High confidence and low accuracy in redundancy masking

Fazilet Zeynep Yildirim, Bilge Sayim

High confidence and low accuracy in redundancy masking

Fazilet Zeynep Yildirim[1]*, Bilge Sayim[1, 2]

[1]Institute of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland

[2]SCALab - Sciences Cognitives et Sciences Affectives, CNRS, UMR 9193, University of Lille, 59000 Lille, France

***Corresponding author:** Fazilet Zeynep Yildirim

Email: fazilet.yildirim@psy.unibe.ch

Address: Fabrikstrasse 8, 3012, Bern, Switzerland

Telephone: +41 31 631 38 05

ORCIDs

Fazilet Zeynep Yildirim: https://orcid.org/0000-0002-8754-8137

Bilge Sayim: https://orcid.org/0000-0002-7589-5385

Keywords

Confidence, redundancy masking, rich visual consciousness, peripheral vision, metacognition

Author Contributions

B.S. and F.Z.Y. designed the study and prepared the manuscript. F.Z.Y. collected and analyzed the data and prepared the figures. Both authors interpreted the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Abstract

Visual scenes typically contain redundant information. One mechanism by which the visual system compresses such redundancies is ‘redundancy masking’ – the reduction of the perceived number of items in repeating patterns. For example, when presented with three lines in the periphery, observers frequently report only two lines. Redundancy masking is strong in radial arrangements and absent in tangential arrangements. Previous studies suggested that redundancy-masked percepts predominate in stimuli susceptible to redundancy masking. Here, we investigated whether strong redundancy masking is associated with high confidence in perceptual judgements. Observers viewed three to seven radially or tangentially arranged lines at 10° eccentricity. They first indicated the number of lines, and then rated their confidence in their responses. As expected, redundancy masking was strong in radial

arrangements and weak in tangential arrangements. Importantly, with radial arrangements, observers were more confident in their responses when redundancy masking occurred (i.e., lower number of lines reported) than when it did not occur (i.e., correct number of lines reported). Hence, observers reported higher confidence for erroneous than for correct judgments. In contrast, with tangential arrangements, observers were similarly confident in their responses whether redundancy masking occurred or not. The inversion of confidence in the radial condition (higher confidence when accuracy was low and lower confidence when accuracy was high) suggests that redundancy-masked appearance trumps 'veridical' perception. The often-reported richness of visual consciousness may partly be due to overconfidence in erroneous judgments in visual scenes that are subject to redundancy masking.

Introduction

Humans have a rich subjective impression of their visual environment. However, empirical evidence indicates that visual abilities are strongly limited, suggesting that the impression of "richness" is illusory. For example, phenomena such as change blindness (e.g., O'Regan et al., 1999; Simons & Rensink, 2005), inattentional blindness (e.g., Simons, 2000), and the attentional blink (e.g., Dux & Marois, 2009) indicate that salient information can easily go unnoticed. Furthermore, studies investigating the capacity of visual attention and visual working memory have revealed that only a few items can be processed and maintained at once (Luck & Vogel, 2013; Scimeca & Franconeri, 2015). Such findings raise the question why our subjective impressions of the visual environment are richly detailed even though the empirical evidence strongly suggests otherwise.

One view argues that we have rich impressions because conscious perception overflows the capacities of cognitive mechanisms (e.g., attention, working memory, and decision-making) (the "rich view": Block, 2011; Bronfman et al., 2014; Koch & Tsuchiya, 2007; Vandenbroucke et al., 2014). As a result, according to the rich view, we are aware of more information than we can attend, remember, or report. According to this view, there is no tension between the capacities of conscious perception and the capacities of cognitive functions (such as attention and memory) because they are based on separate mechanisms. By contrast, many researchers argue that without specific evidence, there appears no good scientific reason to believe that consciousness and cognition are based on separate mechanisms (Baars, 1989; Cohen et al., 2016; Cohen & Dennett, 2011; Dehaene, 2014; Kouider et al., 2010; Lau & Rosenthal, 2011; Ward et al., 2016; Ward, 2018). For example, it was suggested that the mechanisms of conscious perception and cognitive processes cannot be separated because awareness is intrinsically linked to cognitive functions and information is not consciously perceived until it is accessed by higher-order systems, such as attention, working memory, and decision-making (the "sparse view": Baars, 1989; Cohen et al., 2016; Cohen & Dennett, 2011; Dehaene, 2014; Kouider et al., 2010; Lau & Rosenthal, 2011; Ward et al., 2016; Ward, 2018). According to the sparse view, the perceived level of detail is rich only at the focus of attention and becomes drastically limited (sparse) outside of the focal point of attention. The tension between rich conscious perception and limited capacities of cognitive functions remains in this case: why do we think we see more if our conscious perception is limited by higher-level cognitive mechanisms?

To answer this question, some researchers suggested that even though perception is limited by cognitive mechanisms, it is not 'sparse' because the visual system can encode considerably more than just a few items by representing groups of items as an ensemble and by summarizing redundant information (Cohen et al., 2016; Jackson-Nielsen et al., 2017). The key idea here is that the visual system exploits redundancies found in real-world scenes to represent large amounts of information, often extending into the visual periphery, as single summary statistics (Alvarez, 2011; Ariely, 2001; Whitney et al., 2014; Whitney & Yamanashi Leib, 2018). This view reconciles the subjective impressions of a richly detailed world and poor visual performance: Ensemble representations store a wide range of information which is assumed to result in rich impressions of the visual world; however, as only summary information is available, we have little or no information about the individual items in the scene (Cohen et al., 2016). Studies showing overestimation of our performance can be regarded as support for this view. For example, people often believe that detecting changes in a change blindness experiment will be easy, and are surprised to find out that it is not (Levin et al., 2000). Similarly, people do not realize how limited their performance is in the visual periphery. For example, observers estimated their performance to be higher in a crowded condition compared to an uncrowded condition in the periphery although their performance was worse in the crowded condition (Odegaard et al., 2018; but see Toscani Mamassian, & Valsecchi, 2021).

Recent studies showed that the visual system does not only represent information as an ensemble but also compresses redundant information substantially by masking individual items in repeating patterns (i.e., 'redundancy masking'; Sayim & Taylor, 2019; Taylor & Sayim, 2020; Yildirim et al., 2019, 2020, 2021, 2022). For example, when presented with three identical items in the periphery, the majority of observers report seeing only two items (Sayim & Taylor, 2019; Yildirim et al., 2021). We have shown that redundancy masking increased with increasing similarity and regularity of the items, decreased with increasing spacing between items, and was larger with radial compared to tangential item arrangements (Yildirim et al., 2020). When redundancy masking occurs, visual space seems to be compressed: Observers estimated the distance between the two perceived of three presented items as smaller than the actual distance between the two outermost of the three presented items, and as larger than the distance between two adjacent items (Yildirim et al., 2019). Taken together, these findings suggest that the visual system summarizes information not only in the form of ensemble perception but also compresses the number of identical items even when only three items are presented. Perceiving only two of three items is such a significant loss of information about the actual stimulus that it makes redundancy masking a promising tool to investigate questions regarding the often-reported richness of visual consciousness. Here we investigated observers' metacognition in conditions where redundancy masking is expected to be strong or to not occur.

Observers were presented with arrays of lines in the visual periphery and asked to rate their confidence with displays usually yielding strong (radially arranged lines) and no (tangentially arranged lines) redundancy masking. The number of lines presented was from three to seven, however, our main focus was on three lines as the relative magnitude of redundancy masking is maximal in this condition (i.e., often only two lines are reported, corresponding to missing ⅓ of the presented lines). Consistent with our previous study (Yildirim et al., 2020), we found that redundancy masking was strong with radially arranged lines and absent with tangentially arranged lines. The analysis of confidence ratings showed that in the radial condition, observers were more confident in their responses when redundancy masking occurred (less lines reported than presented) compared to when no redundancy masking occurred (correct responses). Hence, observers reported higher confidence for erroneous than for correct judgments. In contrast, in the tangential condition, observers' confidence was similar in trials with and without redundancy masking. The inversion of confidence in the radial condition (higher confidence when accuracy was low and lower confidence when accuracy was high) suggests that redundancy-masked appearance trumped 'veridical' perception. We suggest that the often-reported richness of visual consciousness may partly be due to overconfidence in erroneous judgments in visual scenes that are subject to redundancy masking.

Materials and methods

Participants

Thirteen undergraduate students (age range: 21-26 years, seven male) from the University of Bern participated in the experiment in either exchange for course credit or without compensation. The number of participants recruited was determined based on our earlier studies (Yildirim et al., 2020, 2021). All observers reported normal or corrected-to-normal visual acuity. Observers were naïve regarding the aim of the study. Before the experiment, observers signed a consent form and were informed about the general procedure. The experimental protocols were approved by the local ethics committee at the University of Bern. All procedures were in accordance with the Declaration of Helsinki.

Apparatus and stimuli

Stimuli were generated with Psychopy v2.7.11 (Peirce, 2007) and displayed on a 22-in. CRT monitor with a resolution of 1152 × 864 and a refresh rate of 110 Hz. The experiment was conducted in a dimly illuminated room. Observers viewed the monitor from a distance of 57 cm and were supported by a chin and head rest. A black (1 cd/m^2) disc (diameter = 0.2°) was presented at the center of the screen for fixation. Stimuli consisted of black (1 cd/m^2) lines that were 1° in length and 0.04° in width, presented on a uniform gray (42 cd/m^2) background. The line arrays were centered at 10° eccentricity, and presented to either the right or the left of the fixation disc. The number of presented lines ranged from three to seven. The center-to-center spacing between adjacent lines within a line array was 0.85°. There were two spatial arrangements of lines: in the radial condition, vertically oriented lines were horizontally arranged (Figure 1, top left frame), and in the tangential condition, horizontally oriented lines were vertically arranged (Figure 1, top right frame). The conditions are denoted as radial and tangential to refer to how the lines are located relative to a concentric circle around fixation. The position of the line array was slightly varied at random across trials (centered at 10° or jittered 4.7′ either up, down, left, or right). Responses were recorded using a number pad (enumeration task) and a computer mouse (confidence rating task).

Procedure

At the beginning of the experiment, the fixation disc was presented for 1s. Next, a stimulus was presented for 150 ms randomly to the left or the right of fixation. Observers were required to indicate the number of lines they perceived with a key press on the number pad (0–9). Then, they indicated their confidence on their response with a confidence scale from 1 (low) to 4 (high) by using the computer mouse (Figure 1). The center of the confidence scale axis was presented at fixation. Observers were instructed to distribute their confidence responses over the whole response scale. The next stimulus was presented 440 ms after the confidence response.

The stimulus location (left or right of fixation) and the number of lines (three to seven) were randomized and counterbalanced within each block. Presentations were blocked according to the spatial arrangement of the lines (radial and tangential). A block consisted of 80 trials. Observers completed two blocks with each spatial arrangement (a total of 320 trials). The sequence of radial and tangential blocks was pseudorandomized for each observer. A schematic depiction of the procedure is shown in Figure 1.

Before the experiment, for each participant we verified that the spacing between adjacent lines was above their resolution limit. A two-line discrimination task with 100 trials was performed with radial and tangential lines before the main experiment (a total of 200 trials): Two lines with varying spacings were presented on the horizontal meridian at the maximum eccentricity of the lines in the main experiment (radial: 12.1°, tangential: 10°). Participants were asked whether they perceived one or two lines. All observers reported perceiving two lines in all trials with the center-to-center spacing presented in the main experiment (0.85°).

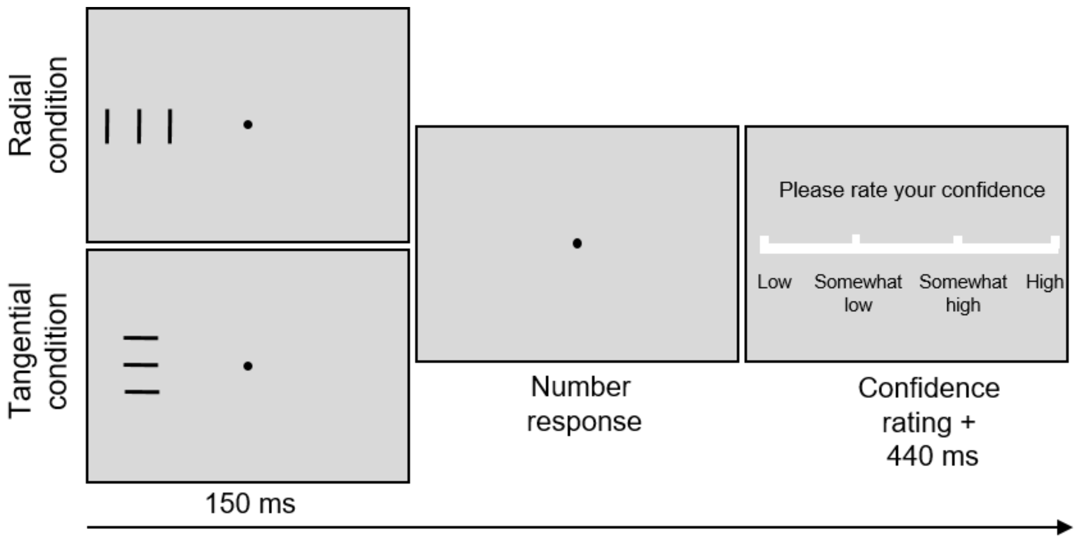

Figure 1. Schematic of the experimental paradigm. Stimuli consisted of three to seven lines, and were presented in separate blocks in radial or tangential arrangements randomly to the left or right side of fixation. Observers first indicated the number of lines, and then rated their confidence on their response. (The text in the last frame appeared in German in the experiment).

Data analyses

Deviation scores: Performance in the enumeration task was defined as the number of lines presented subtracted from the number of lines reported (“deviation”). Hence if the number of lines reported was the same as the number of lines presented, the deviation was zero; reporting more lines than presented yielded deviation scores above zero, and reporting fewer lines than presented yielded deviation scores below zero. Mean deviation scores were analyzed by a generalized linear mixed-effects model specifying the number of lines presented, and spatial arrangement as fixed effects and subject as a random factor using the glmmTMB package (Brooks et al., 2017). For model selection, null models (without the fixed effects) and full models (with the fixed effects) were fitted and hierarchically compared. Similar incremental model building was used to select the minimum degree polynomial that fitted the data. Likelihood-ratio tests with Satterthwaite's approximation for the degrees of freedom were performed for model comparisons, and the Akaike information criterion was used to select the best fitting model (Matuschek et al., 2017). Confidence intervals were calculated with the ggpredict function of the ggeffects package (Lüdecke, 2018). The pseudo *R* squared statistic (R^2) was computed to quantify how well the fixed and random factors explained the performance using r.squaredGLMM() function of the MuMIn package (Barton & Barton, 2015; Johnson, 2014). Assumptions underlying the models were checked with diagnostic plots of residuals using the DHARMa package (Hartig, 2017). A second-degree polynomial regression was used to fit the deviation scores on the number of lines presented. The random effect structure included random slopes and random intercepts for each subject. The full model was selected (R^2 = 0.43).

Confidence ratings: Mean confidence ratings were analyzed by a generalized linear mixed-effects model specifying the number of lines presented, and spatial arrangement as fixed effects and subject as a random factor using the glmmTMB package (Brooks et al., 2017). Model selection and checking of the assumptions were the same as in the analysis of the deviation scores. The full model was selected (R^2 = 0.37).

Response speeds: The inverse transformations of reaction times (1/RT; response speed) were analyzed by a generalized linear mixed-effects model specifying the deviation scores, the number of lines presented, and the spatial arrangements as fixed effects, and subject as a random factor using the glmmTMB package (Brooks et al., 2017). Model selection and assumptions checking were the same as in the analysis of deviation scores. The full model was selected (R^2 = 0.15).

Outlier detection and removal: Trials on which RTs were longer than 10 s were excluded from the analyses (Baranski & Petrusic, 1994). One trial was excluded from the analyses because of excessively fast response speed (0.92 ms) for an outlier response in that condition (i.e., the number was only reported once by a single observer). Additionally, trials on which deviation scores exceeded ±3 were excluded from the analyses. In total, 0.47% of the trials were removed.

Results

Mean deviation scores as a function of the number of lines presented for the radial and tangential conditions are shown in Figure 2. As expected, we found strong redundancy masking in the radial and no redundancy masking in the tangential condition. Both main effects of spatial arrangement ($\chi^2(1)$ = 21.49, p = 0.0000036) and the number of lines presented ($\chi^2(2)$ = 12.55, p = 0.0019), and the interaction ($\chi^2(2)$ = 8.53, p = 0.014) between them were significant. Mean deviation scores were clearly below zero in the radial condition (–0.66 ± SD 0.42; strong redundancy masking), and they were above zero in the tangential condition (0.34 ± SD 0.69; the opposite of redundancy masking, i.e., overestimation). The deviation scores varied with the number of lines presented in a quadratic (inverted-U) manner in the both conditions: an initial positive slope of the deviation scores for small numbers of lines was followed by a negative slope with larger numbers of lines (the quadratic fit outperformed linear fits). In the radial condition, the deviation scores were all below zero, and in the tangential condition, they were close to zero with three lines, and above zero for all other numbers of lines. Descriptive statistics of deviation scores for each number of lines and spatial arrangement condition are shown in the Supplementary Table 1a.

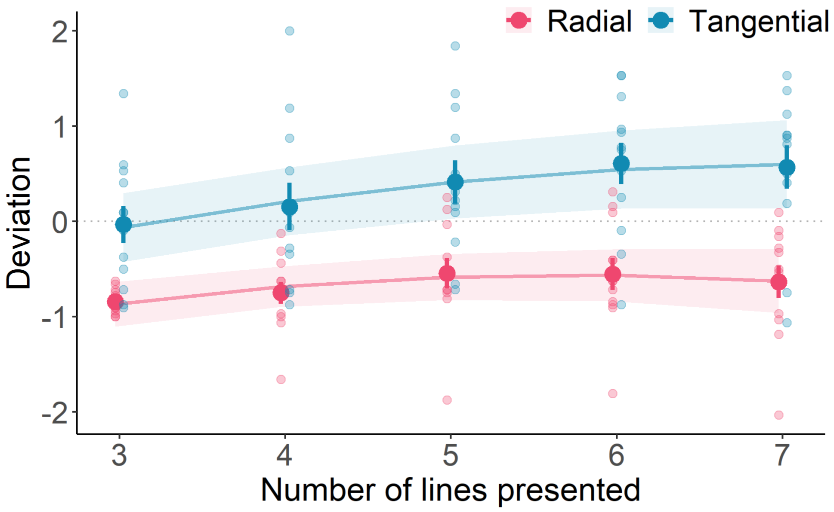

Figure 2. Deviation scores as a function of the number of lines presented in the radial and tangential conditions. The large pink and blue data points show mean deviation scores (±*SEM*) for the radial and tangential condition, respectively. The small pink and blue data points show mean deviation scores for each individual observer. The lines and shaded regions show the model fits and confidence intervals, respectively (±1.96 * *SEM*).

Confidence ratings as a function of deviation scores for the radial and tangential conditions are shown in Figure 3. Our main focus was the condition with three lines as redundancy masking is most pronounced in this condition. The result of the polynomial regression analysis suggested that observers were more confident in the redundancy masked report (deviation score -1; two lines reported) compared to the correct report (deviation score 0; three lines reported) in the radial condition (Fig. 3, three lines). In comparison, they were equally confident in both reports in the tangential condition (Fig. 3, three lines). An analysis of the local slopes in the polynomial regressions confirmed this finding. The slope between the deviation score of -1 (two lines) and 0 (three lines) in the radial condition was significantly lower than zero (β = -0.61, 95% CI [-0.84, -0.38], $p < .0001$) whereas the corresponding slope in the tangential condition did not differ from zero (β = 0.008, 95% CI [-0.25, 0.27], p =.953). Importantly, the slopes between the deviation scores of -1 (two lines) and 0 (three lines) differed between the radial and the tangential condition ($t(380) = -3.45$, p = 0.003). Confidence dropped significantly from two lines to three lines in the radial condition (Fig. 3; three lines); in contrast, it did not change between two and three lines in the tangential condition (Fig. 3, three lines). Descriptive statistics of confidence ratings for each number of lines and spatial arrangement condition are shown in the Supplementary Table 1b.

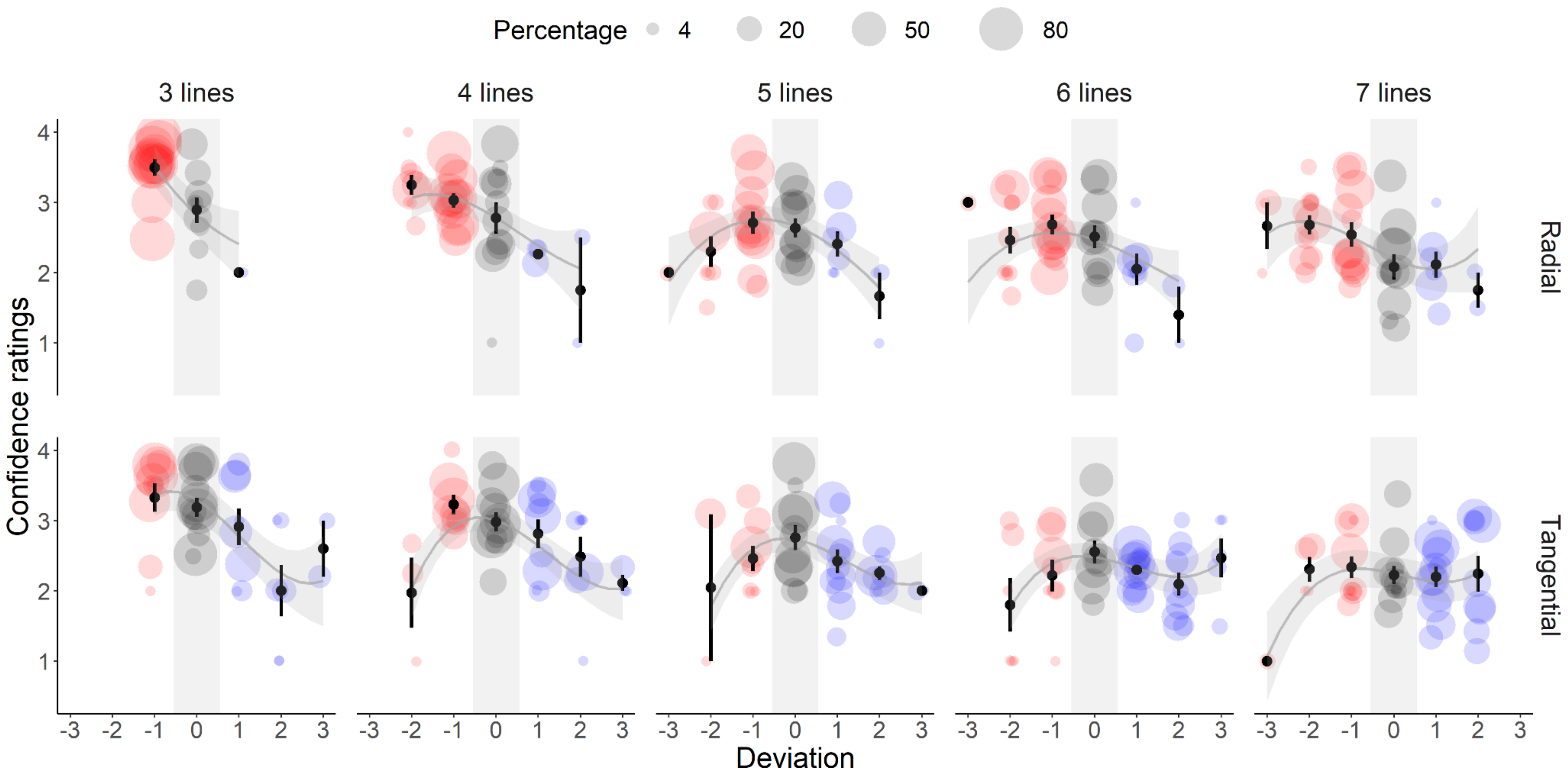

Figure 3. Confidence ratings as a function of deviation scores shown for each number of lines separately for the radial (top row) and tangential conditions (bottom row). Each red, grey, or blue colored data point represents the responses of one participant in a given condition. Red, grey, and blue points denote deviation scores lower than, equal to, and higher than zero, respectively. The size of the data points represents the percentage of the participant's response with a particular number (see legend on top). The highlighted rectangles show the correct enumeration responses (deviation score of zero). The black data points with error bars (±*SEM*) show the mean confidence ratings. The grey lines and shaded regions around the lines show the predicted values and confidence intervals for the predicted values based on the standard errors (± 1.96 * *SEM*), respectively.

We also assessed the response speeds in the radial and tangential conditions to explore whether they followed the same pattern as the confidence ratings. Response speed as a function of deviation scores for the radial and tangential conditions are shown in Figure 4. As in the confidence ratings analysis, we were mainly interested in the conditions with three lines. The polynomial regressions showed that the slope between deviation score of -1 (two lines) and 0 (three lines) in the radial condition was significantly lower than zero (β = -0.21, 95% CI [-0.30, -0.12], p < .0001) whereas the corresponding slope in the tangential condition did not differ from zero (β = -0.03, 95% CI [-0.13, 0.07], p =.524). Importantly, the slopes between the deviation scores of -1 (two lines) and 0 (three lines) differed between the radial and the tangential condition ($t(379)$ = -3.78, p = 0.0009): Response speeds were significantly faster for two than three lines in the radial condition; in contrast, there was no difference of response speeds between two and three lines in the tangential condition (Fig. 4; three lines). Systematic differences based on key locations or fingers used may have contributed to the different response speeds for indicating two and three lines. However, while there was a difference between two and three lines in the radial condition, no such difference occurred in the tangential condition. Descriptive statistics of the response speeds for each number of lines and spatial arrangement condition are shown in the Supplementary Table 1c.

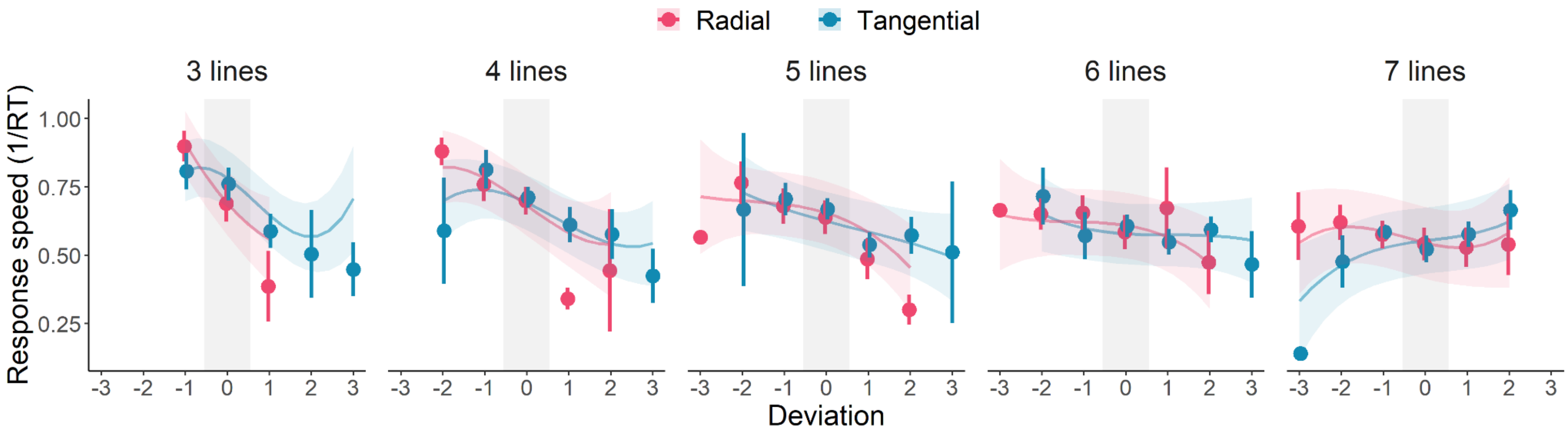

Figure 4. Response speed (1/RT) as a function of deviation scores shown for each number of lines and spatial arrangement conditions. The pink and blue data points with error bars (±*SEM*) show the mean response speed for the radial and tangential conditions, respectively. The pink and blue lines show the predicted values from the mixed model for the radial and tangential conditions, respectively. Shaded regions represent confidence intervals for the predicted values based on the standard errors (± 1.96 * *SEM*).

Discussion

A long-standing puzzle in vision science is that we seem to be unaware of how poor our peripheral vision actually is. Here, we used redundancy masking, the loss of items in repeating patterns, to probe observers' metacognition of peripheral vision. We found an inversion of the usual relation of performance and confidence: Participants were more confident in their responses when erroneously reporting two instead of three lines. This was the case in the radial condition where redundancy masking was expected, but not in the tangential condition where no – or less - redundancy masking was expected. We suggest that stimulus appearance in the radial condition was less ambiguous when redundancy masking occurred compared to when it did not occur. This is also reflected in the proportion of trials that were redundancy-masked: In the radial condition, observers reported two lines in 80% of the trials when three lines were presented (Fig. 3, three lines, radial condition). By contrast, in the tangential condition, in which there was no difference in confidence between redundancy-masked and correct trials, participants reported two lines in only 32% of the trials (Fig. 3, three lines, tangential condition). The evaluation of response speeds also supports this interpretation: Not only were participants more confident in their responses when redundancy masking occurred in the radial condition, they were also faster. Hence, all measures, confidence, frequency, and response speeds indicated that three lines appeared more strongly like two than three lines. Taken together, our results suggest that redundancy-masked appearance trumps 'veridical' perception when items are radially arranged.

Unlike previous studies in which observers' metacognition was compared across conditions that differed strongly, such as foveal versus peripheral vision (Solovey et al., 2015), or crowded versus uncrowded stimuli (e.g., Odegaard et al., 2018), we aimed to investigate metacognition with identical stimuli and presentation conditions. Comparisons between foveal and peripheral presentation are challenging because many factors may cause differences between the conditions. In particular, different criteria to evaluate accuracy of one's performance may produce differences between foveal and peripheral presentation, even when performance in the two conditions is matched. For example, a recent study showed that observers' metacognitive sensitivity was lower when comparing perceptual decisions of stimuli presented at different eccentricities (central and peripheral) compared to the same eccentricity (central with central, and peripheral with peripheral) (Toscani et al., 2021). Comparing crowded with uncrowded stimuli is limited by similar constraints: Differences in stimulus complexity (Zhang, Zhang, Xue, Liu, & Yu, 2009; Sayim & Wagemans, 2017), attentional demands (Strasburger, 2005; He, Intriligator, & Cavanagh, 1996; Rummens & Sayim, 2021), internal and external noise (Sun, Chung, & Tjan, 2010) render it difficult to extract the variable(s) that underlie variations of confidence. Here, we used a redundancy masking paradigm to investigate confidence when appearance varied with physically identical stimuli under identical presentation conditions. Thus, any differences between confidence judgments were not due to variations of the stimuli or presentation conditions, but reflect changes in stimulus *appearance* independent of these confounding factors. Hence, our results complement previous studies that investigated confidence in peripheral vision (Baldassi et al., 2006; Odegaard et al., 2018; Otten et al., 2017; Rosenthal, 2019; Solovey et al., 2015). For example, Odegaard et al. (2018) found higher confidence ratings for erroneous responses when the target was crowded compared to when it was not crowded. This finding was taken as an indication that the rich visual experience is due to 'subjective inflation' in the periphery. However, as their conclusions are based on comparisons between different stimulus conditions (i.e., crowded vs uncrowded), other factors related to these conditions (e.g., stimulus complexity, attentional demands) might confound their results. Here, we avoid these confounding factors by comparing confidence ratings of erroneous (redundancy-masked) responses and correct responses of the same stimulus. By comparing confidence ratings within the same stimulus, we show that confidence followed the 'appearance' of a stimulus, not performance.

Our main focus was the condition with three lines as the relative magnitude of redundancy masking is greatest when the number of items is small (Yildirim et al., 2020; 2021; see also the relative magnitude of errors in Supplementary Figure S1). In agreement with previous studies (Rummens & Sayim, 2022; Sayim & Taylor, 2019; Yildirim et al., 2020, 2021, 2022), there was strong redundancy masking in the radial condition: when three lines were presented, participants reported two lines in 80% of the trials, with high confidence (on average 3.5±0.39 on a scale from 1 to 4). In comparison, participants reported the correct number, three lines, in only 18% of the trials with lower confidence (2.9±0.57). Contrary to findings of most studies investigating confidence in perceptual judgements, observers were more confident in their incorrect responses than their correct responses. Only two observers reported four lines (1.2% of the trials with three lines), and their confidence was low (2±0.00). In the tangential condition, by contrast, observers reported two lines in 32%, three in 47%, four in 16%, and five or six in 5% of the trials with three lines. Importantly, confidence did not differ between "two" (3.3±0.64) and "three" (3.2±0.46) responses in the tangential condition (and confidence was lower when observers reported four (2.9±0.73), five (2.0±0.89), and six (2.6±0.57) lines). Hence, confidence in redundancy-masked reports was relatively high with radial, but not with tangential arrangements.

Enumeration of small quantities (1 to ~3) is typically fast and accurate (i.e., subitizing) whereas enumeration of larger quantities (~4 and larger) is slow and error-prone (i.e., estimation) (Kaufman et al., 1949; Mandler & Shebo, 1982). Subitizing performance in peripheral vision is similar as with free/central viewing for attended and widely spaced items, but has been shown to decline under high attentional load (Railo et al., 2008) and close inter-item spacings (Chakravarthi & Herbert, 2019). We suggest that in the current study observers subitized, albeit mostly erroneously, the number of lines when presented with three lines, resulting in reports of either two or three lines in most trials. In contrast, when presented with more lines (~4 to 7), subitizing was no longer possible, and observers had to estimate the number of lines, which led to a wider range of deviation scores (ranging mostly between -2 to +2 around the presented number of lines) (see also Yildirim et al., 2021).

In the radial condition, redundancy masking was not only strong with three lines, but also with larger numbers of lines. As expected, confidence ratings decreased with the number of lines presented, showing the typical increase of uncertainty about response accuracy when estimating the number of items in large sets (see Supplementary Figure S2) (Kaufman et al., 1949; Railo et al., 2008). In line with the lower confidence ratings for larger numbers of items, the variance of responses increased (ranging mostly between -2 to +1 around the presented number) and the response speed decreased. However, while there was only a small decrease of the magnitude of (negative) deviation scores with the increase of the number of lines (from three to five lines; Figure 2 and Supplementary Table 1a), relative deviation scores decreased more strongly with increasing numbers of lines (Supplementary Figure S1), showing how redundancy masking is most pronounced with smaller numbers of items. Hence, in the radial condition, the strongest relative deviation went hand in hand with the highest confidence: Observers were most confident when the relative deviation from the correct response was largest. In the tangential condition, by contrast, there was neither redundancy masking for small nor for large numbers of lines. While responses were on average accurate for three lines, (positive) deviation scores increased with the number of lines, showing that observers overestimated the number when large numbers of lines were presented. The observed decrease of confidence ratings (and response speed), and the increase of variance with increasing numbers of lines were - as in the radial condition - unsurprising. Similarly, overestimation with tangentially arranged lines has been reported previously (Yildirim et al., 2020).

Estimating large numbers of items (above the subitizing range) typically results in underestimation (Burr et al., 2010; Vetter et al., 2008). However, some studies found overestimation of the number of items (Alam et al., 1986; Ginsburg, 1976, 1978; Li, Reynvoet, & Sayim, in revision, 2021). For example, when observers were presented with regular and random sets of 7, 19, 37, 61, or 91 dots, they underestimated the numerosities in the random dot patterns (except for the smallest number which was accurately reported) (Ginsburg, 1978). By contrast, the number of dots in regular patterns was overestimated (see also 'regular-random numerosity illusion', Ginsburg 1980). In a recent study, overestimation of large numbers of items was also found with irregular arrangements (Li et al., in revision, 2021). Instead of asking observers to compare two stimuli and choose the more (or less) numerous as is common in numerosity experiments, the task was to directly indicate the number of items in the display (as in the present study). Overestimation was stronger for tangentially arranged discs compared to radially arranged discs, similar to the present results, however, with very different stimuli (irregular dot clouds) and only for larger numbers (>31 items). Our stimuli were highly regular, which is a necessary condition for redundancy masking (Rummens & Sayim, 2022; Yildirim et al., 2020). However, we only found overestimation for large numbers of tangentially arranged lines and underestimation for radially arranged lines. Hence, regularity per se cannot explain the pattern of results found here. Also direct estimation (in contrast to discrimination) is not sufficient to explain this pattern of results as, again, the task was the same in the radial and tangential condition. The stronger overestimation of tangentially arranged items in Li et al. (2021) was attributed to the anisotropic interference ('crowding') zone around targets in peripheral vision (Greenwood et al., 2017; Petrov & Meleshkevich, 2011; Toet & Levi, 1992), which may similarly underlie the difference between the radial and tangential condition in the present study (see also Yildirim et al., 2020 for similar results). We suggest that the pattern of results we found here is due to an interplay of redundancy masking (the reduction of the perceived number of radially arranged lines), and tendencies to overestimate regular (Ginsburg 1980) and tangentially (Li et al., in revision, 2021; Yildirim et al., 2020) arranged patterns.

Higher confidence in redundancy-masked stimuli may well play a role in the seemingly rich representation of the visual environment. Given the usually strong link between high confidence and good performance (e.g., Barthelmé & Mamassian, 2010), poor performance yielding high confidence remains the exception to the rule. Redundancy masking seems to be an example that inverts this relationship. Importantly, in contrast to previous studies that found high confidence when performance was poor (i.e., inflation) based on comparisons with other stimuli (e.g., uncrowded versus crowded, Odegaard et al., 2018) or presentation conditions (e.g., fovea versus periphery, Solovey et al., 2015), here we directly compared confidence judgments on the very same stimuli and under the same presentation conditions. To evaluate whether redundancy masking is a case of inflation, comparisons can be made between redundancy-masked and correct trials, and between smaller and larger numbers of lines. We already outlined the difference in confidence for redundancy-masked and correct trials (in particular, in the radial condition with three lines). The high confidence in reporting two lines when presented with three lines is indicative of inflation: Observers were confident that their erroneous responses were correct. However, confidence was lower for correct than for incorrect responses. This reversal of the typical relationship between confidence and performance shows that observers were largely incapable of accurately evaluating their performance. Importantly, confidence was relatively high for both redundancy-masked (3.5±0.39, for the three radial lines) and correct trials (2.9±0.57, for the three radial lines) - two categorically different responses for one and the same stimulus, suggesting again that observers' confidence judgments were inflated. Confidence was overall higher with smaller compared to larger numbers of lines (see also Supplementary Figure S2), and deviation scores indicated only slightly stronger redundancy masking for three lines compared to the other numbers of lines in the radial condition (and stronger overestimation with increasing numbers of lines in the tangential condition). This was expected, and it alone does not show inflation. However, due to redundancy masking, the *relative* error was much higher for three lines than larger numbers of lines (five, six, and seven; again in the radial condition) (see Supplementary Figure S1). The high confidence ratings for three lines in conjunction with the highest relative deviation error shows that observers' confidence judgments were inflated.

Our results suggest that redundancy-masked trials are a good representation of how a stimulus appeared to observers. Hence, redundancy masking may contribute to the impression of a rich visual world by creating a convincing illusion that what is perceived is accurately capturing what is present in the stimulus. Importantly, redundancy masking would only contribute to the impression of a rich visual world when the 'erroneous' nature of the percept goes unnoticed, as high confidence in the accuracy of the percept is required. While future studies still need to investigate to what extent the occurrence of redundancy masking is not detected when attending to a stimulus in the periphery before fixating it, there is some evidence that it is not easily detected. Despite many decades of investigations of visual crowding (Bernard & Chung, 2011; Bouma, 1970, 1973; Flom et al., 1963; Herzog et al., 2015; Korte, 1923; Levi et al., 1985, 2002; Manassi et al., 2012; Melnik et al., 2018, 2020; Pelli et al., 2004; Rummens & Sayim, 2019, 2021; Sayim & Cavanagh, 2013, Sayim et al., 2014, Sayim & Wagemans, 2017; Strasburger et al., 1991; Yeshurun & Rashal, 2010), where usually three more or less similar items are presented (a target and two flankers), redundancy masking was only discovered recently (Sayim & Taylor, 2019; Taylor & Sayim, 2018, 2020; Yildirim et al., 2019, 2020, 2021, 2022). Interestingly, a philosophical debate investigating the phenomenon of 'identity crowding' where the target and the flankers are the same (similar to three radial lines in the present study), has been focused on the question why performance on the central of the three identical items is so unexpectedly *good* (Block, 2012; 2013; Prettyman, 2018; Richards, 2016; Taylor & Sayim, 2018, 2020), and not unexpectedly *poor* as shown in redundancy masking paradigms. The seemingly good performance has been proposed to be evidence for the capacity to see without attention (Block 2012, 2013). Evidence from redundancy masking experiments suggests that the premises of this debate, based to a large part on purely phenomenological approaches, were erroneous, supporting the proposal that redundancy masking easily goes unnoticed. Taken together, we showed high confidence for redundancy-masked stimuli, suggesting that a compressed, non-veridical representation of the actual stimulus better represents stimulus appearance. The impression of a rich visual world may partly be driven by the inability to notice the loss of information in visual scenes susceptible to redundancy masking, and high confidence that one's phenomenology is an accurate representation of the observed scene.

Data and code availability

The datasets and R analysis scripts generated during the study are available on OSF (https://osf.io/vs5te/).

Acknowledgment

The authors thank Kim Molinari for her help with the data collection.

Funding

This work was supported by the Swiss National Science Foundation (PP00P1_163723 to Bilge Sayim).

References

Alam, S., Luccio, R., & Vardabasso, F. (1986). Regularity, Exposure Time and Perception of Numerosity. *Perceptual and Motor Skills*, *63*(2), 883–888. https://doi.org/10.2466/pms.1986.63.2.883

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. https://doi.org/10.1016/j.tics.2011.01.003

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162. https://doi.org/10.1111/1467-9280.00327

Baars, B. J. (1989). *A cognitive theory of consciousness*. Cambridge University Press.

Baldassi, S., Megna, N., & Burr, D. C. (2006). Visual clutter causes high-magnitude errors. *PLoS Biology*, *4*(3), e56. https://doi.org/10.1371/journal.pbio.0040056

Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, *55*(4), 412–428. https://doi.org/10.3758/bf03205299

Barthelmé, S., & Mamassian, P. (2010). Flexible mechanisms underlie the evaluation of visual confidence. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(48), 20834–20839. https://doi.org/10.1073/pnas.1007704107

Barton, K., & Barton, M. K. (2015). *Package 'MuMIn'. Version, 1, 18*.

Bernard, J.-B., & Chung, S. T. L. (2011). The dependence of crowding on flanker complexity and target-flanker similarity. *Journal of Vision*, *11*(8). https://doi.org/10.1167/11.8.1

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, *15*(12), 567–575. https://doi.org/10.1016/j.tics.2011.11.001

Block, N. (2012). The Grain of Vision and the Grain of Attention. *Thought: A Journal of Philosophy*, *1*(3), 170–184. https://doi.org/10.1002/tht3.28

Block, N. (2013). Seeing and Windows of Integration. *Thought: A Journal of Philosophy*, *2*(1), 29–39. https://doi.org/10.1002/tht3.62

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177–178. https://doi.org/10.1038/226177a0

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, *13*(4), 767–782. https://doi.org/10.1016/0042-6989(73)90041-2

Bronfman, Z. Z., Brezis, N., Jacobson, H., & Usher, M. (2014). We see more than we can report: "cost free" color phenomenality outside focal attention. *Psychological Science*, *25*(7), 1394–1403. https://doi.org/10.1177/0956797614532656

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, *9*(2), 378-400.https://doi.org/10.3929/ethz-b-000240890

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision*, *10*(6), 20–20. https://doi.org/10.1167/10.6.20

Chakravarthi, R., & Herbert, A. (2019). Two's company, three's a crowd: Individuation is necessary for object recognition. *Cognition*, *184*, 69–82. https://doi.org/10.1016/j.cognition.2018.12.008

Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, *15*(8), 358–364. https://doi.org/10.1016/j.tics.2011.06.008

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the Bandwidth of Perceptual Experience? *Trends in Cognitive Sciences*, *20*(5), 324–335. https://doi.org/10.1016/j.tics.2016.03.006

Dehaene, S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Penguin.

Dux, P. E., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention, Perception, & Psychophysics*, *71*(8), 1683–1700. https://doi.org/10.3758/APP.71.8.1683

Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual Resolution and Contour Interaction. *Journal of the Optical Society of America*, *53*, 1026–1032. https://doi.org/10.1364/josa.53.001026

Ginsburg, N. (1976). Effect of item arrangement on perceived numerosity: Randomness vs regularity. *Perceptual and Motor Skills*, *42*(43), 663–668. https://doi.org/10.2466/pms.1976.43.2.663

Ginsburg, N. (1978). Perceived numerosity, item arrangement, and expectancy. *The American Journal of Psychology*, *91*(2), 267–273.

Ginsburg, N. (1980). The regular-random numerosity illusion: Rectangular patterns. *The Journal of General Psychology*, *103*(2d Half), 211–216. https://doi.org/10.1080/00221309.1980.9921000

Greenwood, J. A., Szinte, M., Sayim, B., & Cavanagh, P. (2017). Variations in crowding, saccadic precision, and spatial localization reveal the shared topology of spatial vision. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(17), E3573–E3582. https://doi.org/10.1073/pnas.1615504114

Hartig, F. (2017). *DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models*. (R package version 0.1) [Computer software].

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334–337. https://doi.org/10.1038/383334a0

Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *Journal of Vision*, *15*(6), 5. https://doi.org/10.1167/15.6.5

Jackson-Nielsen, M., Cohen, M. A., & Pitts, M. A. (2017). Perception of ensemble statistics requires attention. *Consciousness and Cognition*, *48*, 149–160. https://doi.org/10.1016/j.concog.2016.11.007

Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. Methods in Ecology and Evolution, 5(9), 944–946. https://doi.org/10.1111/2041-210X.12225

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*, 498–525. https://doi.org/10.2307/1418556

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. https://doi.org/10.1016/j.tics.2006.10.012

Korte, W. (1923). Uber die Gestaltauffassung im indirekten Sehen [On the apprehension of Gestalt in indirect vision]. *Zeitschrift Für Psychologie*, *93*, 17–82.

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, *14*(7), 301–307. https://doi.org/10.1016/j.tics.2010.04.006

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, *15*(8), 365–373. https://doi.org/10.1016/j.tics.2011.05.009

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, *2*(2), 167–177. https://doi.org/10.1167/2.2.3

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*(7), 963–977. https://doi.org/10.1016/0042-6989(85)90207-x

Levin, D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change Blindness Blindness: The Metacognitive Error of Overestimating Change-detection Ability. *Visual Cognition*, *7*(1–3), 397–412. https://doi.org/10.1080/135062800394865

Li, M., Reynvoet, B., & Sayim, B. (2021). More crowded, less numerous: Crowding reduces the number of perceived items in numerosity perception. Virtual Vision Science Society, May 2021.

Li, M., Reynvoet, B., & Sayim, B. (in revision). Visual numerosity estimation is based on anisotropic representations of visual space.

Luck, S. J., & Vogel, E. K. (2013). Visual Working Memory Capacity: From Psychophysics and Neurobiology to Individual Differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, *3*(26), 772. https://joss.theoj.org/papers/10.21105/joss.00772

Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision*, *12*(10), 13. https://doi.org/10.1167/12.10.13

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22. https://doi.org/10.1037/0096-3445.111.1.1

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305-315.https://doi.org/10.1016/j.jml.2017.01.001

Melnik, N., Coates, D. R., & Sayim, B. (2020). Emergent features break the rules of crowding. *Scientific Reports*, *10*(1), 406. https://doi.org/10.1038/s41598-019-57277-y

Melnik, N., Coates, D. R., & Sayim, B. (2018). Emergent features in the crowding zone: When target-flanker grouping surmounts crowding. *Journal of Vision*, *18*(9), 19. https://doi.org/10.1167/18.9.19

Odegaard, B., Chang, M. Y., Lau, H., & Cheung, S.-H. (2018). Inflation versus filling-in: Why we feel we see more than we actually do in peripheral vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *373*(1755). https://doi.org/10.1098/rstb.2017.0345

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of "mudsplashes." *Nature*, *398*(6722), 34. https://doi.org/10.1038/17953

Otten, M., Pinto, Y., Paffen, C. L. E., Seth, A. K., & Kanai, R. (2017). The Uniformity Illusion: Central Stimuli Can Determine Peripheral Perception. *Psychological Science*, *28*(1), 56–68. https://doi.org/10.1177/0956797616672270

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, *162*(1-2), 8-13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12), 1136–1169. https://doi.org/10.1167/4.12.12

Petrov, Y., & Meleshkevich, O. (2011). Asymmetries and idiosyncratic hot spots in crowding. *Vision Research*, *51*(10), 1117–1123. https://doi.org/10.1016/j.visres.2011.03.001

Prettyman, A. (2018). Seeing the Forest and the Trees: A Response to the Identity Crowding Debate. *Thought: A Journal of Philosophy*, *7*(1), 20–30. https://doi.org/10.1002/tht3.265

Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition*, *107*(1), 82–104. https://doi.org/10.1016/j.cognition.2007.08.004

Richards, B. (2016). Attention and seeing objects: The identity-crowding debate. *Philosophical Psychology*, *29*(5), 743–758. https://doi.org/10.1080/09515089.2016.1150449

Rosenthal, D. (2019). Consciousness and confidence. *Neuropsychologia*, *128*, 255–265. https://doi.org/10.1016/j.neuropsychologia.2018.01.018

Rummens, K., & Sayim, B. (2022). Multidimensional feature interactions in visual crowding: When configural cues eliminate the polarity advantage. *Journal of Vision*, *22*(6), 2. https://doi.org/10.1167/jov.22.6.2

Rummens, K., & Sayim, B. (2021). Broad attention uncovers benefits of stimulus uniformity in visual crowding. *Scientific Reports*, *11*(1), 23976. https://doi.org/10.1038/s41598-021-03258-z

Rummens, K., & Sayim, B. (2019). Disrupting uniformity: Feature contrasts that reduce crowding interfere with peripheral word recognition. *Vision Research*, *161*, 25–35. https://doi.org/10.1016/j.visres.2019.05.006

Sayim, B., & Cavanagh, P. (2013). Grouping and crowding affect target appearance over different spatial scales. *PloS One*, *8*(8), e71188. https://doi.org/10.1371/journal.pone.0071188

Sayim, B., Greenwood, J. A., & Cavanagh, P. (2014). Foveal target repetitions reduce crowding. *Journal of Vision*, *14*(6), 4–4. https://doi.org/10.1167/14.6.4

Sayim, B., & Wagemans, J. (2017). Appearance changes and error characteristics in crowding revealed by drawings. *Journal of Vision*, *17*(11), 8. https://doi.org/10.1167/17.11.8

Sayim, B., & Taylor, H. (2019). Letters Lost: Capturing Appearance in Crowded Peripheral Vision Reveals a New Kind of Masking. *Psychological Science*, *30*(7), 1082–1086. https://doi.org/10.1177/0956797619847166

Scimeca, J. M., & Franconeri, S. L. (2015). Selecting and tracking multiple objects. *WIREs Cognitive Science*, *6*(2), 109–118. https://doi.org/10.1002/wcs.1328

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences*, *4*(4), 147–155. https://doi.org/10.1016/S1364-6613(00)01455-8

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Solovey, G., Graney, G. G., & Lau, H. (2015). A decisional account of subjective inflation of visual perception at the periphery. *Attention, Perception & Psychophysics*, *77*(1), 258–271. https://doi.org/10.3758/s13414-014-0769-1

Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, *49*(6), 495–508. https://doi.org/10.3758/bf03212183

Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, *5*(11), 1024–1037. https://doi.org/10.1167/5.11.8

Sun, G. J., Chung, S. T. L., & Tjan, B. S. (2010). Ideal observer analysis of crowding and the reduction of crowding through learning. *Journal of Vision*, *10*(5), 16. https://doi.org/10.1167/10.5.16

Taylor, H., & Sayim, B. (2018). Crowding, attention and consciousness: In support of the inference hypothesis. *Mind & Language*, *33*(1), 17–33. https://doi.org/10.1111/mila.12169

Taylor, H., & Sayim, B. (2020). Redundancy masking and the identity crowding debate. *Thought: A Journal of Philosophy*, *9*(4), 257–265. https://doi.org/10.1002/tht3.469

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*(7), 1349–1357. https://doi.org/10.1016/0042-6989(92)90227-a

Toscani, M., Mamassian, P., & Valsecchi, M. (2021). Underconfidence in peripheral vision. *Journal of Vision*, *21*(6), 2. https://doi.org/10.1167/jov.21.6.2

Vandenbroucke, A. R. E., Sligte, I. G., Barrett, A. B., Seth, A. K., Fahrenfort, J. J., & Lamme, V. A. F. (2014). Accurate metacognition for visual sensory memory representations. *Psychological Science*, *25*(4), 861–873. https://doi.org/10.1177/0956797613516146

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *PloS One*, *3*(9), e3269. https://doi.org/10.1371/journal.pone.0003269

Ward, E. J., Bear, A., & Scholl, B. J. (2016). Can you perceive ensembles without perceiving individuals?: The role of statistical perception in determining whether awareness overflows access. *Cognition*, *152*, 78–86. https://doi.org/10.1016/j.cognition.2016.01.010

Ward, E. J. (2018). Downgraded phenomenology: How conscious overflow lost its richness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1755), 20170355. https://doi.org/10.1098/rstb.2017.0355

Whitney, D., Haberman, J., & Sweeny, T. (2014). From textures to crowds: Multiple levels of summary statistical perception. *The New Visual Neurosciences*, 695–710.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology*, *69*, 105–129. https://doi.org/10.1146/annurev-psych-010416-044232

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision*, *10*(10), 16. https://doi.org/10.1167/10.10.16

Yildirim, F. Z., Coates, D. R., & Sayim, B. (2019). Lost lines in warped space: Evidence for spatial compression in crowded displays. *Journal of Vision*, *19*(10), 13c–13c. https://doi.org/10.1167/19.10.13c

Yildirim, F. Z., Coates, D. R., & Sayim, B. (2020). Redundancy masking: The loss of repeated items in crowded peripheral vision. *Journal of Vision*, *20*(4), 14. https://doi.org/10.1167/jov.20.4.14

Yildirim, F. Z., Coates, D. R., & Sayim, B. (2021). Hidden by bias: How standard psychophysical procedures conceal crucial aspects of peripheral visual appearance. *Scientific Reports*, *11*(1), 4095. https://doi.org/10.1038/s41598-021-83325-7

Yildirim, F. Z., Coates, D. R., & Sayim, B. (2022). Atypical visual field asymmetries in redundancy masking. *Journal of Vision*, *22*(5), 4. https://doi.org/10.1167/jov.22.5.4

Zhang, J.-Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2009). Legibility of Chinese characters in peripheral vision and the top-down influences on crowding. *Vision Research*, *49*(1), 44–53. https://doi.org/10.1016/j.visres.2008.09.021