



HAL
open science

High confidence and low accuracy in redundancy masking

Fazilet Zeynep Yildirim, Bilge Sayim

► **To cite this version:**

Fazilet Zeynep Yildirim, Bilge Sayim. High confidence and low accuracy in redundancy masking. *Consciousness and Cognition*, 2022, 102, pp.103349. 10.1016/j.concog.2022.103349 . hal-03904529

HAL Id: hal-03904529

<https://hal.univ-lille.fr/hal-03904529>

Submitted on 7 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 High confidence and low accuracy in redundancy masking

2 Fazilet Zeynep Yildirim^{1*}, Bilge Sayim^{1,2}

3 ¹Institute of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland

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

6 ***Corresponding author:** Fazilet Zeynep Yildirim

7 **Email:** fazilet.yildirim@psy.unibe.ch

8 **Address:** Fabrikstrasse 8, 3012, Bern, Switzerland

9 **Telephone:** +41 31 631 38 05

10

11 **ORCIDs**

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

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

14

15 **Keywords**

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

17

18 **Author Contributions**

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

21

22 **Competing interests**

23 The authors declare no competing interests.

24

25 **Abstract**

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

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

44

45

46

47 **Introduction**

48 Humans have a rich subjective impression of their visual environment. However, empirical evidence
49 indicates that visual abilities are strongly limited, suggesting that the impression of "richness" is illusory.
50 For example, phenomena such as change blindness (e.g., O'Regan et al., 1999; Simons & Rensink,
51 2005), inattention blindness (e.g., Simons, 2000), and the attentional blink (e.g., Dux & Marois, 2009)
52 indicate that salient information can easily go unnoticed. Furthermore, studies investigating the capacity
53 of visual attention and visual working memory have revealed that only a few items can be processed
54 and maintained at once (Luck & Vogel, 2013; Scimeca & Franconeri, 2015). Such findings raise the
55 question why our subjective impressions of the visual environment are richly detailed even though the
56 empirical evidence strongly suggests otherwise.

57 One view argues that we have rich impressions because conscious perception overflows the
58 capacities of cognitive mechanisms (e.g., attention, working memory, and decision-making) (the "rich
59 view": Block, 2011; Bronfman et al., 2014; Koch & Tsuchiya, 2007; Vandenbroucke et al., 2014). As a
60 result, according to the rich view, we are aware of more information than we can attend, remember, or
61 report. According to this view, there is no tension between the capacities of conscious perception and
62 the capacities of cognitive functions (such as attention and memory) because they are based on
63 separate mechanisms. By contrast, many researchers argue that without specific evidence, there
64 appears no good scientific reason to believe that consciousness and cognition are based on separate
65 mechanisms (Baars, 1989; Cohen et al., 2016; Cohen & Dennett, 2011; Dehaene, 2014; Kouider et al.,
66 2010; Lau & Rosenthal, 2011; Ward et al., 2016; Ward, 2018). For example, it was suggested that the
67 mechanisms of conscious perception and cognitive processes cannot be separated because awareness
68 is intrinsically linked to cognitive functions and information is not consciously perceived until it is
69 accessed by higher-order systems, such as attention, working memory, and decision-making (the
70 "sparse view": Baars, 1989; Cohen et al., 2016; Cohen & Dennett, 2011; Dehaene, 2014; Kouider et al.,
71 2010; Lau & Rosenthal, 2011; Ward et al., 2016; Ward, 2018). According to the sparse view, the
72 perceived level of detail is rich only at the focus of attention and becomes drastically limited (sparse)
73 outside of the focal point of attention. The tension between rich conscious perception and limited

74 capacities of cognitive functions remains in this case: why do we think we see more if our conscious
75 perception is limited by higher-level cognitive mechanisms?

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

93 Recent studies showed that the visual system does not only represent information as an ensemble
94 but also compresses redundant information substantially by masking individual items in repeating
95 patterns (i.e., 'redundancy masking'; Sayim & Taylor, 2019; Taylor & Sayim, 2020; Yildirim et al., 2019,
96 2020, 2021, 2022). For example, when presented with three identical items in the periphery, the majority
97 of observers report seeing only two items (Sayim & Taylor, 2019; Yildirim et al., 2021). We have shown
98 that redundancy masking increased with increasing similarity and regularity of the items, decreased with
99 increasing spacing between items, and was larger with radial compared to tangential item arrangements
100 (Yildirim et al., 2020). When redundancy masking occurs, visual space seems to be compressed:
101 Observers estimated the distance between the two perceived of three presented items as smaller than
102 the actual distance between the two outermost of the three presented items, and as larger than the
103 distance between two adjacent items (Yildirim et al., 2019). Taken together, these findings suggest that
104 the visual system summarizes information not only in the form of ensemble perception but also
105 compresses the number of identical items even when only three items are presented. Perceiving only
106 two of three items is such a significant loss of information about the actual stimulus that it makes
107 redundancy masking a promising tool to investigate questions regarding the often-reported richness of
108 visual consciousness. Here we investigated observers' metacognition in conditions where redundancy
109 masking is expected to be strong or to not occur.

110 Observers were presented with arrays of lines in the visual periphery and asked to rate their
111 confidence with displays usually yielding strong (radially arranged lines) and no (tangentially arranged
112 lines) redundancy masking. The number of lines presented was from three to seven, however, our main
113 focus was on three lines as the relative magnitude of redundancy masking is maximal in this condition
114 (i.e., often only two lines are reported, corresponding to missing $\frac{1}{3}$ of the presented lines). Consistent

115 with our previous study (Yildirim et al., 2020), we found that redundancy masking was strong with radially
116 arranged lines and absent with tangentially arranged lines. The analysis of confidence ratings showed
117 that in the radial condition, observers were more confident in their responses when redundancy masking
118 occurred (less lines reported than presented) compared to when no redundancy masking occurred
119 (correct responses). Hence, observers reported higher confidence for erroneous than for correct
120 judgments. In contrast, in the tangential condition, observers' confidence was similar in trials with and
121 without redundancy masking. The inversion of confidence in the radial condition (higher confidence
122 when accuracy was low and lower confidence when accuracy was high) suggests that redundancy-
123 masked appearance trumped 'veridical' perception. We suggest that the often-reported richness of
124 visual consciousness may partly be due to overconfidence in erroneous judgments in visual scenes that
125 are subject to redundancy masking.

126

127

128 **Materials and methods**

129 *Participants*

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

137 *Apparatus and stimuli*

138 Stimuli were generated with Psychopy v2.7.11 (Peirce, 2007) and displayed on a 22-in. CRT monitor
139 with a resolution of 1152 × 864 and a refresh rate of 110 Hz. The experiment was conducted in a dimly
140 illuminated room. Observers viewed the monitor from a distance of 57 cm and were supported by a chin
141 and head rest. A black (1 cd/m²) disc (diameter = 0.2°) was presented at the center of the screen for
142 fixation. Stimuli consisted of black (1 cd/m²) lines that were 1° in length and 0.04° in width, presented
143 on a uniform gray (42 cd/m²) background. The line arrays were centered at 10° eccentricity, and
144 presented to either the right or the left of the fixation disc. The number of presented lines ranged from
145 three to seven. The center-to-center spacing between adjacent lines within a line array was 0.85°. There
146 were two spatial arrangements of lines: in the radial condition, vertically oriented lines were horizontally
147 arranged (Figure 1, top left frame), and in the tangential condition, horizontally oriented lines were
148 vertically arranged (Figure 1, top right frame). The conditions are denoted as radial and tangential to
149 refer to how the lines are located relative to a concentric circle around fixation. The position of the line
150 array was slightly varied at random across trials (centered at 10° or jittered 4.7' either up, down, left, or
151 right). Responses were recorded using a number pad (enumeration task) and a computer mouse
152 (confidence rating task).

153

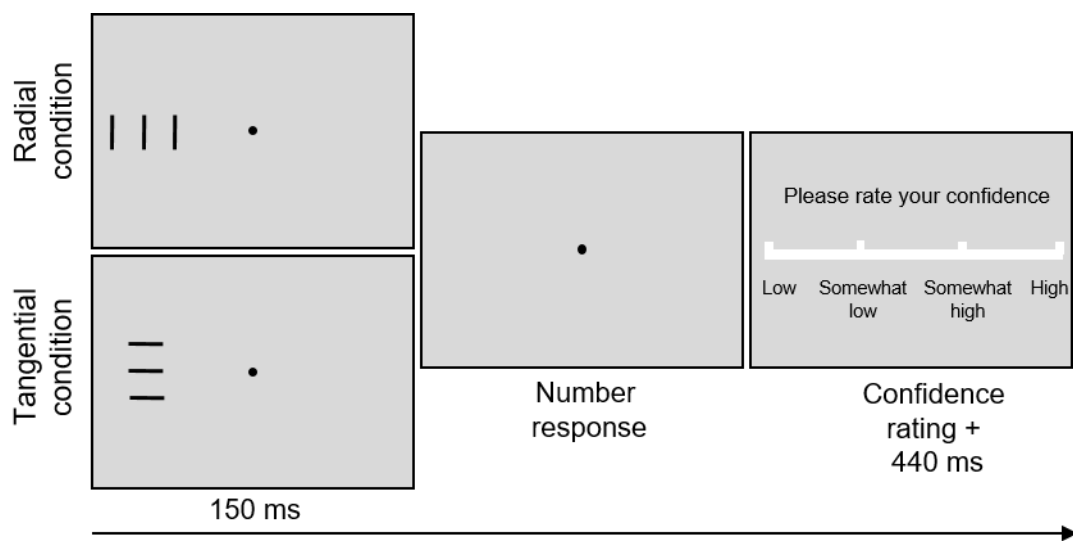
154 *Procedure*

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

162 The stimulus location (left or right of fixation) and the number of lines (three to seven) were
163 randomized and counterbalanced within each block. Presentations were blocked according to the spatial
164 arrangement of the lines (radial and tangential). A block consisted of 80 trials. Observers completed two
165 blocks with each spatial arrangement (a total of 320 trials). The sequence of radial and tangential blocks
166 was pseudorandomized for each observer. A schematic depiction of the procedure is shown in Figure
167 1.

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

175
176
177
178



179
180 **Figure 1.** Schematic of the experimental paradigm. Stimuli consisted of three to seven lines, and were
181 presented in separate blocks in radial or tangential arrangements randomly to the left or right side of
182 fixation. Observers first indicated the number of lines, and then rated their confidence on their response.
183 (The text in the last frame appeared in German in the experiment).
184

185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224

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 (Lüdtke, 2018). The pseudo R^2 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 ($R^2 = 0.43$).

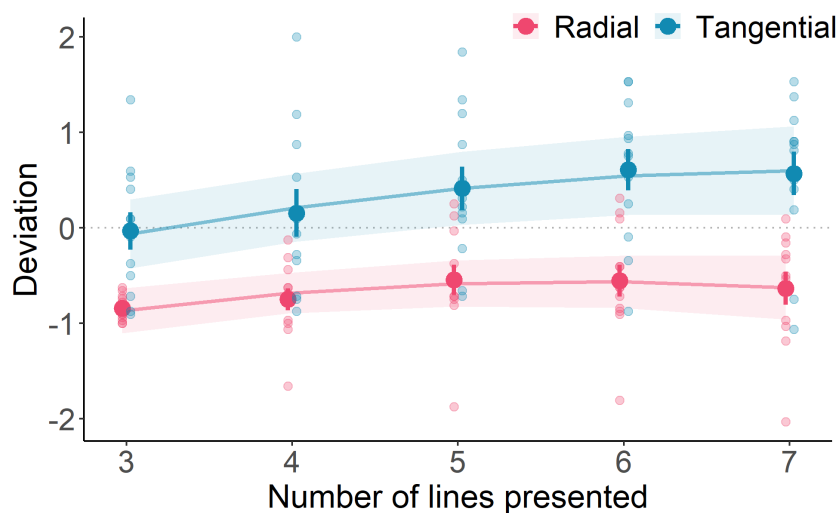
Confidence ratings: Mean confidence ratings were analyzed by a generalized linear mixed-effects model specifying the number of lines presented, and spatial arrangement as fixed effects and subject as a random factor using the glmmTMB package (Brooks et al., 2017). Model selection and checking of the 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.

225 **Results**

226 Mean deviation scores as a function of the number of lines presented for the radial and tangential
227 conditions are shown in Figure 2. As expected, we found strong redundancy masking in the radial and
228 no redundancy masking in the tangential condition. Both main effects of spatial arrangement ($\chi^2(1) =$
229 $21.49, p = 0.0000036$) and the number of lines presented ($\chi^2(2) = 12.55, p = 0.0019$), and the interaction
230 ($\chi^2(2) = 8.53, p = 0.014$) between them were significant. Mean deviation scores were clearly below zero
231 in the radial condition ($-0.66 \pm SD 0.42$; strong redundancy masking), and they were above zero in the
232 tangential condition ($0.34 \pm SD 0.69$; the opposite of redundancy masking, i.e., overestimation). The
233 deviation scores varied with the number of lines presented in a quadratic (inverted-U) manner in the
234 both conditions: an initial positive slope of the deviation scores for small numbers of lines was followed
235 by a negative slope with larger numbers of lines (the quadratic fit outperformed linear fits). In the radial
236 condition, the deviation scores were all below zero, and in the tangential condition, they were close to
237 zero with three lines, and above zero for all other numbers of lines. Descriptive statistics of deviation
238 scores for each number of lines and spatial arrangement condition are shown in the Supplementary
239 Table 1a.



240

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

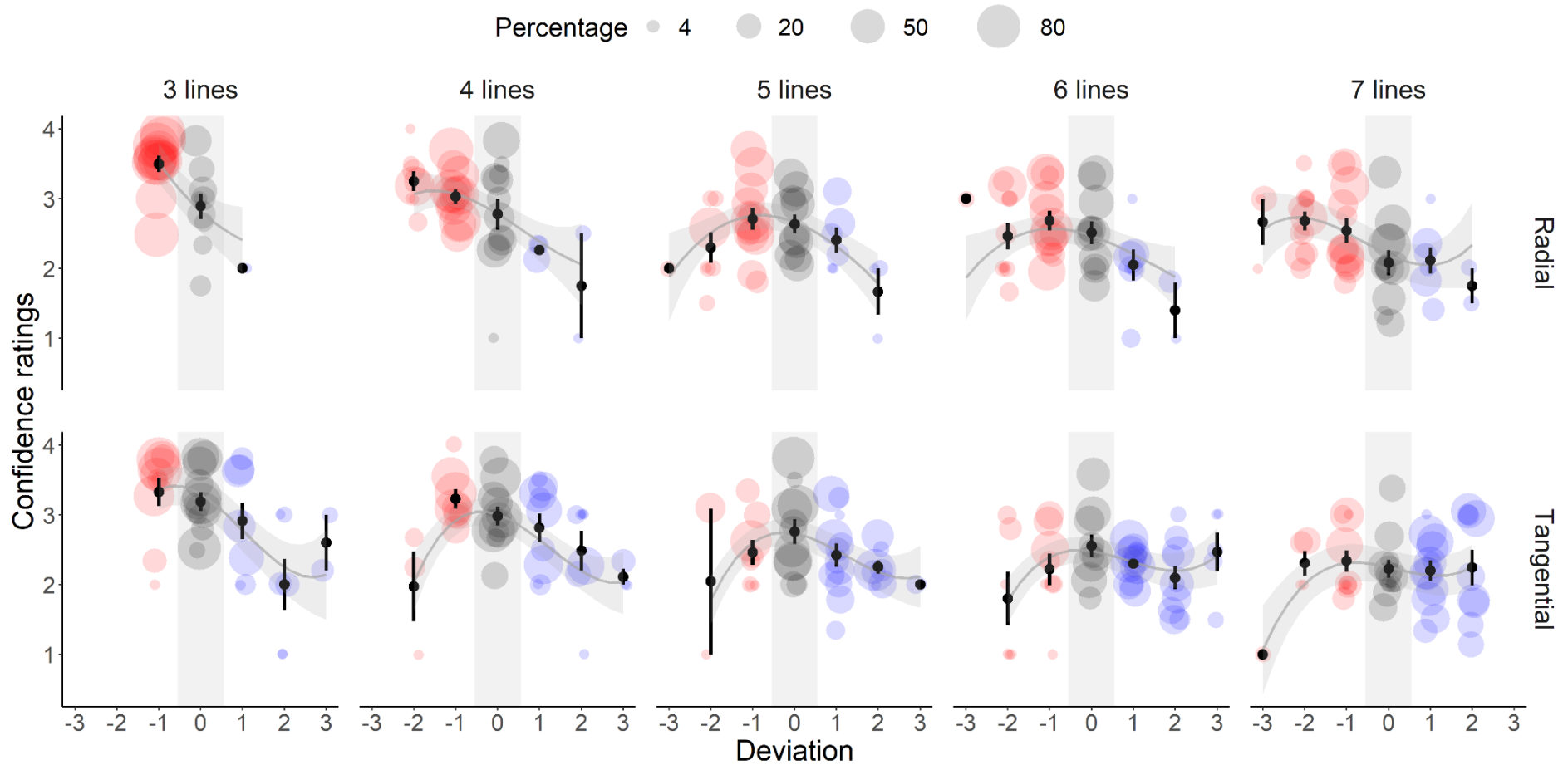
246

247
248 Confidence ratings as a function of deviation scores for the radial and tangential conditions are
249 shown in Figure 3. Our main focus was the condition with three lines as redundancy masking is most
250 pronounced in this condition. The result of the polynomial regression analysis suggested that observers
251 were more confident in the redundancy masked report (deviation score -1; two lines reported) compared
252 to the correct report (deviation score 0; three lines reported) in the radial condition (Fig. 3, three lines).
253 In comparison, they were equally confident in both reports in the tangential condition (Fig. 3, three lines).
254 An analysis of the local slopes in the polynomial regressions confirmed this finding. The slope between
255 the deviation score of -1 (two lines) and 0 (three lines) in the radial condition was significantly lower than
256 zero ($\beta = -0.61, 95\% CI [-0.84, -0.38], p < .0001$) whereas the corresponding slope in the tangential

257 condition did not differ from zero ($\beta = 0.008$, 95% CI [-0.25, 0.27], $p = .953$). Importantly, the slopes
258 between the deviation scores of -1 (two lines) and 0 (three lines) differed between the radial and the
259 tangential condition ($t(380) = -3.45$, $p = 0.003$): Confidence dropped significantly from two lines to three
260 lines in the radial condition (Fig. 3; three lines); in contrast, it did not change between two and three
261 lines in the tangential condition (Fig. 3, three lines). Descriptive statistics of confidence ratings for each
262 number of lines and spatial arrangement condition are shown in the Supplementary Table 1b.

263

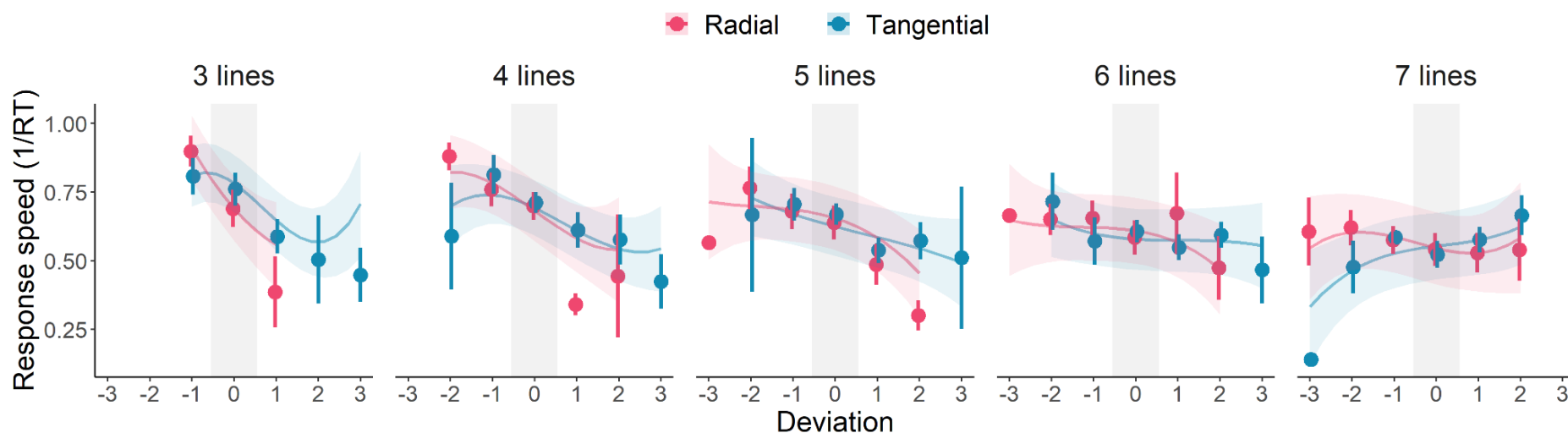
264



265
 266
 267
 268
 269
 270
 271
 272

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 ($\pm SEM$) 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 ($\pm 1.96 * SEM$), respectively.

273 We also assessed the response speeds in the radial and tangential conditions to explore whether they followed the same pattern as the confidence ratings.
 274 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
 275 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
 276 lines) in the radial condition was significantly lower than zero ($\beta = -0.21$, 95% CI [-0.30, -0.12], $p < .0001$) whereas the corresponding slope in the tangential
 277 condition did not differ from zero ($\beta = -0.03$, 95% CI [-0.13, 0.07], $p = .524$). Importantly, the slopes between the deviation scores of -1 (two lines) and 0 (three
 278 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
 279 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).
 280 Systematic differences based on key locations or fingers used may have contributed to the different response speeds for indicating two and three lines. However,
 281 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
 282 the response speeds for each number of lines and spatial arrangement condition are shown in the Supplementary Table 1c.
 283



284
 285 **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
 286 points with error bars ($\pm SEM$) show the mean response speed for the radial and tangential conditions, respectively. The pink and blue lines show the predicted
 287 values from the mixed model for the radial and tangential conditions, respectively. Shaded regions represent confidence intervals for the predicted values based
 288 on the standard errors ($\pm 1.96 * SEM$).

289 Discussion

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

307 Unlike previous studies in which observers' metacognition was compared across conditions that
308 differed strongly, such as foveal versus peripheral vision (Solovey et al., 2015), or crowded versus
309 uncrowded stimuli (e.g., Odegaard et al., 2018), we aimed to investigate metacognition with identical
310 stimuli and presentation conditions. Comparisons between foveal and peripheral presentation are
311 challenging because many factors may cause differences between the conditions. In particular, different
312 criteria to evaluate accuracy of one's performance may produce differences between foveal and
313 peripheral presentation, even when performance in the two conditions is matched. For example, a recent
314 study showed that observers' metacognitive sensitivity was lower when comparing perceptual decisions
315 of stimuli presented at different eccentricities (central and peripheral) compared to the same eccentricity
316 (central with central, and peripheral with peripheral) (Toscani et al., 2021). Comparing crowded with
317 uncrowded stimuli is limited by similar constraints: Differences in stimulus complexity (Zhang, Zhang,
318 Xue, Liu, & Yu, 2009; Sayim & Wagemans, 2017), attentional demands (Strasburger, 2005; He,
319 Intriligator, & Cavanagh, 1996; Rummens & Sayim, 2021), internal and external noise (Sun, Chung, &
320 Tjan, 2010) render it difficult to extract the variable(s) that underlie variations of confidence. Here, we
321 used a redundancy masking paradigm to investigate confidence when appearance varied with physically
322 identical stimuli under identical presentation conditions. Thus, any differences between confidence
323 judgments were not due to variations of the stimuli or presentation conditions, but reflect changes in
324 stimulus *appearance* independent of these confounding factors. Hence, our results complement
325 previous studies that investigated confidence in peripheral vision (Baldassi et al., 2006; Odegaard et al.,
326 2018; Otten et al., 2017; Rosenthal, 2019; Solovey et al., 2015). For example, Odegaard et al. (2018)
327 found higher confidence ratings for erroneous responses when the target was crowded compared to
328 when it was not crowded. This finding was taken as an indication that the rich visual experience is due
329 to 'subjective inflation' in the periphery. However, as their conclusions are based on comparisons

330 between different stimulus conditions (i.e., crowded vs uncrowded), other factors related to these
331 conditions (e.g., stimulus complexity, attentional demands) might confound their results. Here, we avoid
332 these confounding factors by comparing confidence ratings of erroneous (redundancy-masked)
333 responses and correct responses of the same stimulus. By comparing confidence ratings within the
334 same stimulus, we show that confidence followed the ‘appearance’ of a stimulus, not performance.

335 Our main focus was the condition with three lines as the relative magnitude of redundancy masking
336 is greatest when the number of items is small (Yildirim et al., 2020; 2021; see also the relative magnitude
337 of errors in Supplementary Figure S1). In agreement with previous studies (Rummens & Sayim, 2022;
338 Sayim & Taylor, 2019; Yildirim et al., 2020, 2021, 2022), there was strong redundancy masking in the
339 radial condition: when three lines were presented, participants reported two lines in 80% of the trials,
340 with high confidence (on average 3.5 ± 0.39 on a scale from 1 to 4). In comparison, participants reported
341 the correct number, three lines, in only 18% of the trials with lower confidence (2.9 ± 0.57). Contrary to
342 findings of most studies investigating confidence in perceptual judgements, observers were more
343 confident in their incorrect responses than their correct responses. Only two observers reported four
344 lines (1.2% of the trials with three lines), and their confidence was low (2 ± 0.00). In the tangential
345 condition, by contrast, observers reported two lines in 32%, three in 47%, four in 16%, and five or six in
346 5% of the trials with three lines. Importantly, confidence did not differ between “two” (3.3 ± 0.64) and
347 “three” (3.2 ± 0.46) responses in the tangential condition (and confidence was lower when observers
348 reported four (2.9 ± 0.73), five (2.0 ± 0.89), and six (2.6 ± 0.57) lines). Hence, confidence in redundancy-
349 masked reports was relatively high with radial, but not with tangential arrangements.

350 Enumeration of small quantities (1 to ~3) is typically fast and accurate (i.e., subitizing) whereas
351 enumeration of larger quantities (~4 and larger) is slow and error-prone (i.e., estimation) (Kaufman et
352 al., 1949; Mandler & Shebo, 1982). Subitizing performance in peripheral vision is similar as with
353 free/central viewing for attended and widely spaced items, but has been shown to decline under high
354 attentional load (Railo et al., 2008) and close inter-item spacings (Chakravarthi & Herbert, 2019). We
355 suggest that in the current study observers subitized, albeit mostly erroneously, the number of lines
356 when presented with three lines, resulting in reports of either two or three lines in most trials. In contrast,
357 when presented with more lines (~4 to 7), subitizing was no longer possible, and observers had to
358 estimate the number of lines, which led to a wider range of deviation scores (ranging mostly between -
359 2 to +2 around the presented number of lines) (see also Yildirim et al., 2021).

360 In the radial condition, redundancy masking was not only strong with three lines, but also with larger
361 numbers of lines. As expected, confidence ratings decreased with the number of lines presented,
362 showing the typical increase of uncertainty about response accuracy when estimating the number of
363 items in large sets (see Supplementary Figure S2) (Kaufman et al., 1949; Railo et al., 2008). In line with
364 the lower confidence ratings for larger numbers of items, the variance of responses increased (ranging
365 mostly between -2 to +1 around the presented number) and the response speed decreased. However,
366 while there was only a small decrease of the magnitude of (negative) deviation scores with the increase
367 of the number of lines (from three to five lines; Figure 2 and Supplementary Table 1a), relative deviation
368 scores decreased more strongly with increasing numbers of lines (Supplementary Figure S1), showing
369 how redundancy masking is most pronounced with smaller numbers of items. Hence, in the radial
370 condition, the strongest relative deviation went hand in hand with the highest confidence: Observers

371 were most confident when the relative deviation from the correct response was largest. In the tangential
372 condition, by contrast, there was neither redundancy masking for small nor for large numbers of lines.
373 While responses were on average accurate for three lines, (positive) deviation scores increased with
374 the number of lines, showing that observers overestimated the number when large numbers of lines
375 were presented. The observed decrease of confidence ratings (and response speed), and the increase
376 of variance with increasing numbers of lines were - as in the radial condition - unsurprising. Similarly,
377 overestimation with tangentially arranged lines has been reported previously (Yildirim et al., 2020).

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

403 Higher confidence in redundancy-masked stimuli may well play a role in the seemingly rich
404 representation of the visual environment. Given the usually strong link between high confidence and
405 good performance (e.g., Barthelmé & Mamassian, 2010), poor performance yielding high confidence
406 remains the exception to the rule. Redundancy masking seems to be an example that inverts this
407 relationship. Importantly, in contrast to previous studies that found high confidence when performance
408 was poor (i.e., inflation) based on comparisons with other stimuli (e.g., uncrowded versus crowded,
409 Odegaard et al., 2018) or presentation conditions (e.g., fovea versus periphery, Solovey et al., 2015),
410 here we directly compared confidence judgments on the very same stimuli and under the same
411 presentation conditions. To evaluate whether redundancy masking is a case of inflation, comparisons

412 can be made between redundancy-masked and correct trials, and between smaller and larger numbers
413 of lines. We already outlined the difference in confidence for redundancy-masked and correct trials (in
414 particular, in the radial condition with three lines). The high confidence in reporting two lines when
415 presented with three lines is indicative of inflation: Observers were confident that their erroneous
416 responses were correct. However, confidence was lower for correct than for incorrect responses. This
417 reversal of the typical relationship between confidence and performance shows that observers were
418 largely incapable of accurately evaluating their performance. Importantly, confidence was relatively high
419 for both redundancy-masked (3.5 ± 0.39 , for the three radial lines) and correct trials (2.9 ± 0.57 , for the
420 three radial lines) - two categorically different responses for one and the same stimulus, suggesting
421 again that observers' confidence judgments were inflated. Confidence was overall higher with smaller
422 compared to larger numbers of lines (see also Supplementary Figure S2), and deviation scores indicated
423 only slightly stronger redundancy masking for three lines compared to the other numbers of lines in the
424 radial condition (and stronger overestimation with increasing numbers of lines in the tangential
425 condition). This was expected, and it alone does not show inflation. However, due to redundancy
426 masking, the *relative* error was much higher for three lines than larger numbers of lines (five, six, and
427 seven; again in the radial condition) (see Supplementary Figure S1). The high confidence ratings for
428 three lines in conjunction with the highest relative deviation error shows that observers' confidence
429 judgments were inflated.

430 Our results suggest that redundancy-masked trials are a good representation of how a stimulus
431 appeared to observers. Hence, redundancy masking may contribute to the impression of a rich visual
432 world by creating a convincing illusion that what is perceived is accurately capturing what is present in
433 the stimulus. Importantly, redundancy masking would only contribute to the impression of a rich visual
434 world when the 'erroneous' nature of the percept goes unnoticed, as high confidence in the accuracy of
435 the percept is required. While future studies still need to investigate to what extent the occurrence of
436 redundancy masking is not detected when attending to a stimulus in the periphery before fixating it,
437 there is some evidence that it is not easily detected. Despite many decades of investigations of visual
438 crowding (Bernard & Chung, 2011; Bouma, 1970, 1973; Flom et al., 1963; Herzog et al., 2015; Korte,
439 1923; Levi et al., 1985, 2002; Manassi et al., 2012; Melnik et al., 2018, 2020; Pelli et al., 2004; Rummens
440 & Sayim, 2019, 2021; Sayim & Cavanagh, 2013, Sayim et al., 2014, Sayim & Wagemans, 2017;
441 Strasburger et al., 1991; Yeshurun & Rashal, 2010), where usually three more or less similar items are
442 presented (a target and two flankers), redundancy masking was only discovered recently (Sayim &
443 Taylor, 2019; Taylor & Sayim, 2018, 2020; Yildirim et al., 2019, 2020, 2021, 2022). Interestingly, a
444 philosophical debate investigating the phenomenon of 'identity crowding' where the target and the
445 flankers are the same (similar to three radial lines in the present study), has been focused on the
446 question why performance on the central of the three identical items is so unexpectedly *good* (Block,
447 2012; 2013; Prettyman, 2018; Richards, 2016; Taylor & Sayim, 2018, 2020), and not unexpectedly *poor*
448 as shown in redundancy masking paradigms. The seemingly good performance has been proposed to
449 be evidence for the capacity to see without attention (Block 2012, 2013). Evidence from redundancy
450 masking experiments suggests that the premises of this debate, based to a large part on purely
451 phenomenological approaches, were erroneous, supporting the proposal that redundancy masking
452 easily goes unnoticed. Taken together, we showed high confidence for redundancy-masked stimuli,

453 suggesting that a compressed, non-veridical representation of the actual stimulus better represents
454 stimulus appearance. The impression of a rich visual world may partly be driven by the inability to notice
455 the loss of information in visual scenes susceptible to redundancy masking, and high confidence that
456 one's phenomenology is an accurate representation of the observed scene.

457

458 **Data and code availability**

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

461 **Acknowledgment**

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

463 **Funding**

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

466

467

468

469

470

471

472

473

474

475

476

477

478

479

481 **References**

- 482 Alam, S., Luccio, R., & Vardabasso, F. (1986). Regularity, Exposure Time and Perception of
 483 Numerosity. *Perceptual and Motor Skills*, 63(2), 883–888.
 484 <https://doi.org/10.2466/pms.1986.63.2.883>
- 485 Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition.
 486 *Trends in Cognitive Sciences*, 15(3), 122–131. <https://doi.org/10.1016/j.tics.2011.01.003>
- 487 Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, 12(2),
 488 157–162. <https://doi.org/10.1111/1467-9280.00327>
- 489 Baars, B. J. (1989). *A cognitive theory of consciousness*. Cambridge University Press.
- 490 Baldassi, S., Megna, N., & Burr, D. C. (2006). Visual clutter causes high-magnitude errors. *PLoS*
 491 *Biology*, 4(3), e56. <https://doi.org/10.1371/journal.pbio.0040056>
- 492 Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual
 493 judgments. *Perception & Psychophysics*, 55(4), 412–428. <https://doi.org/10.3758/bf03205299>
- 494 Barthelmé, S., & Mamassian, P. (2010). Flexible mechanisms underlie the evaluation of visual
 495 confidence. *Proceedings of the National Academy of Sciences of the United States of*
 496 *America*, 107(48), 20834–20839. <https://doi.org/10.1073/pnas.1007704107>
- 497 Barton, K., & Barton, M. K. (2015). *Package ‘MuMIn’*. Version, 1, 18.
- 498 Bernard, J.-B., & Chung, S. T. L. (2011). The dependence of crowding on flanker complexity and
 499 target-flanker similarity. *Journal of Vision*, 11(8). <https://doi.org/10.1167/11.8.1>
- 500 Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*,
 501 15(12), 567–575. <https://doi.org/10.1016/j.tics.2011.11.001>
- 502 Block, N. (2012). The Grain of Vision and the Grain of Attention. *Thought: A Journal of Philosophy*,
 503 1(3), 170–184. <https://doi.org/10.1002/tht3.28>
- 504 Block, N. (2013). Seeing and Windows of Integration. *Thought: A Journal of Philosophy*, 2(1), 29–39.
 505 <https://doi.org/10.1002/tht3.62>
- 506 Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226(5241), 177–178.
 507 <https://doi.org/10.1038/226177a0>
- 508 Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words.
 509 *Vision Research*, 13(4), 767–782. [https://doi.org/10.1016/0042-6989\(73\)90041-2](https://doi.org/10.1016/0042-6989(73)90041-2)
- 510 Bronfman, Z. Z., Brezis, N., Jacobson, H., & Usher, M. (2014). We see more than we can report: “cost
 511 free” color phenomenality outside focal attention. *Psychological Science*, 25(7), 1394–1403.
 512 <https://doi.org/10.1177/0956797614532656>

- 513 Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... &
514 Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-
515 inflated generalized linear mixed modeling. *The R journal*, 9(2), 378-
516 400. <https://doi.org/10.3929/ethz-b-000240890>
- 517 Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires
518 attentional resources. *Journal of Vision*, 10(6), 20–20. <https://doi.org/10.1167/10.6.20>
- 519 Chakravarthi, R., & Herbert, A. (2019). Two's company, three's a crowd: Individuation is necessary for
520 object recognition. *Cognition*, 184, 69–82. <https://doi.org/10.1016/j.cognition.2018.12.008>
- 521 Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in*
522 *Cognitive Sciences*, 15(8), 358–364. <https://doi.org/10.1016/j.tics.2011.06.008>
- 523 Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the Bandwidth of Perceptual
524 Experience? *Trends in Cognitive Sciences*, 20(5), 324–335.
525 <https://doi.org/10.1016/j.tics.2016.03.006>
- 526 Dehaene, S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*.
527 Penguin.
- 528 Dux, P. E., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention,*
529 *Perception, & Psychophysics*, 71(8), 1683–1700. <https://doi.org/10.3758/APP.71.8.1683>
- 530 Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual Resolution and Contour Interaction.
531 *Journal of the Optical Society of America*, 53, 1026–1032.
532 <https://doi.org/10.1364/josa.53.001026>
- 533 Ginsburg, N. (1976). Effect of item arrangement on perceived numerosity: Randomness vs regularity.
534 *Perceptual and Motor Skills*, 42(43), 663–668. <https://doi.org/10.2466/pms.1976.43.2.663>
- 535 Ginsburg, N. (1978). Perceived numerosity, item arrangement, and expectancy. *The American Journal*
536 *of Psychology*, 91(2), 267–273.
- 537 Ginsburg, N. (1980). The regular-random numerosity illusion: Rectangular patterns. *The Journal of*
538 *General Psychology*, 103(2d Half), 211–216. <https://doi.org/10.1080/00221309.1980.9921000>
- 539 Greenwood, J. A., Szinte, M., Sayim, B., & Cavanagh, P. (2017). Variations in crowding, saccadic
540 precision, and spatial localization reveal the shared topology of spatial vision. *Proceedings of*
541 *the National Academy of Sciences of the United States of America*, 114(17), E3573–E3582.
542 <https://doi.org/10.1073/pnas.1615504114>
- 543 Hartig, F. (2017). *DHARMA: Residual diagnostics for hierarchical (multi-level/mixed) regression*
544 *models*. (R package version 0.1) [Computer software].
- 545 He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness.
546 *Nature*, 383(6598), 334–337. <https://doi.org/10.1038/383334a0>
- 547 Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object
548 recognition: A matter of appearance. *Journal of Vision*, 15(6), 5. <https://doi.org/10.1167/15.6.5>

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

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

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

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

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

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

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

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

568 Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical
569 magnification. *Vision Research*, 25(7), 963–977. [https://doi.org/10.1016/0042-](https://doi.org/10.1016/0042-6989(85)90207-x)
570 [6989\(85\)90207-x](https://doi.org/10.1016/0042-6989(85)90207-x)

571 Levin, D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change Blindness Blindness: The
572 Metacognitive Error of Overestimating Change-detection Ability. *Visual Cognition*, 7(1–3),
573 397–412. <https://doi.org/10.1080/135062800394865>

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

- 617 Rosenthal, D. (2019). Consciousness and confidence. *Neuropsychologia*, 128, 255–265.
618 <https://doi.org/10.1016/j.neuropsychologia.2018.01.018>
- 619 Rummens, K., & Sayim, B. (2022). Multidimensional feature interactions in visual crowding: When
620 configural cues eliminate the polarity advantage. *Journal of Vision*, 22(6), 2.
621 <https://doi.org/10.1167/jov.22.6.2>
- 622 Rummens, K., & Sayim, B. (2021). Broad attention uncovers benefits of stimulus uniformity in visual
623 crowding. *Scientific Reports*, 11(1), 23976. <https://doi.org/10.1038/s41598-021-03258-z>
- 624 Rummens, K., & Sayim, B. (2019). Disrupting uniformity: Feature contrasts that reduce crowding
625 interfere with peripheral word recognition. *Vision Research*, 161, 25–35.
626 <https://doi.org/10.1016/j.visres.2019.05.006>
- 627 Sayim, B., & Cavanagh, P. (2013). Grouping and crowding affect target appearance over different
628 spatial scales. *PloS One*, 8(8), e71188. <https://doi.org/10.1371/journal.pone.0071188>
- 629 Sayim, B., Greenwood, J. A., & Cavanagh, P. (2014). Foveal target repetitions reduce crowding.
630 *Journal of Vision*, 14(6), 4–4. <https://doi.org/10.1167/14.6.4>
- 631 Sayim, B., & Wagemans, J. (2017). Appearance changes and error characteristics in crowding
632 revealed by drawings. *Journal of Vision*, 17(11), 8. <https://doi.org/10.1167/17.11.8>
- 633 Sayim, B., & Taylor, H. (2019). Letters Lost: Capturing Appearance in Crowded Peripheral Vision
634 Reveals a New Kind of Masking. *Psychological Science*, 30(7), 1082–1086.
635 <https://doi.org/10.1177/0956797619847166>
- 636 Scimeca, J. M., & Franconeri, S. L. (2015). Selecting and tracking multiple objects. *WIREs Cognitive
637 Science*, 6(2), 109–118. <https://doi.org/10.1002/wcs.1328>
- 638 Simons, D. J. (2000). Attentional capture and inattention blindness. *Trends in Cognitive Sciences*,
639 4(4), 147–155. [https://doi.org/10.1016/S1364-6613\(00\)01455-8](https://doi.org/10.1016/S1364-6613(00)01455-8)
- 640 Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in
641 Cognitive Sciences*, 9(1), 16–20. <https://doi.org/10.1016/j.tics.2004.11.006>
- 642 Solovey, G., Graney, G. G., & Lau, H. (2015). A decisional account of subjective inflation of visual
643 perception at the periphery. *Attention, Perception & Psychophysics*, 77(1), 258–271.
644 <https://doi.org/10.3758/s13414-014-0769-1>
- 645 Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric
646 characters in direct and eccentric view. *Perception & Psychophysics*, 49(6), 495–508.
647 <https://doi.org/10.3758/bf03212183>
- 648 Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision.
649 *Journal of Vision*, 5(11), 1024–1037. <https://doi.org/10.1167/5.11.8>

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

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

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

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

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

660 Vandenbroucke, A. R. E., Sligte, I. G., Barrett, A. B., Seth, A. K., Fahrenfort, J. J., & Lamme, V. A. F.
661 (2014). Accurate metacognition for visual sensory memory representations. *Psychological*
662 *Science*, 25(4), 861–873. <https://doi.org/10.1177/0956797613516146>

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

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

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

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

674 Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology*, 69,
675 105–129. <https://doi.org/10.1146/annurev-psych-010416-044232>

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

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

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

683 Yildirim, F. Z., Coates, D. R., & Sayim, B. (2021). Hidden by bias: How standard psychophysical
684 procedures conceal crucial aspects of peripheral visual appearance. *Scientific Reports*, 11(1),
685 4095. <https://doi.org/10.1038/s41598-021-83325-7>

686 Yildirim, F. Z., Coates, D. R., & Sayim, B. (2022). Atypical visual field asymmetries in redundancy
687 masking. *Journal of Vision*, 22(5), 4. <https://doi.org/10.1167/jov.22.5.4>
688 Zhang, J.-Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2009). Legibility of Chinese characters in peripheral
689 vision and the top-down influences on crowding. *Vision Research*, 49(1), 44–53.
690 <https://doi.org/10.1016/j.visres.2008.09.021>
691