	0.02	0.899	<0.01	-	-
2	Exp	1	56	0.67	0.417	0.01	-	-
3	Strength	6	336	401.02	**<0.001**	0.40	0.42	**<0.001**
4	Att	2	112	11.86	**<0.001**	0.01	0.84	**<0.001**
5	Site:Exp	1	56	1.12	0.294	0.02	-	-
6	Site:Strength	6	336	0.97	0.447	<0.01	0.42	0.398
7	Exp:Strength	6	336	12.95	**<0.001**	0.02	0.42	**<00.001**
8	Site:Att	2	112	1.82	0.167	<0.01	0.84	0.174
9	Exp:Att	2	112	5.21	**0.007**	<0.01	0.84	**0.010**
10	Strength:Att	12	672	58.27	**<0.001**	0.04	0.65	**<0.001**
11	Site:Exp:Strength	6	336	3.41	**0.003**	<0.01	0.42	**0.026**
12	Site:Exp:Att	2	112	0.05	0.948	<0.01	0.84	0.924
13	Site:Strength:Att	12	672	1.05	0.403	<0.01	0.65	0.399
14	Exp:Strength:Att	12	672	1.49	0.121	<0.01	0.65	0.159
15	Site:Exp:Strength:Att	12	672	2.02	**0.020**	<0.01	0.65	**0.044**
	Experiment 1							
1	Site	1	28	0.64	0.429	0.02	-	-
2	Strength	6	168	264.35	**<0.001**	0.33	0.31	**<0.001**
3	Att	2	56	6.76	**0.002**	0.02	0.83	**0.004**
4	Site:Strength	6	168	2.35	0.033	<0.01	0.31	0.109
5	Site:Att	2	56	0.54	0.588	<0.01	0.83	0.556
6	Strength:Att	12	336	35.14	**<0.001**	0.04	0.56	**<0.001**
7	Site:Strength:Att	12	336	0.93	0.522	<0.01	0.56	0.485
	Experiment 3							
1	Site	1	28	0.48	0.495	0.01	-	-
2	Strength	6	168	183.35	**<0.001**	0.47	0.43	**<0.001**
3	Att	2	56	13.61	**<0.001**	0.01	-	**-**
4	Site:Strength	6	168	2.12	0.053	0.01	0.43	0.113
5	Site:Att	2	56	2.08	0.135	<0.01	-	-
6	Strength:Att	12	336	27.09	**<0.001**	0.05	-	-
7	Site:Strength:Att	12	336	1.86	**0.038**	<0.01	-	-

Supplementary Table 9. Unequal variance stimulus detection sensitivity (related to Supplementary Figure 6a).

	Effect	DF_n	DF_d	*F*	*p*	η^2_G	ε	*p*[GG]
	All experiments							
1	Site	1	56	2.76	0.102	0.02	-	-
2	Exp	1	56	5.50	0.023	0.04	-	-
3	Strength	6	336	209.87	**<0.001**	0.50	0.44	**<0.001**
4	Att	2	112	201.24	**<0.001**	0.36	0.82	**<0.001**
5	Site:Exp	1	56	0.27	0.603	<0.01	-	-
6	Site:Strength	6	336	0.49	0.818	<0.01	0.44	0.667
7	Exp:Strength	6	336	7.29	**<0.001**	0.03	0.44	**<0.001**
8	Site:Att	2	112	2.38	0.098	<0.01	0.82	0.109
9	Exp:Att	2	112	4.53	**0.013**	0.01	0.82	**0.019**
10	Strength:Att	12	672	23.79	**<0.001**	0.08	0.75	**<0.001**
11	Site:Exp:Strength	6	336	0.29	0.941	<0.01	0.44	0.806
12	Site:Exp:Att	2	112	1.24	0.292	<0.01	0.82	0.288
13	Site:Strength:Att	12	672	1.14	0.323	<0.01	0.75	0.332
14	Exp:Strength:Att	12	672	1.35	0.185	<0.01	0.75	0.208
15	Site:Exp:Strength:Att	12	672	1.58	0.094	<0.01	0.75	0.120
	Experiment 1							
1	Site	1	28	2.76	0.108	0.03	-	-
2	Strength	6	168	106.35	**<0.001**	0.49	0.43	**<0.001**
3	Att	2	56	107.61	**<0.001**	0.44	0.82	**<0.001**
4	Site:Strength	6	168	0.25	0.957	<0.01	0.43	0.831
5	Site:Att	2	56	2.84	0.07	0.02	0.82	0.079
6	Strength:Att	12	336	14.21	**<0.001**	0.09	0.50	**<0.001**
7	Site:Strength:Att	12	336	0.69	0.765	<0.01	050	0.660
	Experiment 3							
1	Site	1	28	0.57	0.456	<0.01	-	-
2	Strength	6	168	110.36	**<0.001**	0.52	0.40	**<0.001**
3	Att	2	56	95.30	**<0.001**	0.28	0.80	**<0.001**
4	Site:Strength	6	168	0.50	0.811	<0.01	0.40	0.645
5	Site:Att	2	56	0.15	0.861	<0.01	0.80	0.815
6	Strength:Att	12	336	11.27	**<0.001**	0.08	0.72	**<0.001**
7	Site:Strength:Att	12	336	1.90	0.034	0.01	0.72	0.056

Supplementary Table 10. Unequal variance stimulus detection criterion (related to Supplementary Figure 6b).

	Effect	DF_n	DF_d	*F*	*p*	η_G^2	ε	*p*[GG]
	All experiments							
1	Site	1	56	0.30	0.586	<0.01	-	-
2	Exp	1	56	0.42	0.520	<0.01	-	-
3	Strength	6	336	250.30	**<0.001**	0.38	0.42	**<0.001**
4	Att	2	112	4.65	**0.012**	<0.01	0.70	**0.023**
5	Site:Exp	1	56	0.03	0.864	<0.01	-	-
6	Site:Strength	6	336	3.88	**0.001**	<0.01	0.42	**0.016**
7	Exp:Strength	6	336	24.11	**<0.001**	0.06	0.42	**<0.001**
8	Site:Att	2	112	6.20	**0.003**	0.01	0.70	**0.008**
9	Exp:Att	2	112	5.57	**0.005**	<0.01	0.70	**0.012**
10	Strength:Att	12	672	33.19	**<0.001**	0.04	0.61	**<0.001**
11	Site:Exp:Strength	6	336	1.36	0.231	<0.01	0.42	0.261
12	Site:Exp:Att	2	112	2.08	0.129	<0.01	0.70	0.146
13	Site:Strength:Att	12	672	1.40	0.162	<0.01	0.61	0.202
14	Exp:Strength:Att	12	672	2.55	**0.003**	<0.01	0.61	**0.013**
15	Site:Exp:Strength:Att	12	672	1.44	0.143	<0.01	0.51	0.184
	Experiment 1							
1	Site	1	28	0.24	0.626	<0.01	-	-
2	Strength	6	168	99.71	**<0.001**	0.25	0.41	**<0.001**
3	Att	2	56	6.88	**0.002**	0.03	0.71	**0.006**
4	Site:Strength	6	168	1.43	0.205	<0.01	0.41	0.244
5	Site:Att	2	56	6.20	**0.004**	0.02	0.71	**0.010**
6	Strength:Att	12	336	11.09	**<0.001**	0.03	0.43	**<0.001**
7	Site:Strength:Att	12	336	1.83	0.043	<0.01	0.43	0.109
	Experiment 3							
1	Site	1	28	0.08	0.786	<0.01	-	-
2	Strength	6	168	157.43	**<0.001**	0.50	0.36	**<0.001**
3	Att	2	56	2.21	0.120	<0.01	0.68	0.139
4	Site:Strength	6	168	3.26	**0.005**	0.02	0.36	**0.041**
5	Site:Att	2	56	0.78	0.466	<0.01	0.68	0.421
6	Strength:Att	12	336	24.84	**<0.001**	0.06	0.49	**<0.001**
7	Site:Strength:Att	12	336	1.00	0.452	<0.01	0.49	0.429

Supplementary Table 11. Unequal variance feature detection sensitivity (related to Supplementary Figure 6a).

	Effect	DF_n	DF_d	*F*	*p*	η^2_G	ε	*p*[GG]
	All experiments							
1	Site	1	56	0.15	0.703	<0.01	-	-
2	Exp	1	56	1.89	0.174	0.01	-	-
3	Strength	6	336	169.03	**<0.001**	0.41	0.46	**<0.001**
4	Att	2	112	107.66	**<0.001**	0.26	0.80	**<0.001**
5	Site:Exp	1	56	0.12	0.731	<0.01	-	-
6	Site:Strength	6	336	0.25	0.960	<0.01	0.46	0.847
7	Exp:Strength	6	336	4.51	**<0.001**	0.02	0.46	**0.006**
8	Site:Att	2	112	2.05	0.134	<0.01	0.80	0.144
9	Exp:Att	2	112	6.66	**0.002**	0.02	0.80	**0.004**
10	Strength:Att	12	672	8.83	**<0.001**	0.03	0.61	**<0.001**
11	Site:Exp:Strength	6	336	0.50	0.812	<0.01	0.46	0.670
12	Site:Exp:Att	2	112	3.69	**0.028**	0.01	0.80	**0.038**
13	Site:Strength:Att	12	672	1.24	0.248	<0.01	0.61	0.275
14	Exp:Strength:Att	12	672	1.25	0.245	<0.01	0.61	0.272
15	Site:Exp:Strength:Att	12	672	1.15	0.316	<0.01	0.61	0.330
	Experiment 1							
1	Site	1	28	<0.01	0.981	<0.01	-	-
2	Strength	6	168	75.51	**<0.001**	0.33	0.43	**<0.001**
3	Att	2	56	57.77	**<0.001**	0.34	0.80	**<0.001**
4	Site:Strength	6	168	0.29	0.939	<0.01	0.43	0.798
5	Site:Att	2	56	3.90	**0.026**	0.03	0.80	**0.036**
6	Strength:Att	12	336	6.27	**<0.001**	0.04	0.48	**<0.001**
7	Site:Strength:Att	12	336	0.84	0.610	<0.01	0.48	0.536
	Experiment 3							
1	Site	1	28	0.31	0.584	<0.01	-	-
2	Strength	6	168	94.67	**<0.001**	0.49	0.44	**<0.001**
3	Att	2	56	55.57	**<0.001**	0.17	0.80	**<0.001**
4	Site:Strength	6	168	0.43	0.862	<0.01	0.44	0.712
5	Site:Att	2	56	0.22	0.807	<0.01	0.80	0.758
6	Strength:Att	12	336	4.19	**<0.001**	0.04	0.59	**<0.001**
7	Site:Strength:Att	12	336	1.45	0.144	0.01	0.59	0.188

Supplementary Table 12. Unequal variance feature detection criterion (related to Supplementary Figure 6b).

	Effect	DF_n	DF_d	*F*	*p*	η_G^2	ε	*p*[GG]
	All experiments							
1	Site	1	56	<0.01	0.973	<0.01	-	-
2	Exp	1	56	0.38	0.539	<0.01	-	-
3	Strength	6	336	322.36	**<0.001**	0.40	0.47	**<0.001**
4	Att	2	112	3.42	**0.036**	<0.01	0.84	**0.045**
5	Site:Exp	1	56	0.63	0.430	<0.01	-	-
6	Site:Strength	6	336	0.98	0.441	<0.01	0.47	0.402
7	Exp:Strength	6	336	13.90	**<0.001**	0.03	0.47	**<0.001**
8	Site:Att	2	112	0.30	0.741	<0.01	0.84	0.702
9	Exp:Att	2	112	3.65	**0.029**	<0.01	0.84	**0.037**
10	Strength:Att	12	672	29.21	**<0.001**	0.05	0.65	**<0.001**
11	Site:Exp:Strength	6	336	2.04	0.059	<0.01	0.47	0.114
12	Site:Exp:Att	2	112	0.53	0.588	<0.01	0.84	0.557
13	Site:Strength:Att	12	672	0.82	0.633	<0.01	0.65	0.584
14	Exp:Strength:Att	12	672	2.20	**0.010**	<0.01	0.65	**0.028**
15	Site:Exp:Strength:Att	12	672	1.45	0.139	<0.01	0.65	0.177
	Experiment 1							
1	Site	1	28	0.28	0.598	<0.01	-	-
2	Strength	6	168	175.96	**<0.001**	0.32	0.43	**<0.001**
3	Att	2	56	1.88	0.162	<0.01	0.84	0.170
4	Site:Strength	6	168	0.89	0.504	<0.01	0.43	0.437
5	Site:Att	2	56	0.03	0.972	<0.01	0.84	0.954
6	Strength:Att	12	336	11.90	**<0.001**	0.03	0.41	**<0.001**
7	Site:Strength:Att	12	336	0.42	0.957	<0.01	0.41	0.834
	Experiment 3							
1	Site	1	28	0.35	0.559	<0.01	-	-
2	Strength	6	168	164.18	**<0.001**	0.48	0.46	**<0.001**
3	Att	2	56	10.31	**<0.001**	0.01	-	-
4	Site:Strength	6	168	1.82	0.097	0.01	0.46	0.154
5	Site:Att	2	56	2.01	0.144	<0.01	-	-
6	Strength:Att	12	336	18.81	**<0.001**	0.07	0.56	**<0.001**
7	Site:Strength:Att	12	336	1.72	0.062	<0.01	0.56	0.110

Supplementary Table 13. AUC of stimulus visibility (related to Figure 2d and Figure 5 top row).

		Effect	DF_r	DF_d	*F*	*p*	η_G^2	ε	*p*[GG]
All experiments	1	Site	1	111	0.26	0.613	<0.01	-	-
	2	Exp	3	111	67.57	**<0.001**	0.61	-	-
	3	Att	2	222	117.24	**<0.001**	0.14	0.91	**<0.001**
	4	Site:Exp	3	111	3.52	**0.017**	0.08	-	-
	5	Site:Att	2	222	0.81	0.447	<0.01	0.91	0.436
	6	Exp:Att	6	222	1.75	0.110	<0.01	0.91	0.118
	7	Site:Exp:Att	6	222	1.46	0.193	<0.01	0.91	0.200
Experiment 1	1	Site	1	28	4.79	**0.037**	0.14	-	-
	2	Att	2	56	36.06	**<0.001**	0.10	0.77	**<0.001**
	3	Site:Att	2	56	3.03	0.056	0.01	0.77	0.071
Experiment 2	1	Site	1	28	0.04	0.840	<0.01	-	-
	2	Att	2	56	49.41	**<0.001**	0.42	0.61	**<0.001**
	3	Site:Att	2	56	0.50	0.608	<0.01	0.61	0.52
Experiment 3	1	Site	1	28	1.77	0.194	0.05	-	-
	2	Att	2	56	15.84	**<0.001**	0.06	-	-
	3	Site:Att	2	56	1.03	0.362	<0.01	-	-
Experiment 4	1	Site	1	27	0.34	0.568	<0.01	-	-
	2	Att	2	54	39.80	**<0.001**	0.37	-	-
	3	Site:Att	2	54	0.84	0.438	0.01	-	-

Supplementary Table 14. AUC of feature visibility (related to Figure 3d and Figure 5 bottom row).

		Effect	DF_r	DF_d	*F*	*p*	η_G^2	ε	*p*[GG]
All experiments	1	Site	1	84	0.42	0.520	<0.01	-	-
	2	Exp	2	84	7.26	**<0.001**	0.13	-	-
	3	Att	2	168	6.54	**0.002**	<0.01	0.78	**0.004**
	4	Site:Exp	3	84	1.80	0.172	0.04	-	-
	5	Site:Att	2	168	0.32	0.728	<0.01	0.78	0.674
	6	Exp:Att	4	168	7.67	**<0.001**	0.02	0.78	**<0.001**
	7	Site:Exp:Att	4	168	1.58	0.183	<0.01	0.78	0.196
Experiment 1	1	Site	1	28	2.01	0.168	0.05	-	-
	2	Att	2	56	14.63	**<0.001**	0.09	0.62	**<0.001**
	3	Site:Att	2	56	1.18	0.314	<0.01	0.62	0.297
Experiment 2	1	Site	1	28	1.27	0.269	0.04	-	-
	2	Att	2	56	2.22	0.118	<0.01	0.64	0.140
	3	Site:Att	2	56	1.48	0.238	<0.01	0.64	0.239
Experiment 3	1	Site	1	28	1.14	0.294	0.04	-	-
	2	Att	2	56	5.27	**0.008**	0.02	0.89	**0.011**
	3	Site:Att	2	56	0.72	0.492	<0.01	0.89	0.477

Supplementary Table 15. AMI (related to Figure 6).

		Effect	DF_r	DF_d	*F*	*p*	η_G^2
Stimulus-level	1	Site	1	110	0.064	0.800	<0.01
	2	Strength regime	1	110	197.65	**<0.001**	0.64
	3	Stimulus type	1	110	3.58	0.061	0.03
	4	Site:Strength regime	1	110	<0.01	0.963	<0.01
	5	Site:Stimulus type	1	110	0.03	0.876	<0.01
	6	Strength regime:Stimulus type	1	110	0.15	0.697	<0.01
	7	Site:Strength regime:Stimulus type	1	110	1.41	0.238	0.01
Feature-level	1	Site	1	84	0.55	0.459	<0.01
	2	Exp	2	84	20.43	**<0.001**	0.33
	3	Site:Exp	2	84	0.73	0.485	0.02

Supplementary Table 16. AMI by analytic pipeline (related to Supplementary Figure 8).

		Effect	DF_r	DF_d	*F*	*p*	η_G^2
Stimulus-level	1	Site	1	220	0.74	0.390	<0.01
	2	Exp	3	220	130.23	**<0.001**	0.64
	3	Pipe	1	220	0.02	0.895	<0.01
	4	Site:Exp	3	220	0.75	0.525	<0.01
	5	Site:Pipe	1	220	0.26	0.613	<0.01
	6	Exp:Pipe	3	220	0.15	0.931	<0.01
	7	Site:Exp:Pipe	3	220	0.07	0.976	<0.01
Feature-level	1	Site	1	168	2.15	0.145	0.01
	2	Exp	2	168	45.11	**<0.001**	0.35
	3	Pipe	1	168	0.20	0.657	<0.01
	4	Site:Exp	2	168	2.13	0.122	0.03
	5	Site:Pipe	1	168	0.14	0.706	<0.01
	6	Exp:Pipe	2	168	0.09	0.911	<0.01
	7	Site:Exp:Pipe	2	168	0.06	0.943	<0.01

Supplementary Table 17. Thresholds.

		Effect	DF_r	DF_d	*F*	*p*	$\boldsymbol{\eta_G^2}$
All experiments	1	Site	1	112	0.22	0.637	<0.01
	2	Exp	3	112	36.31	**<0.001**	0.49
	3	Site:Exp	3	112	2.53	0.061	0.06
Experiment 1	1	Site	1	28	0.10	0.752	<0.01
Experiment 2	1	Site	1	28	1.51	0.229	0.05
Experiment 3	1	Site	1	28	2.26	0.144	0.08
Experiment 4	1	Site	1	28	1.84	0.185	0.06